We work better together: using an ensemble of natural language processing and forced choice to measure conscientiousness

by

Xavier Heron

B.S., Iowa State University, 2020

A THESIS

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Psychological Sciences
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Dr. Tianjun Sun

Copyright

Abstract

The primary aim of this thesis was to implement a research design which collected forced choice data and text data simultaneously, then examine the psychometric properties of an *ensemble* model generated from both the forced choice data the text data. Forced choice (FC) and language models offer psychometric advantages, as well as advantages in selection settings when compared to Likert scales – the traditional method of measuring personality traits (Ozer & Benet-Martinez, 2006). The present research provides psychometric evidence of the validity of a multi-method ensemble approach to measuring conscientiousness.

Over 45 trials, participants selected one of two statements in a FC block and then provided a short text explanation of their choice. The FC data was scored using the Generalized Thurstonian Unfolding Model (Zhang et al., 2023). The text data was embedded using pre-trained BERT-base-uncased (Devlin et al., 2018) and trained to predict conscientiousness scores generated from the Chernyshenko Conscientiousness Scale (CCS; Chernyshenko, 2002) using ridge regression. FC scores and Natural Language Processing (NLP) scores were ensembled using linear regression to predict scores generated from the CCS.

Construct validity was evaluated using a Multi-Trait Multi-Method table comparing the CCS scores, the NLP-derived scores, the forced choice scores, and the ensemble model scores. Incremental validity of the ensemble model above and beyond the other models was evaluated through two hierarchical regressions predicting GPA and subjective well-being.

The ensemble model yielded good criterion-related validity and convergent validity, acceptable discriminant validity, but poor reliability – likely due to unstable forced choice estimates. The ensemble model did not demonstrate incremental validity above and beyond the other estimates.

Table of Contents

List of Figures

List of Tables

Acknowledgements

Thank you to my wonderful wife, Lizzie, my sweet boy, Henry (3), and my beautiful baby girl, Zoey (3 months). Everything is always family and God. These last two years have been an adventure and doing it with my amazing wife and family has been a blessing and joy that I don't deserve and am grateful for. The best is yet to come.

Thank you to my advisor, Dr. Tianjun Sun, perhaps the smartest person I've ever met (and I have an ego and own a mirror). If not for her patience for my endless questions and grace for all my missed deadlines, I would not have the opportunity to do what I love. The best is yet to come.

Thank you to my committee, Dr. Jin Lee and Dr. Rachael Xiong. I couldn't have asked for (or probably found) a smarter, or kinder, group of scholars who I am certain will accomplish whatever they want. The best is yet to come.

Dedication

To Lizzie, Henry, and Zoey.

Preface

This line of research is about methods and measurement applied to personality. The personality trait *conscientiousness,* which include the facets industriousness, orderliness, self-control, responsibility, virtue, and traditionalism, was chosen for this research because of its relationship with positive work outcomes such as higher job satisfaction, lower turnover intentions, and higher subjective well-being (Roberts et al., 2007). The focus of this research and thesis will be the methods and measurement, not the construct itself.

Personality

Everyone is a unique and stunningly bright personality,

but some of us have been trained to hide.

When we hide from the world, the world gets a little bit darker.

Worse, we unintentionally train those around us to also hide.

But when we ignore the training and trainers and stop hiding,

we unconsciously give those around us permission to also stop hiding.

And the dark starts to look a little brighter.

And the cold starts to feel a little warmer.

And the world starts to feel a little less lonely.

****while these are my words, much is drawn from a 1972 Marianne Williamson quote discussing what she referred to as 'Our Deepest Fear'****

Chapter 1 - Introduction

Personality and Models

Personality researchers want to measure and model personality. To do so, there are three fundamental assumptions to this research, which will be discussed in turn. First, everyone's personality is unique. Second, everything is a model. Third, "better" models are more accurate representations of what they are attempting to model.

Everyone Has a Unique Personality

The question of the role of nature and nurture in who we are has been a debate in psychology for a long time (McCrae et al., 2000). The world we live in is responsible, at least in part, for who we are. The number of environmental factors and the combinations of those factors is infinite. Where we were born, when we were born, who we were born to, how we're born, and what we're born with all interact in unique combinations and magnitudes to contribute to our personality.

Location

Where we are born matters – there are cross-cultural [observable] differences in personality (McCrae, 2002). Even within a country there are regional differences, which are not as easily observed but likely present. While not while studied in psychological research, in the United States, "the South" is often seen as a distinct culture, or at least as having distinct cultural aspects (Cox, 2009). The people we're around, the norms of the community we are a part of, and the policies of the governments we empower are all factors in the environment. That environment, the one we're born into and raised in, but didn't choose, has an effect on who we are, which, sometimes, is large enough to be observable, even with our current measurement techniques. These environments are different for everyone, even if only subtle ways.

Sidebar - Philosophy of Effect Sizes

The nuanced environmental differences may or may not produce observable differences, but that can certainly be ascribed to the sensitivity of the measurement tools. Detecting an effect requires the measurement tool be sensitive enough. Easily forgotten by researchers is the fact that "insignificant" effects are, in fact, simply, effects that are not detectable by the current measurement instruments and *not* nonexistent effects. An unfortunate side effect of null hypothesis testing is that many researchers assume the null to be true until proven false, which is a categorically false belief. An effective test for identifying a dogmatic researcher is whether they endorse the idea that a psychological phenomenon can be truly orthogonal to another psychological phenomenon. Any relationship between anything that anyone thinks of has an effect size; whether it's statistically significant, practically significant, or poorly formed is another matter.

Time-Period

When we were born matters. The time in history we're born into plays a role in our personality. For example, there's evidence that cell phone use plays a role in personality (Roberts et al., 2015). Cell phone usage is a unique factor to people alive today that was not a factor for most of history. Each time period will have its own unique factors. For example, after WWII, a large proportion of soldiers coming home with PTSD had an effect on the society at large, an effect which is difficult to quantify (Jones & Wessely, 2006). The unique things that happen to different generations will have an effect on personality.

Genetics

What we're born with, our genetics, while not well understood, play a role in our personality (Penke et al., 2007). At the very least, men and women get treated differently in the

workplace (Heilman, 2012) and the response of the environment to the individual is going to have an effect on who they are. Even more subtle things, such as height and hair color, are also going change how the environment reacts to the individual and have a, usually, unobservable effect on who we are.

Parents

Who we're born to, who our parents are, also has an effect on personality. For example, socio-economic status plays a role in our personality (Ayoub et al., 2018). Our parents have a role in our personality (Harris, 2000). The advantages parents provide for their children and disadvantages they impose on their children are so strong that some of these differences can be observed.

Starts From Birth

There's even some evidence to suggest how we're born matters. People born via c-section might have different personalities as adults (Verdult, 2009). How we're born is an additional differential factor that has an effect on our personality.

Everything Interacts

These are just some examples and a way of categorizing the unquantifiable number of environmental and genetic factors that shape our personalities. Importantly, each of these nature and nurture factors interact in different ways and different magnitudes for everyone. The number of factors is unimaginably large, but the number of combinations of those factors with differential magnitudes is infinite – and it makes us who we are. Before getting into measurable, observable, and scientifically verifiable relationships of factors and personality, it is obvious, clear, and empirically supported that there are unique factors in everyone's lives which contribute to who they are. The goal of personality research should be to understand those

differences, identify the largest effects, and apply those findings in meaningful ways to people's lives. Not to question whether the environmental differences matter depending on their quantifiability.

Everyone has a unique personality that the world has not seen before and will not see again.

Everything is a Model; Better Models Are More Accurate Representations

We know with absolute certainty that everyone has a unique personality, which needs to be built on. To study personality psychologists need to measure it – quantify it, which is challenging. Personality is invisible and the true score, the theta (θ), is unobservable. While the trait itself can't be directly observed, how it manifests can be observed and measured. That information can then be used to infer the personality trait using inferential statistics.

Likert Scales and Modeling

Traditionally, as researchers, we measure personality by asking participants to rate their agreement to statements based on their self-perception using a Likert scale (e.g., BFI-2; Soto & John, 2017; IPIP-50; Goldberg 1992). Likert scales artificially discretize continuous, unique personalities into just a few categories. A colorful, dynamic personality gets filtered into a score that is limited in range.

Still, even a mediocre representation of people is extremely useful, which attests to just how bright the light is that shines in each of us is; it even shines through when it's filtered through messy measurement tools. To name just a few empirically supported relationships - personality predicts subjective well-being, job satisfaction, turnover intentions, GPA, and career success (Roberts et al., 2007; Noftle & Robings, 2007; Cheng & Ickes, 2009; Judge et al., 1999).

The current tools (i.e., Likert-based scales) to measure personality are validated and effective, even if they have flaws (e.g., Soto & John, 2017; Chernyshenko, 2002; Drasgow et al., 2012). After using an established tool to quantify personality, the next step is to model it.

A model is a simplification. For example, pictures, maps, language, and human perception are all models. A picture is a model of the scene it is capturing. It is not literally that scene, it's a simplification. A map is a simplification of a landscape. A "perfect" map of the entire world would be a second Earth. Language is a simplification of the thoughts and feelings and ideas the speaker is trying to communicate – there's information lost between what the speaker is trying to communicate and the words chosen. Human perception is a model of the environment – we don't see all wave lengths of light or hear every pitch of sound and we all process each stimuli differently (Gregory, 1972).

After measuring the construct and obtaining numerical representations of the personality trait for each person in the data set, the next step is to model, or represent, all of those scores. Once a model is generated, inferences about the population from which the sample is representative of can be made. The quality of those inferences depend on the validity and reliability of the measure and quality of the model. The measure will be flawed and the model will be a simplification of the measure. The degree to which each of these are flawed is precisely the degree to which the inferences from the model can be relied upon.

Reliability and Validity

Reliability is the ceiling of a model's validity. The accuracy of a model's representation of the measure is ceiling of the validity of inferences that can be drawn from the model. Better models are more accurate representations of the measure they're modeling. The "best" model would be one that perfectly represents the data it's modeling – it's a perfectly accurate

representation. However, inferential statistics depend on the model being generalizable to the population from which the sample was drawn (e.g., the model needs to perform well out of sample). Overfitting, or fitting the 'random quirks' of the sample oftentimes reduces out-of-sample model performance, which is a concern when attempting to discover generalizable truths of human psychology. This is the bias/variance trade-off problem The best models are the most accurate representations of the rest of the population from which the sample was drawn. The purpose of the sample is to generate a model that represents the population from which the sample was drawn, and not just the sample.

Problem 1 - Uniqueness

Every single person is unique and has a unique personality which the world has never seen before. Personality researchers want to measure and model personality. The traditional method of measuring personality traits, Likert scales, will produce scores which suggest two people have the exact same level of conscientiousness. The personality scale used in this research was specifically designed for conscientiousness (the Chernyshenko Conscientiousness Scale; CCS; Cherynshenko, 2002) and will produce scores that will suggest two people have the exact same level of conscientiousness. The CCS is a 60-item scale measuring conscientiousness on a 1-4 Likert response scale. This scale can only produce 240 unique scores, which does not leave room for the uniqueness of personality.

Problem 2 – Faking

The powerfully predictive ability of personality has led organizations to use it as a component of their selection processes, which has yielded criticism about the vulnerability to faking that is indissoluble from traditional self-report personality measures (Roulin & Krings, 2020; Goffin & Boyd, 2009). Additionally, there is evidence to suggest traditional self-report

personality measures also suffer from low validity in selection contexts (Morgeson et al., 2007a, 2007b). These shortcomings make sense, since most people can "strongly agree" with the statement "I am always on time" if they choose to. Many people will choose to do so when there's a job on the line, which, to be clear, is not a necessarily insidious action. However, it seems to be the case that when candidates are competing for a position, everyone is very conscientious, emotionally stable, and agreeable.

Personality is often measured with Likert response scales (Ozer & Reise, 1994), despite the fact participants can easily inflate socially desirable traits (e.g., conscientiousness) to appear more attractive for a job role – especially in high stakes situations. In high stakes situations, it is easy to inflate socially desirable traits on Likert scales. Putting on a "best face" for a job interview is normal and part of why selection systems try to use more concrete indicators, such as ability tests (Schmidt & Hunter, 1998). However, most people have worked with someone who is capable at the job, but impossible to get along with. Having a way to get an accurate personality estimate in the selection system would help screen out those types of individuals. Faking is a vulnerability of Likert scales which make it less useful in selection settings than it has the potential to be.

Solution 1 - Forced Choice

Forced choice (FC) is a scale format requiring participants to choose the statement they agree with more between two statements. FC data measuring non-cognitive ability, such as personality, is psychometrically best analyzed using an ideal-point Item Response Theory (IRT) model-- a statistical technique which maps the relationships between items, traits, and respondents (Drasgow et al., 2010; Chernyshenko et al., 2007). This is juxtaposed with analysis of cognitive ability, which is better analyzed with dominance models. Ideal-point models don't

assume a “best” answer, they assume that when endorsement of a scale-point occurs, it’s due to the relationship between the item and the trait level. A dominance model assumes that stronger endorsement corresponds to higher levels of the trait. For non-cognitive assessments, such as personality, there isn’t a “best” or “correct” answer, there’s only the answer that matches the trait the closest, which is what an ideal-point IRT model unfolding process allows for. Whereas for cognitive assessments, there is a best answer and responses closer to that answer correspond to higher levels of the trait.

FC addresses some of the limitations with Likert scales. With FC, there isn’t a need for participants to understand the scale in a uniform way and neutral responding is eliminated. Importantly, previous research has provided evidence that forced choice is faking resistant, when compared to traditional scales (Cao et al., 2019). When FC blocks are assembled, they are balanced for social desirability. *Social desirability* refers to participants’ tendencies to respond in a way that makes them appear better to others (Falk & Cai, 2016). So, when participants are motivated to inflate their scores, selecting the FC statement that makes them look appear to be a better candidate is a more challenging task than selecting a stronger agreement on a Likert scale.

As discussed earlier, there are only 240 possible responses to the CCS, which limits the extent to which a personality measure can capture the uniqueness of an individual. Forced choice ameliorates this limitation. A two-alternative forced choice block, which is what was used in this study, over 45 blocks, increases response options from 240 to 45^2 or 2,025, a clear improvement. It is also worth noting that, as a normally distributed variable, most of those 240 scores will cluster toward the middle. Whereas for FC, the high quantity of responses should be closer to evenly spread across all 2,025 possibilities.

An important aspect of FC is the underlying assumption of Thurstone's Law of Comparative judgement, which, as it relates to FC, states that when an individual endorses a statement, it's the relationship between the statement and the individual's trait level that is causing the endorsement and not the presence of the other statement, which is the required to satisfy the assumption of local independence in Item Response Theory (IRT; Brown, 2016) Understanding the decision-making process underlying FC could provide insight into this relationship. In fact, the design for this research lends itself very well to understanding that relationship. Participants are providing a rationale for their decision, which could provide insight into that decision-making process. A supplementary analysis was conducted to explore the decision-making process underlying forced choice (See Appendix B).

Solution 2 - Natural Language Processing

Natural language processing (NLP) is a process of converting text data into quantitative data, which can be used to create a model to estimate, in this case, conscientiousness and its underlying facets by finding the best numerical representation of text features to predict an outcome variable. Since creating a story is inherently more challenging than simply selecting a different degree of endorsement on a traditional scale, it would make sense, in theory, that NLP-derived trait estimations might be faking resistant. Furthermore, empirical evidence supports the idea that NLP-derived personality models are resistant to faking (Heron et al., 2024). Additionally, an NLP-based approach offers the benefit of resembling a peer-report measure, which can further enhance the validity of the responses. Perhaps NLP can be conceptualized as a behavior/peer-report measure of personality - unlike traditional scales and forced-choice (FC) scales NLP taps into grammar, word choice, vocabulary, and similar features of writing. Behavior measures of personality have been compared and contrasted with self-report measures

of personality for a long time and each have benefits (Golfriend & Kent, 1972). Traditionally, participants are asked to match their trait level to the item they're presented with. However, with NLP narratives, the narrative provided by the participant is evaluated by the algorithm in a pseudo-peer-report fashion, which can assess these linguistic features more effectively. The use of NLP in personality research is a rapidly developing field and there are a lot of possibilities.

The number of unique scores generated by an NLP-powered model is, quite literally, infinite. Using text data to derive a personality score allows for the uniqueness of personality to be captured with a statistical model. Machine-derived scores have demonstrated acceptable reliability and convergent validity, but poor discriminant validity (Fan et al., 2023; Sun, 2021).

Ensemble

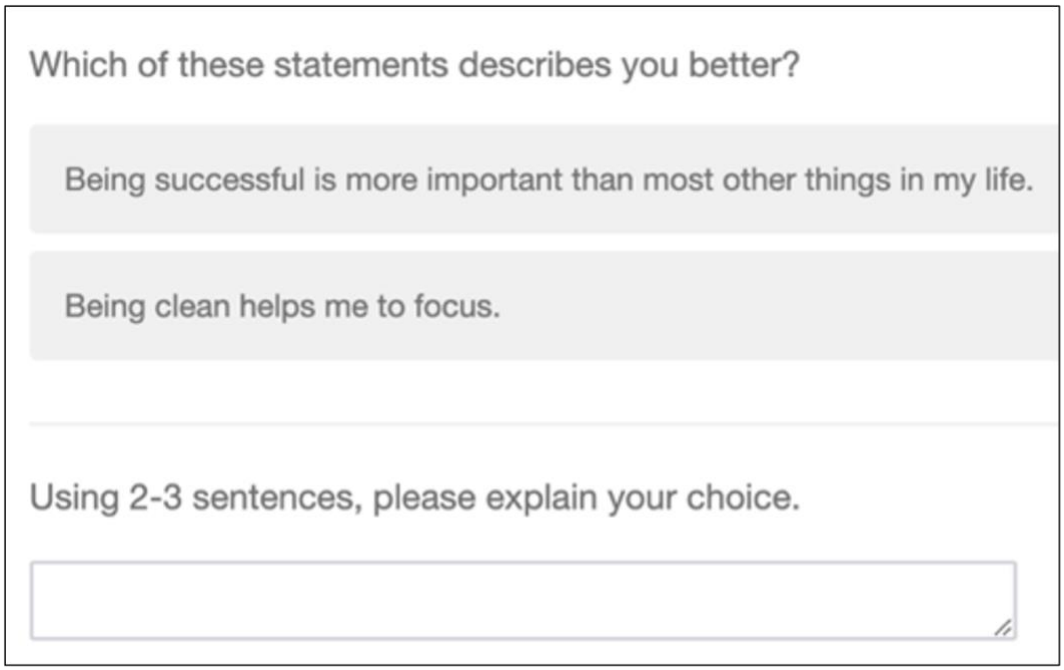
Both NLP and FC are established methods of measuring personality traits and help address the faking and response option problems associated with measuring personality with Likert scales. A model with both with estimates should yield a score that is more resistant to faking and unique than either method on its own, especially when compared to traditional self-report measures. The important question this research is gathering evidence for is whether or not an ensemble model is valid.

Ensemble is, technically, a machine learning technique of combining weak learners to better estimate an outcome of interest. Ensemble models might provide an optimal solution to the bias/variance trade-off - complicated models tend to have less bias but overfit, whereas simpler models tend to have more bias but underfit (Orru et al., 2020). In the present research, among the many ensemble methods available, linear regression was chosen for its parsimony, accessibility, and explainability. This chosen method utilizes both the FC and the NLP trait estimates as predictors in a regression model to predict the CCS scores.

The number of possible scores produced by the ensemble is nearly unquantifiable, which is appropriate given the uniqueness of personality. While the present research doesn't examine this, perhaps this multi-method approach will be more faking resistant than NLP or FC alone. Before considering that aspect, evidence for the validity of the ensemble model must be gathered, which is what the present research attempts to do.

The data for NLP and FC are collected simultaneously (see Figure 1.1)

Figure 1.1. Example of FC pair with a text box for explanation of their choice.



Participants endorse a FC statement then provide a text explanation of their choice. The greater the variance in the words used, including both the quantity and variety of words, the more effectively the model can perform. This method creates more variability across participants. As an example of the importance of variability, consider a scenario where, out of five hundred participants, only five used the word 'albatross' and all received high scores on a personality test. In this case, that word would become a significant predictor for those five individuals. However, if the word 'albatross' were used by everyone, it would fail to provide any distinctive

information. While this example is relatively oversimplified—it doesn't account for the context or manner in which a word is used—it illustrates a key point: the larger the variance in the words used, the more accurately the model can predict the outcomes it is trained on. Participants choosing which statement in the FC block to respond to will increase the variability of the words used by participants, necessarily. Participants are now responding to different prompts and different combinations of prompts. They're still responding to the same block so there's going to be commonalities, but the choice to provide text about the block or on the one they chose (or didn't choose) is information in itself. Having the choice itself should increase variability. FC and NLP are complimentary methods and an ensemble could prove to be an effective method of using them both together.

While an ensemble of NLP and FC has many potential upsides and can potentially further ameliorate some of the problems of measuring personality with Likert scales, there is no evidence that an ensemble would produce valid scores, which is precisely what the present research attempts to gather. As such, there are two guiding research questions to examine the psychometric properties and criterion-related validity.

The first research question (RQ1) is **what is the evidence to suggest the ensemble model yields scores demonstrating construct validity?** And the second research question (RQ2) is **what is the evidence the ensemble model yields scores that provide evidence of criterion-related validity?** These are the fundamental questions. In the process of conducting all these complex and computational expensive analyses, the final model needs to still be measuring conscientiousness. As such, the psychometric properties, (e.g., reliability, convergent validity, and discriminant validity) of the ensemble model will be examined to answer the first research

question. The resulting trait estimation also needs to be related to the other variables it's theoretically and empirically related to, which addresses the second research question.

Chapter 2 - Methods

Preface

The purpose of this research was to gather evidence as to whether forced choice and natural language processing estimates can be combined, or *ensembled*, and produce a valid estimate. There are many methods of embedding and training language models (e.g., the BERT family (roBERTa, distilBERT, etc.); Devlin et al., 2018), but not much evidence as to which is optimal. The same is true for ensembles. Each of these decisions have important downstream effects on the results and need to be considered. However, for this thesis, the purpose was to gather evidence as to whether an ensemble can be applied to NLP and FC data to produce a valid estimate. The guiding criteria for choosing materials and analytical methods was accessibility and parsimony. The analytical methods selected are among the most parsimonious available and, to maximize accessibility, all analysis was conducted in R using publicly available packages.

Research Questions

(RQ1) Is there evidence to suggest the ensemble model yields scores demonstrating construct validity?

The primary purpose of this research is to gather evidence that could suggest how valid this method of combining measures is. To examine construct validity a Multi-Trait Multi-Method (MTMM) Table was constructed to compare the correlations between the ensemble score, the CSS score, the FC score, and the NLP score using some of the variance decomposition procedures outlined in Woehr et al., (2012). Comparing correlations from the same trait and facets across the FC, CCS, NLP, and ensemble models will provide evidence as to whether the ensemble model is actually measuring what's it's supposed to.

(RQ2) Will the ensemble model yield scores that provide evidence of criterion-related validity?

Another consideration is the criterion-related validity for the ensemble model. The strength of the relationship between two variables partially depends on the quality of both measurements. For example, theory and previous research suggest that individuals higher in conscientiousness should have better GPAs (Noftle & Robins, 2007; Cheng & Ickes, 2009). The strength of that relationship is, in part, dependent on the amount of measurement error. In theory, if the ensemble model were a more accurate measure of conscientiousness, then the positive correlation between GPA and conscientiousness would be stronger. So, if the correlation between GPA and the ensemble score is stronger than the FC and CCS scores, then that would provide evidence of an improved measurement. The same should be true of subjective well-being, which is positively related to conscientiousness (Roberts et al., 2007). To examine whether the ensemble method-based trait estimates demonstrates criterion-related validity, a hierarchical linear regression will be constructed to examine incremental validity on the outcome variables subjective well-being and GPA.

Participants

N = 466 participants were collected from both a large midwestern university (n = 405) and the crowdsourcing website Prolific (n = 61). The sample was 69% women, 78% white, $mean_{age}$ = 22, $median_{age}$ = 19. Participants recruited through the student pool were compensated with course credit. Participants on Prolific were compensated with $8/hour.

Before analysis was conducted a total of n = 176 participants were removed. Of those, n = 143 participants were removed for failing to complete the entire survey (the highest completion percentage for removed participants was 69%). Additionally, n = 9 participants were

removed for completing the survey too quickly (less than 15 minutes for the highest removed participant). Furthermore, n = 23 participants were removed for failing both attention checks and another n = 1 participant was removed for failing a single attention check and not providing useable data for the outcome variables. However, participants who only failed a single attention check were left in the dataset.

The sample size (n = 466) is based on considerations of model complexity. These considerations are based primarily on IRT considerations, since the language model is supplementary to it. Stark et al. (2006), provides evidence that, to minimize differential item functioning and obtain sufficient power when using 15 forced choice pairs, 500 participants are needed for dichotomous IRT models (the present research uses 45 pairs). Dalal et al. (2014) provides evidence that complex models similar to the present research requires closer 700. These general sample size recommendations is reinforced by other research on sample size requirements for IRT models and structural equation models (Wolf et al., 2013; Lance et al., 2010). However, the quantity of items in the current research should generate enough power to compensate for the smaller sample size.

Materials

The personality trait conscientiousness, which, in the materials chosen for this research, includes the facets of industriousness, orderliness, self-control, responsibility, virtue, and tradition, has been linked to positive outcomes such as higher job satisfaction, lower turnover intentions, and higher subjective well-being (Roberts et al., 2007). It could be argued that conscientiousness is one of the most indispensable personality traits in the workplace, as highly conscientious individuals tend to be diligent, planful, and responsible (Barrick et al., 1991).

Two validated conscientiousness scales, which were developed with ideal-point IRT models in mind, were used for this research (Chernyshenko Conscientiousness Scale (CCS); Cherynshenko, 2002; Tailored Adaptive Personality Assessment System (TAPAS); Drasgow et al., 2012). The CCS is a 60 item scale measuring each of the six facets ($\alpha = .91$) and there are 90 TAPAS items measuring each of the six facets ($\alpha = .90$). Both of these scales ask participants to rate their agreement on a four-point scale with 1 = Strongly Disagree, 2 = Disagree, 3 = Agree, and 4 = Strongly Agree (all items for both scales are available in Appendix A). Previous research has provided evidence that 5-point scales are inappropriate for ideal-point IRT models (Dalal et al., 2014; Sun et al., 2019). The neutral response option makes it challenging for the ideal-point model from finding a probabilistic account of responding. TAPAS and CCS are on the same scale, constructed with ideal-point models in mind, and allowed for the use of different validated scales for assembling forced choice pairs and for training the language model. Conscientiousness is a valuable personality trait for the work force and previous research has developed an abundance of materials that fit the purpose of this research.

Assembly of the 45 forced choice pairs was conducted using the *autoFC* package in R (Li et al., 2022). The pairs were optimized on dimensionality, social desirability, and extremity – in that priority order. Balancing for social desirability has been shown to reduce social desirably responding (Christiansen, 2005). All pairs contained statements from different facets (e.g., if the first statement measured virtue, the second statement would be from one of the other five facets). The pairs were balanced for social desirability to reduce social desirability bias. Finally, the pairs were balanced for extremity (e.g., if the first statement measured the lower end of the facet virtue then it's pair would also measure on the low end). All pairs are multidimensional (they are from different facets), balanced for social desirability, and balanced for item extremity.

The outcome variables chosen for this research are all theoretically related to conscientiousness. See Table 2.1 for a complete list of outcome variables collected. Basic demographic information was also collected. All scales and items used in this research are in Appendix A.

Table 2.1. Summary of Outcome Variables Collected

Scale	Number of Items	Citation
Emotion-Regulation Self-Efficacy	5	Wang et al., 2011
Identity with the Job	4	Farmer & Van Dyne, 2010
Subjective Well-Being	5	Diener, et al., 1985
Counterproductive Work Behavior	10	Spector et al., 2010
Organizational Citizenship Behavior	10	Spector et al., 2010
Job Satisfaction in General	8	Russell et al., 2004
GPA	2	-----

Procedures

The survey took approximately 90 minutes to complete. Participants completed the entire study through Qualtrics. First, participants were presented with FC pairs and asked to provide a short two- to three-sentence explanation of their choice (Figure 1). Participants could not move on to the next question until a statement in the pair was selected and a text response of at least 25 characters was provided, which ensured participants provided some information while allowing them to answer concisely. Selecting a pair then providing a text response was repeated for 45 trials. Participants also responded to the single statement (SS) version of each FC item, which was used for item calibration. To estimate the trait (theta) from FC blocks, the difficulty and discrimination parameters for each item need to be known. The order of the SS block and the FC block was randomized since they contained the same items. The FC and SS items are from the Tailored Adaptive Personality Assessment System (TAPAS) research items pool (Drasgow et al.,

2012) and FC pairs were assembled using the *autoFC* package in R, which allows forced choice blocks to be balanced for social desirability and item extremity (Li et al., 2022).

Next, participants completed the 60-item Chernyshenko Conscientiousness Scale (CCS; Cherynshenko, 2002; Hill & Roberts, 2011), in which respondents responded on a 4-point Likert-type scale. Finally, participants completed a set of demographic and criterion measures (e.g., subjective well-being, job satisfaction; see Table 2.1 for outcome measures; see Appendix A for a complete list of all measured variables and items).

Analytical Procedures

Forced Choice

To get trait estimates from the FC items, the single statement (SS) items were calibrated with the Generalized Graded Unfolding Model using the *ggum* package in R (GGUM; Roberts et al., 2000). Then, using the parameters alpha (item discrimination – the amount of information the item provides at a given trait level, or, in other words, how well the item can discriminate among participants at different trait levels) and delta (item difficulty – where along the trait continuum the item measures, or, in other words, where along the trait continuum the probability of endorsing the item is 50%), the FC pairs were estimated using the Generalized Thurstonian Unfolding Model using the *fcscoring* package in R (GTUM; Zhang et al., 2023). The GTUM model appropriately models the unfolding process, which assumes item endorsement is a function of the relationship between the item and the underlying latent trait level. It's important to note, the estimates generated in this thesis were unstable, meaning they had low reliability and poor relationships with other conscientiousness measures. This is likely due to the facets of conscientiousness being too closely correlated with each other for the trait estimates to be accurately estimated.

The GTUM model produced a trait estimate for each of the six facets; industriousness, orderliness, virtue, traditionalism, responsibility, and self-control for each of the 466 participants. The absolute value of these six facet estimates were averaged to produce a conscientiousness trait estimate for each person. The final product is trait estimate for each of conscientiousness, industriousness, orderliness, virtue, traditionalism, responsibility, and self-control for each of the 466 participants. In the results section, these scores will be referred to as the FC model and it is these scores that will go into the ensemble model as a single "item".

Text Data – Language Model

To get trait estimates from the text data, the NLP algorithm used a pre-trained Bidirectional Encoder Representations from Transformers (BERT-base) model (Devlin et al., 2018). Specifically, the embeddings were generated in R with the *text* package using BERT-base-uncased (Kjell et al., 2023). Before the embeddings were generated, pre-processing of the text data was automatically conducted. Prior to tokenization, all words were converted to lowercase, punctuation was removed, and white space was used to break the input text down into preliminary words. Then, the texts are tokenized using the WordPiece tokenization method, breakingthe sentences into words that BERT has been trained on. After the tokenization, the input is padded or truncated to make all inputs? the same number of tokens. Then, additional embeddings and tokens are added to differentiate segments of text and to identify positions of the text. Since BERT is pre-trained, it doesn't require tuning and can work effectively on small sample sizes, like the one in this research.

Before tokenization, all the words in the input texts were converted to lowercase, and all the special characters such as punctuation were removed. Subsequently, the texts were tokenized using the WordPiece method, which breaks them down into words recognized by BERT.

Following tokenization, the inputs are either padded or truncated to ensure uniformity in the number of tokens. To further enhance text analysis, additional embeddings and tokens are incorporated to distinguish between segments and to pinpoint specific positions within the texts. Since BERT is pre-trained, it requires no additional tuning and performs effectively even with the small sample sizes used as in this study

The embeddings contained 768 columns, each representing a feature of the text data that was identified from the BERT-base-uncased embedding process. Using these embeddings and identified features, seven models were trained simultaneously on CCS scores for conscientiousness, industriousness, orderliness, virtue, traditionalism, responsibility, and self-control, using ridge regression as the training method and was also done through the *text* package in R (Kjell et al., 2023). Each model was trained on all 466 participants. Since the purpose of this research is not to generate a model which can be generalized to the population, but rather to compare the psychometric properties of the ensemble model to its subsidiary models, out-of-sample validation or cross validation of the models was not conducted.

The resulting NLP model contained 466 rows representing each participant and 315 columns representing an NLP-model-predicted score for each of the 45 responses across the seven CCS scores (45 x 7 = 315). In other words, a predicted score for each participant's conscientiousness, industriousness, orderliness, virtue, traditionalism, responsibility, and self-control was generated for each of their 45 responses to each of the FC blocks. To generate a single score for each of the seven CCS scores, the 45 predicted scores were aggregated using linear regression with the corresponding CCS score as the outcome. The final product of this process was an NLP trait estimate for each of conscientiousness, industriousness, orderliness, virtue, traditionalism, responsibility, and self-control for each of the 466 participants. In the

results section, these scores will be referred to as the NLP model and it is these scores that will go into the ensemble model as a single "item".

Additional Language Model – Forced Choice Statement Text

A second NLP was generated using precisely the same procedures as outlined above. The first NLP model used the response text where participants explained their decision. The second NLP model used the forced statement the participant selected from the pair. For example, one of the FC blocks contained the statements, "I hardly ever finish the tasks I start" and "Taking on responsibility is just asking for trouble". Each participant selected the one that matched them better and, for this second model, the selected statements across the 45 blocks was used as the text for an additional language model.

This second language model will go into the ensemble but it does not appear in the results. The decision to only include it in the ensemble model was based on the large aim for the research. This extra model can only exist because of the simultaneous data collection of FC and NLP. It doesn't belong to FC because it's a language model, but it also doesn't belong to the NLP-R model because it's FC. It belongs to the interaction of the NLPFC format, which is why it appears only in the ensemble model. The resulting model produced a score a trait estimate for each of conscientiousness, industriousness, orderliness, virtue, traditionalism, responsibility, and self-control for each of the 466 participants. Those scores will be each used as an "item" in the ensemble model.

Ensemble

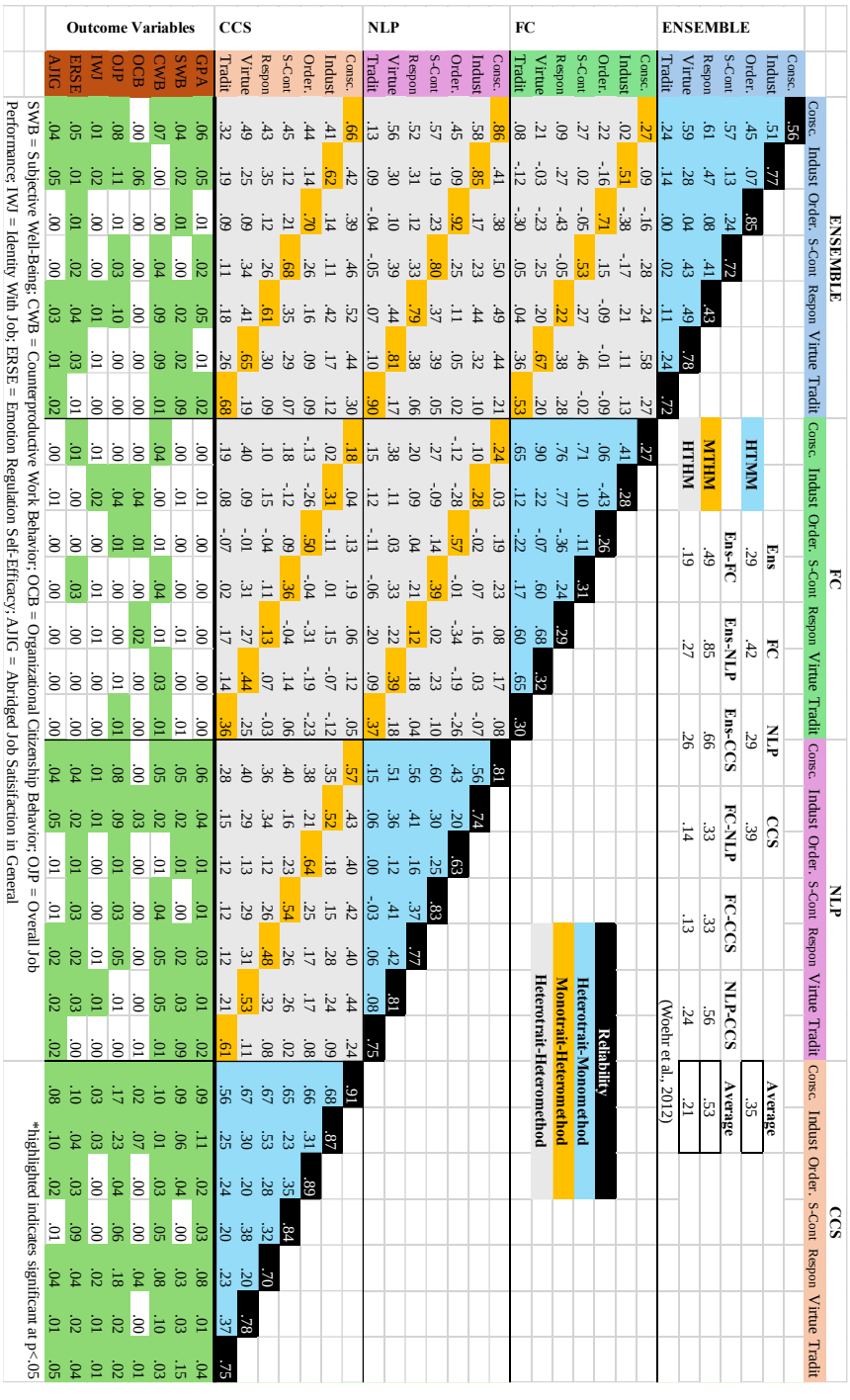
Each of the 466 participants had three scores for each of conscientiousness, industriousness, orderliness, virtue, traditionalism, responsibility, and self-control, a total of 21

scores. The goal was to have a single score for each person on each of the seven dimensions, or, seven total scores for everyone using the information gathered from each of the three scores. The first score was the forced choice generated model – FC. The second score was derived from the response text participants provided to each of the blocks – NLP-R. The third score was derived from the text of the forced choice statement participants selected – NLP-S. To ensemble these three scores each score was considered an "item" and the three "items" were put into a linear regression to predict the corresponding CCS score. This regression formula was applied to each of the seven dimensions - [Ensemble Score= FC_x + NLP-R_y + NLP-S_z + intercept]. The final result of this process was a trait estimate for each of conscientiousness, industriousness, orderliness, virtue, traditionalism, responsibility, and self-control for each of the 466 participants. These scores will be referred to as the ensemble model in the results section.

Chapter 3 - Results

MTMM Table Results

Figure 3.1. MTMM Table comparing conscientiousness and its facets across four measures and with criterion outcome variables



To answer RQ1, a Multi-Trait Multi-Method (MTMM) table was constructed with each of the six facets and the conscientiousness estimate across the CCS scale, the NLP model, the FC estimate, and the ensemble model (See Table 3.1). Each facet and trait estimate are correlated with each other facet and trait estimate across all methods. The reliabilities for the CCS are Cronbach's alpha. The reliabilities for the FC scores are Person Separation Index (PSI) reliability coefficients. PSI is an IRT estimate of the distance between the estimated trait level and the true score and is comparable to Cronbach's alpha (Lin, 2021). The NLP reliability is standardized Cronbach's alpha for the item-level NLP estimates. The ensemble reliability is the standardized Cronbach's alpha for the model level.

The black boxes in MTMM table along the diagonal show the reliability of each trait or facet estimate for each measurement method. The light blue boxes show the correlation between different traits measured with the same method (HTMM; hetero trait mono method). The grey boxes show the correlation between different traits with different methods (HTHM; hetero trait hetero method). The orange boxes show the correlation between the same trait measured with different methods (MTHM; mono trait hetero method).

Reliability

The reliability of the ensemble model for the trait level *conscientiousness* is .56 and .43 for the facet *responsibility*, which are low. While the other five facets are all above .7, that is still low. Ideally, a reliability of .8 is considered acceptable. The low reliability is explainable by the unstable forced choice estimates, which have very low reliability. The estimates are perhaps unstable because of the high correlation between facets. The HTMM correlations for the CCS facets are mostly in the .2 - .4 range, which makes sense since, collectively, they make up the personality trait *conscientiousness*. However, for the GTUM model, it appears that the

intercorrelation of facets made it difficult for the model to converge. Even using a larger (bootstrapped) sample and tighter priors, the estimates were still unstable. This was unforeseen design limitation which resulted in the poor performance of the FC model. For the ensemble model, the reliability was calculated using three "items". These items were the FC score, the NLP score, and the NLP-derived FC statement score. The low reliability of the FC model caused the reliability of the ensemble to decrease.

HTMM and HTHM

For different traits measured with the same method (indicated in blue), they should be more correlated to each other than different traits measured with different methods (Woehr et al., 2011). Trait estimates produced from the same method should have a stronger relationship to each other than trait estimates produced from different methods. For the ensemble model, the average HTMM correlation is .29 and the average HTHM correlation is .19.

MTHM

For the same trait measured with different methods, they should correlated closely. MTHM scores are indicated in orange. CCS is the benchmark to compare with. On average, the ensemble model is more closely correlated with CCS (.66) than either NLP (.56) or FC (.33). The improvement from the NLP model to the ensemble model is almost entirely due to the NLP-FC statement model being added in.

Criterion Variables in MTMM Table

In the far right columns of the MTMM table are the R^2 results of each trait from each measure predicting each outcome. This is different from the results of the hierarchical regression, which considers whether the ensemble model predicts above and beyond what is predicted by the other models. The cells that are highlighted in green indicate a significant relationship. The

instability of the FC scores can be seen more clearly here. There are not many significant relationships with the outcomes.

This also visually displays the "ground truth" problem. The CCS is our ground truth, which each model was trained on. The magnitude and quantity of significant relationships between the CCS to NLP is diminished. Then, again, between the ensemble and NLP models, the magnitude and quantity of significant relationships is diminished. This is, in part, explainable by the poor performance of the FC scores.

Hierarchical Regression Results

To answer RQ2, two hierarchical linear regressions were conducted. The purpose of these is to examine whether the ensemble model explains the variance in the related outcome variable *above and beyond* what is explained by the CCS. Step 1 is the base model, CCS predicting SWB/GPA. Step 2, FC scores get added in. In step 3, the NLP scores get added in. In step 4, the ensemble scores get added in. The change in R^2 provides evidence as to whether the ensemble model is explaining more of the variance in the outcome variables.

The scores for GPA were adjusted in a few ways. Participants provided their GPA and the highest possible GPA that could be earned at their institution. Those numbers were divided against each other to ensure everyone was on the same scale. A ceiling of 1 was imposed (some participants indicated a GPA above a perfect GPA, which some schools allow for but would confound comparisons to other institutions). Finally, participants who indicated an impossible GPA (7.0 on a 4.0 scale, for example) were removed from the analysis. Between those who didn't provide a response and those who provided impossible responses, a total of $n = 41$ observations were removed from the GPA analysis.

Table 3.1. Hierarchical Regression with Subjective Well-Being as the outcome

SWB	Step 1		Step 2		Step 3		Step 4		ΔR^2
Measure	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	
CCS	-0.08	0.01	-0.08	0.01	-.07	0.01	-0.08	0.01	
FC			-.00	0.11	0.01	0.11	0.00	0.11	
NLP					-0.03	0.02	-0.06	0.04	
Ensemble							0.04	0.04	
R^2	.09		0.09		0.10		0.10		0.00

*Each scale is a conscientiousness score; CCS = Chernyshenko Conscientiousness Scale; FC = Forced Choice; NLP = Natural Language Processing

Table 3.2. Hierarchical Regression with GPA as the outcome

GPA	Step 1		Step 2		Step 3		Step 4		ΔR^2
Measure	*B*	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*	
CCS	-0.08	0.01	-0.08	0.01	-.07	0.01	-0.08	0.01	
FC			-.00	0.11	0.02	0.11	0.00	0.11	
NLP					-0.03	0.02	-0.06	0.04	
Ensemble							0.04	0.04	
R^2	.09		0.09		0.10		0.10		0.00

*Each scale is a conscientiousness score; CCS = Chernyshenko Conscientiousness Scale; FC = Forced Choice; NLP = Natural Language Processing

Chapter 4 - Discussion

The primary aim of this thesis was to implement a research design which collected forced choice data and text data simultaneously, then examine the psychometric properties of an *ensemble* model generated from the FC data the text data. The present research evaluated the psychometric evidence of the validity of a multi-method ensemble approach to measuring conscientiousness.

The ensemble model yielded good criterion-related validity and convergent validity, acceptable discriminant validity, but poor reliability. The ensemble model did not demonstrate incremental validity above and beyond the other estimates – it didn't explain more of the variance in GPA or subjective well-being than the Chernyshenko Conscientiousness Scores, the NLP-derived scores, or the FC scores. The poor reliability of the ensemble model can be contributed to the poor performance of the FC scores. The lack of incremental validity is confounded by the poor performance of the FC scores and it's impossible to say how the ensemble would have performed with better FC scores. Results, limitations, implications, and next steps are discussed.

RQ1 Results

Overall, the results of the MTMM table are encouraging. Despite the poor performance of the FC model, the ensemble model yielded improvements over the NLP model, mostly due to the inclusion of the FC statements in the NLP model. This, in itself, is encouraging and creates promising future directions, which will be discussed later.

RQ2 Results

The results of the hierarchical regression are attenuated by the poor performance of the FC model. Adding the NLP model added incremental validity, but the ensemble model did not

demonstrate incremental validity above the what the NLP and CCS models did alone. In fact, the betas and standard errors are almost exactly the same for both GPA and subjective well-being. If the FC model was better then, perhaps, the ensemble model would perform better as well. It is unclear whether the ensemble would provide incremental validity above and beyond the NLP and FC models if the FC model contained stable estimates.

Limitations

Ensemble reliability

Traditionally, reliability is considered the ceiling for validity. In this case, the reliability might be better considered a pseudo-reliability. The reliability was generated by using the three scores as "items", which means the poor reliability of the FC model is bringing down the reliability of the ensemble model. However, the rest of the MTMM table doesn't suggest there are validity concerns with the ensemble model. The ensemble method of linear regression appears to be buffering the model from overweighting the FC scores.

FC Pair Assembly

Part of the low reliability and poor FC performance could also be due to how the FC pairs were assembled. They were balanced for social desirability and item extremity. However, there's new evidence that using mixed extremity could have helped (Li et al., 2024). One challenge with this is that social desirability and item extremity tend to be highly correlated. Additionally, exploring methods of using items with similar discrimination and difficulty parameters might help with estimating.

Model overfitting

The method of ensembling was linear regression, and it was used at multiple stages. First, the text data was trained on the CCS score, which generated predicted scores for each of 90 items

– 45 forced choice statements and 45 responses to the statement. The response data was aggregated by using the 45 responses to predict the CCS score. Then, predicted scores were generated and the process was repeated for the FC statements that were selected. Using those two sets of predicted scores and the FC trait estimates, they were regressed onto the CCS scores and the predicted scores that were generated were the ensemble predicted scores. This process has the danger of overfitting and a validation dataset could help to evaluate model performance out of sample. However, this research has provided evidence for and against the validity of an ensemble model. The design limitation of the treating facets as separate dimensions needs to be addressed, and it will in future iterations of this research.

Adding in the FC statement NLP scores to Ensemble model

An important decision that was made during the analysis process was whether to use the text from FC statements in the NLP model at all and then whether to include it in the language model part or the ensemble part. It was included in the ensemble part, since the text is from the forced choice statements and a large goal of this research was to see how FC could compliment NLP. So, adding in the NLP-derived FC statement scores to the ensemble allowed the effect of the compliment of FC and NLP to be seen.

Fatigue

This was a 90-minute survey. There's a high likelihood that participants started to get fatigued when taking it and that is confounding some of the results. 143 participants were removed because they didn't finish the survey. Presumably, they simply didn't want to finish. This also creates some bias in the data – it could be participants that low in conscientiousness were screened out. Future versions of this format will need to be shortened.

Quality of Input

Between the different format for gathering text data and the length of the survey, examining the quality of input will be important for future research. This format could be producing more variability in the text data compared to simply responding to a question. However, it's unclear whether the text response will build on the FC statement or whether it will restate it (not provide much information above and beyond what the FC statement provides). Examining the frequency of terms used and their importance could provide useful insight into the quality of the text data. TF-IDF (Term Frequency-Inverse Document Frequency) is a promising analysis for exactly that. The cosine similarity of the TF-IDF vector generated from the response text and the TF-IDF vector generated from FC statement text can be compared to gauge how related the FC statement is to the response text. In other words, evaluating the similarity between the FC statement and the response text will be indicative of the quality of the text input.

Implications

Results

For RQ1, it appears that the ensemble model is still measuring conscientiousness and its facets, which is encouraging. For RQ2, it appears that the ensemble model does not add incremental validity. The limitation of the FC scores makes it challenging to conclude the efficacy of this method, as of right now. However, there the results are, overall, promising and will guide future research.

Embedding the Forced Choice Statements

Embedding the forced choice statements and training them on the outcome variable appears to add to the convergent validity. The ipsative nature of FC is being leveraged by a language model. Traditionally, these are best scored with IRT, but the ensemble method allows for more information to be extracted from the FC statements. If it can be incorporated into the

model well, it could prove to be a boon for this format. Even aside from this format, it can perhaps be used to supplement a FC assessment.

Uniqueness

The possible scores generated from the ensemble model are functionally infinite, which is theoretically appropriate, given the uniqueness of each person's personality. This is one of the problems discussed earlier. Representing personality as a truly continuous variable where each person can, technically, receive a score that is unique to them is a huge advantage of language models and that advantage is retained when moving to this ensemble method.

Interview Format

This format could be useful for an interview. Providing applicants the choice of prompt to respond to and allowing them to explain themselves will produce a larger variety of responses and might even produce more words. Leveraging the choice itself to help estimate a personality trait, while providing for a more dynamic interview, could prove very useful.

Is the Forced Choice Language Model Format Cognitively Taxing?

A challenge to this format is that it might be more cognitively taxing than a traditional self-report method. This could produce adverse impact or add a confound to the results. Shortening the survey, making it more engaging (e.g., chatbot), and testing user reactions will help ameliorate this concern.

Ground Truth

Another challenge is the ground truth that the models get trained on. The relationships with outcomes shown in the MTMM table highlights this problem. The NLP model performed similarly to what it was trained on (CCS), but performed worse in quantity of significant relationships and magnitude of effect. The ensemble model performed similarly to CCS and

NLP, but worse again yet in a similar way. Each successive model that was trained on the CCS performed worse than the model before it. The model is trying to represent the numbers its trained on and so, necessarily, it will be worse than the model it's modeling. The ensemble process might be exacerbating that. Improving the ground truth these models get trained on is critical for creating a better model of personality.

Word Count

The word count for the responses was ~900 words per person, which exceeds the 500 word per person recommendation to produce stable estimates (Li et al., 2017). However, participants are not responding to a question, they are explaining a choice. So, that could sway the results and produce responses which are not directly related to the trait of interest and more related to decision. This doesn't appear to be the case, for the supplementary research question basic topic modeling was performed and it appears that participants are producing responses which are related to conscientiousness.

Forced Choice vs. Narrative Basis of Personality

The FC format is more similar to a traditional self-report method of measuring personality. Namely, participants are asked to match their trait level the way they see it to the item presented to them in a self-report fashion identical to traditional scales. The narrative format is a little different. It is closer to a peer-report measure of personality, where participants provide a narrative which is evaluated by the chatbot. Including an others-reported personality score into future versions of this could provide more insight into this.

Concluding Remarks

The primary purpose of this research was to gather psychometric evidence for the validity of using an ensemble approach to measure personality. As previously discussed, the findings of

this research are subject to several limitations. Nevertheless, this study represents a significant initial effort towards enhancing the measurement of personality. Personality is an intimate and important topic. While the current design and results present considerable challenges and may be too cumbersome to represent an improvement over existing methods, this research serves effectively as a prototype.

This study proposes to expand upon the EFCNLP (Ensemble of Forced Choice and Natural Language Processing) method. There are many important considerations for generating scores for both the ensemble and text-derived scores that haven't received much research attention and were beyond of the scope of this thesis. These include preprocessing procedures, embedding methods, training methods (for both the ensemble and the language model), what the models are trained on (the ground truth problem), and adverse impact considerations of using machine learning.

The immediate next step is to shorten the length of the survey. Many participants didn't finish the study is most likely because of the length of survey, leading to the high dropout rate. Expanding beyond just conscientiousness and using the Big 5 should help solve some of the issues that arose with the FC scoring. Using mixed extremities, as previously described, should also help.

After collecting data with the updated design, the next step will be to explore methods of ensembling. Linear regression is a simple and effective method, however, examining alternatives and evaluating other model performance will help guide future researchers interested in ensembling. Additionally, exploring other methods of constructing language models and gathering evidence for which are optimal will, similarly, help guide future research.

Each of these next steps are important, but they are focused on optimizing model performance with the information gathered. However, if this method is going to be used in the workplace, adverse impact needs to be considered. Bias in machine learning and AI has been a hot topic. Machine learning models are only as biased as what they are modeling. If the ground truth is biased, then the model will be biased. One of the major advantages of these ML models is that they are scalable. Once developed, an assessment like this one can be administered to a lot more people without an arithmetic increase in cost. With that scaling comes extra risk of harm, the magnitude of the effect of biased model will be larger, even if it's smaller at the individual level. The adverse impact of these models needs to be closely monitored. When adverse impact is detected, it can only be due to real group differences or a biased measurement tool. Models describe the world as it is, not as it ought to be. In other words, it's the responsibility of researchers to ensure the measurement tool is modeling the world as accurately as possible. It's the responsibility of practitioners and policy-makers (and all citizens, really) to make the world less biased, which will make future models less biased. Regardless, monitoring adverse impact and transparency are essential components of building a ML model. Exploring effective methods of doing exactly that is also a part of my future research plan.

This thesis is a prototype of this ENLPFC (Ensemble of Natural Language Processing and Forced Choice) method of modeling personality. Much work remains to be done and this thesis can serve as a guide.

References

Ayoub, M., Gosling, S. D., Potter, J., Shanahan, M. J., & Roberts, B. W. (2018). The relations between parental socioeconomic status, personality, and life outcomes. *Social Psychological & Personality Science*, *9*(3), 338–352. https://doi.org/10.1177/1948550617707018

Barrick, M. R., & Mount, M. K. (1991). THE BIG FIVE PERSONALITY DIMENSIONS AND JOB PERFORMANCE: a META-ANALYSIS. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Brown, A. (2015). Personality Assessment, Forced-Choice. In *Elsevier eBooks* (pp. 840–848). https://doi.org/10.1016/b978-0-08-097086-8.25084-8

Brown, A. (2016). Item Response Models for Forced-Choice questionnaires: A Common Framework. *Psychometrika*, *81*(1), 135–160. https://doi.org/10.1007/s11336-014-9434-9

Cao, M., & Drasgow, F. (2019). Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *Journal of Applied Psychology*, *104*(11), 1347–1368. https://doi.org/10.1037/apl0000414

Cheng, W., & Ickes, W. (2009). Conscientiousness and self-motivation as mutually compensatory predictors of university-level GPA. *Personality and Individual Differences*, *47*(8), 817–822. https://doi.org/10.1016/j.paid.2009.06.029

Chernyshenko, O. S. (2002). Applications of ideal point approaches to scale construction and scoring in personality measurement: The development of a six-faceted measure of conscientiousness. *UIUC Dissertaions*. http://psycnet.apa.org/psycinfo/2003-95010-007

Chernyshenko, O. S., Stark, S., Drasgow, F., & Roberts, B. W. (2007). Constructing personality scales under the assumptions of an ideal point response process: Toward increasing the flexibility of personality measures. *Psychological Assessment*, *19*(1), 88–106. https://doi.org/10.1037/1040-3590.19.1.88

Christiansen, N. D., Burns, G. N., & Montgomery, G. E. (2005). Reconsidering Forced-Choice item formats for applicant personality assessment. *Human Performance*, *18*(3), 267–307. https://doi.org/10.1207/s15327043hup1803_4

Cox, K. L. (2009). The South and Mass Culture. *The Journal of Souther History*, *75*(3), 677. https://www.jstor.org/stable/info/27779032

Dalal, D. K., Carter, N. T., & Lake, C. J. (2013). Middle Response Scale Options are Inappropriate for Ideal Point Scales. *Journal of Business and Psychology*, *29*(3), 463–478. https://doi.org/10.1007/s10869-013-9326-5

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv (Cornell University). https://arxiv.org/pdf/1810.04805v2

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment*, *49*(1), 71–75. https://doi.org/10.1207/s15327752jpa4901_13

Drasgow, F., Chernyshenko, O. S., & Stark, S. (2010). Improving the measurement of psychological variables: ideal point models rock! *Industrial and Organizational Psychology*, *3*(4), 515–520. https://doi.org/10.1111/j.1754-9434.2010.01284.x

Drasgow, F., Stark, S., Chernyshenko, O. S., Nye, C. D., Hulin, C. L., & White, L. A. (2012). Development of the Tailored Adaptive Personality Assessment System (TAPAS) to support army personnel selection and classification decisions [Dataset]. In *PsycEXTRA Dataset*. https://doi.org/10.1037/e559902013-001

Falk, C. F., & Cai, L. (2016). A flexible full-information approach to the modeling of response styles. *Psychological Methods*, *21*(3), 328–347. https://doi.org/10.1037/met0000059

Fan, J., Sun, T., Liu, J., Zhao, T., Zhang, B., Chen, Z., Glorioso, M., & Hack, E. (2023). How well can an AI chatbot infer personality? Examining psychometric properties of machine-inferred personality scores. *Journal of Applied Psychology*, *108*(8), 1277–1299. https://doi.org/10.1037/apl0001082

Farmer, S. M., & Van Dyne, L. (2010). The idealized self and the situated self as predictors of employee work behaviors. *Journal of Applied Psychology*, *95*(3), 503–516. https://doi.org/10.1037/a0019149

Goffin, R. D., & Boyd, A. (2009). Faking and personality assessment in personnel selection: Advancing models of faking. *Canadian Psychology*, *50*(3), 151–160. https://doi.org/10.1037/a0015946

Goldberg, L. R. (1992). Possible questionnaire format for administering the 50-item set of IPIP Big-Five factor markers. *Psychol. Assess*, *4*, 26-42.

Golfried, M. R., & Kent, R. N. (1972). Traditional versus behavioral personality assessment: A comparison of methodological and theoretical assumptions. *Psychological Bulletin, 77*(6), 409–420. https://doi.org/10.1037/h0032714

Gregory, R. L. (1972). HOW WE SEE AND HEAR THINGS. *Journal of the Royal Society of Arts*, 120(5193), 630–635. http://www.jstor.org/stable/41370916

Grün, B., & Hornik, K. (2011). topicmodels: AnRPackage for Fitting Topic Models. *Journal of Statistical Software*, *40*(13). https://doi.org/10.18637/jss.v040.i13

Harris, J. R. (2000). Socialization, personality development, and the child's environments: Comment on Vandell (2000). *Developmental Psychology*, *36*(6), 711–723. https://doi.org/10.1037/0012-1649.36.6.711

Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior*, *32*, 113–135. https://doi.org/10.1016/j.riob.2012.11.003

Heron, X., Sylvara, A., & Sun, T. (2024)*. Can ChatGPT Trick Automated Interviews? Evaluating the Use of AI-assistance to Fake on an AI-Chatbot Personality Interview. In B. Connelly, [Advancing Personality Assessment: Leveraging Classic Theories and Modern Measures]. Symposium. *Presented at the 39th Annual Conference of the Society for Industrial and Organizational Psychology,* Chicago, IL.

Hill, P. L., & Roberts, B. W. (2011). The role of adherence in the relationship between conscientiousness and perceived health. *Health Psychology*, *30*(6), 797–804. https://doi.org/10.1037/a0023860

Hofstede, G., & McCrae, R. R. (2004). Personality and culture revisited: linking traits and dimensions of culture. *Cross-cultural Research*, *38*(1), 52–88. https://doi.org/10.1177/1069397103259443

Jones, E., & Wessely, S. (2006). Psychological trauma: a historical perspective. *Psychiatry*, *5*(7), 217–220. https://doi.org/10.1053/j.mppsy.2006.04.011

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). THE BIG FIVE PERSONALITY TRAITS, GENERAL MENTAL ABILITY, AND CAREER SUCCESS ACROSS THE LIFE SPAN. *Personnel Psychology*, *52*(3), 621–652. https://doi.org/10.1111/j.1744-6570.1999.tb00174.x

Judge, T. A., Higgins, C. A., Thoresen, C. J., & Barrick, M. R. (1999). The big five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, *52*(3), 621–652.

Kjell, O., Giorgi, S., & Schwartz, H. A. (2023). The text-package: An R-package for analyzing and visualizing human language using natural language processing and transformers. *Psychological Methods*, *28*(6), 1478–1498. https://doi.org/10.1037/met0000542

Lance, C. E., Lance, C. E., & Vandenberg, R. J. (2010). The partial revival of a dead horse? Comparing classical test theory and item response theory. *Statistical and methodological myths and urban legends* (pp. 57-80).

Li, J., Zhou, M. X., Yang, H., & Mark, G. (2017, March). Confiding in and listening to virtual agents: The effect of personality [Conference session]. Paper presented at the 22nd annual meeting of the intelligent user interfaces community, Limassol, Cyprus. https://doi.org/10.1145/3025171.3025206

Li, M., Sun, T., & Zhang, B. (2022). AutoFC: an R package for automatic item pairing in Forced-Choice Test construction. *Applied Psychological Measurement*, *46*(1), 70–72. https://doi.org/10.1177/01466216211051726

Li, M., Zhang, B., Li, L., Sun, T., & Brown, A. (2024). Mixed-Keying or Desirability-Matching in the construction of Forced-Choice measures? An empirical investigation and practical recommendations. *Organizational Research Methods*. https://doi.org/10.1177/10944281241229784

Li, W., Wu, C., Hu, X., Chen, J., Fu, S., Wang, F., & Zhang, D. (2022). Quantitative personality predictions from a brief EEG recording. *IEEE Transactions on Affective Computing*, *13*(3), 1514–1527. https://doi.org/10.1109/taffc.2020.3008775

Lin, Y. (2021). Reliability estimates for IRT-Based Forced-Choice assessment scores. *Organizational Research Methods*, *25*(3), 575–590. https://doi.org/10.1177/1094428121999086

McCrae, R. R. (2002). Cross-Cultural research on the Five-Factor model of personality. *Online Readings in Psychology and Culture*, *4*(4). https://doi.org/10.9707/2307-0919.1038

McCrae, R. R., Costa, P. T., Ostendorf, F., Angleitner, A., HřebíČková, M., Avia, M. D., Fernández, J. S., Sánchez-Bernardos, M. L., Kusdil, M., Woodfield, R., Saunders, P. R., & Smith, P. B. (2000). Nature over nurture: Temperament, personality, and life span development. *Journal of Personality and Social Psychology*, *78*(1), 173–186. https://doi.org/10.1037/0022-3514.78.1.173

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007a). RECONSIDERING THE USE OF PERSONALITY TESTS IN PERSONNEL SELECTION CONTEXTS. *Personnel Psychology*, *60*(3), 683–729. https://doi.org/10.1111/j.1744-6570.2007.00089.x

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K. R., & Schmitt, N. (2007b). ARE WE GETTING FOOLED AGAIN? COMING TO TERMS WITH LIMITATIONS IN THE USE OF PERSONALITY TESTS FOR PERSONNEL SELECTION. *Personnel Psychology*, *60*(4), 1029–1049. https://doi.org/10.1111/j.1744-6570.2007.00100.x

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

Orrù, G., Monaro, M., Conversano, C., Gemignani, A., & Sartori, G. (2020). Machine learning in psychometrics and psychological research. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02970

Ozer, D. J., & Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annual Review of Psychology*, *57*(1), 401–421. https://doi.org/10.1146/annurev.psych.57.102904.190127

Ozer, D. J., & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology*, *45*(1), 357-388.

Penke, L., Denissen, J. J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of Personality*, *21*(5), 549–587. https://doi.org/10.1002/per.629

Roberts, B. W., Kuncel, N. R., Shiner, R. L., Caspi, A., & Goldberg, L. R. (2007). The Power of Personality: the comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, *2*(4), 313–345. https://doi.org/10.1111/j.1745-6916.2007.00047.x

Roberts, J. A., Pullig, C., & Manolis, C. (2015). I need my smartphone: A hierarchical model of personality and cell-phone addiction. *Personality and Individual Differences*, *79*, 13–19. https://doi.org/10.1016/j.paid.2015.01.049

Roberts, J. S., Donoghue, J., & Laughlin, J. E. (2000). A general item response theory model for unfolding unidimensional polytomous responses. *Applied Psychological Measurement*, *24*(1), 3–32. https://doi.org/10.1177/01466216000241001

Roulin, N., & Krings, F. (2020). Faking to fit in: Applicants' response strategies to match organizational culture. *Journal of Applied Psychology*, *105*(2), 130–145. https://doi.org/10.1037/apl0000431

Russell, S. S., Spitzmüller, C., Lin, L., Stanton, J. M., Smith, P. C., & Ironson, G. (2004). Shorter can Also be Better: The Abridged Job in General Scale. *Educational and Psychological Measurement*, *64*(5), 878–893. https://doi.org/10.1177/0013164404264841

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262–274. https://doi.org/10.1037/0033-2909.124.2.262

Soto, C. J., & John, O. P. (2016, April 7). The Next Big Five Inventory (BFI-2): Developing and Assessing a Hierarchical Model With 15 Facets to Enhance Bandwidth, Fidelity, and Predictive Power. *Journal of Personality and Social Psychology*. Advance online publication. http:// dx.doi.org/10.1037/pspp0000096

Spector, P. E., Bauer, J. A., & Fox, S. (2010). Measurement artifacts in the assessment of counterproductive work behavior and organizational citizenship behavior: Do we know what we think we know? *Journal of Applied Psychology*, *95*(4), 781–790. https://doi.org/10.1037/a0019477

Sun, T. (2021). Artificial intelligence powered personality assessment: A multidimensional psychometric natural language processing perspective (Doctoral dissertation, University of Illinois at Urbana-Champaign).

Sun, T., Bo, Z., Phan, W. M. J., Drasgow, F., & Roberts, B. W. (2019). “Meh!”: Examining Midpoint Endorsement habitude (MEH) in survey research. *Proceedings - Academy of Management*, *2019*(1), 16421. https://doi.org/10.5465/ambpp.2019.227

Verdult, R. (2009). Caesarean Birth: Psychological Aspects in Adults. *International Journal of Prenatal and Perinatal Psychology and Medicine*, *21 (1/2)*, 17–28. https://www.stroeckenverdult.be/site/upload/docs/Isppm%20tijdschrift%20CAESAREAN%20BIRTH%20adults.pdf

Wang, M., Liao, H., Zhan, Y., & Shi, J. (2011). Daily Customer Mistreatment and Employee Sabotage Against Customers:Examining emotion and resource Perspectives. *Academy of Management Journal/˜ the œAcademy of Management Journal*, *54*(2), 312–334. https://doi.org/10.5465/amj.2011.60263093

Woehr, D. J., Putka, D. J., & Bowler, M. C. (2011). An examination of G-Theory methods for modeling Multitrait–Multimethod data. *Organizational Research Methods*, *15*(1), 134–161. https://doi.org/10.1177/1094428111408616

Wolf, E. J., Harrington, K., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models. *Educational and Psychological Measurement*, *73*(6), 913–934. https://doi.org/10.1177/0013164413495237

Zhang, B., Tu, N., Angrave, L., Zhang, S., Sun, T., Tay, L., & Li, J. (2023). The Generalized Thurstonian Unfolding Model (GTUM): advancing the modeling of Forced-Choice data. *Organizational Research Methods*. https://doi.org/10.1177/10944281231210481

Appendix A - All Items From Survey

TAPAS Items for Forced Choice (Drasgow et. al, 2012) (all 90 items are presented twice to each participant, once in a single statement and again in a forced choice pair. The order is counter balanced. The items are grouped by forced choice pairs)

Please indicate the degree to which you agree with the following statements on a 4-point scale. (1=Strongly disagree, 2=Disagree, 3=Agree, 4=Strongly agree)

1. I don't care very much about the quality of my work
2. Keeping a careful record of things is not my strength

3. If I break a rule I feel really guilty
4. People say I'm trustworthy

5. I find traditional family values to be very important
6. If someone is counting on me I never let them down

7. I don't think that being impulsive is a fault
8. I enjoy the process of doing things and don't care much about the results

9. I am only careful on tasks that are important to me
10. I think rules and regulations should be flexible

11. Being successful is more important than most other things in my life
12. Being clean helps me to focus

13. I can ignore a mess for a long time, but eventually I have to clean it up
14. I try to do the right thing, but it can be very difficult

15. I feel comfortable taking on some responsibility
16. I don't mind waiting for something better to come along

17. Sometimes I wish that everyone was as organized as me
18. I am usually cautious

19. I always want to be better than others in the things I do
20. Getting a job is important, but telling the truth in a job interview is more important

21. I try not to promise too much
22. I would never cheat, but don't mind helping someone else cheat

23. Laws and rules limit the development of human society
24. I frequently try to get others to do things that I am responsible for

25. I usually conform to an accepted standard of right and wrong

26. I always think twice before saying something

27. If I know someone is cheating on an exam, I will turn him/her in
28. When I set my mind on achieving a goal, I can always reach it

29. I don't question authority figures' judgments
30. I plan my time very carefully

31. I have a hard time meeting deadlines
32. Lying seems to make life a little easier

33. I often break rules
34. I can never find the things I want at home

35. Being messy helps my creativity
36. I don't mind if people say I'm irresponsible

37. I have a great sense of responsibility
38. Organizing and arranging things is extremely fulfilling

39. I usually follow a schedule, but occasionally I get bored and don't
40. To me, being moderately successful is enough

41. My mind wanders a lot when I'm working on something
42. I do whatever I want and don't worry about the rules

43. It bothers me a lot when my plans are disturbed
44. Following rules is important to keep everyone safe

45. It's hard for me to keep things in order
46. I prefer working in a team where I don't have a lot of responsibility

47. I only care about my own part of a group project
48. I do not blindly follow rules

49. When I have many things to do, I try to focus on the task with highest priority first
50. I arrive at work on time

51. I don't see a need to follow all the traditions my parents taught me
52. I prefer not to plan ahead and instead take life as it comes

53. People should not sacrifice too much for work
54. An impulsive decision isn't always bad

55. I think that part of being a good citizen is strictly following the law
56. There is too much to be done to waste time relaxing

57. I frequently make up believable excuses for not finishing my work
58. I don't care if I am late

59. I finish my work on time but try not to work more than I have to
60. I try to be a good person, but I sometimes do bad things

61. I follow all the advice my parents give me
62. I am always honest with the people in my life

63. I often make careless mistakes
64. Rules don't dictate my behavior

65. I can keep my concentration only on short tasks
66. Being honest with others is not always good

67. Organizing things is a waste of time
68. I try to do the minimal amount of work possible to maintain my current status

69. I follow a strict daily schedule
70. I don't procrastinate

71. I generally finish my work but it's not always on time
72. I don't usually think before I talk

73. I am competitive and play to win
74. Rarely I do something dishonest

75. I tend to do just what is expected of me when doing a job
76. I try to keep track of my bills, but I'm not too accurate

77. I hardly ever finish the tasks I start
78. Taking on responsibility is just asking for trouble

79. I like to help others, but sometimes I am just too busy to help
80. I try to be punctual but sometimes things get in my way

81. If I am interested in something I don't mind working hard
82. For a better life, it is necessary to have rules and laws

83. I always try to do my best work even when no one will know
84. I make plans if I have enough time

85. I am average at the things I do
86. If a rule doesn't make sense to me I won't follow it

87. I don't really care about being successful
88. I don't feel bad about exaggerating my contribution when working on a group project

89. I like to meet important deadlines in my work, but don't always feel the need
90. I'm about as law-abiding as the average person

Chernyshenko Conscientiousness Scale Items (Cherynshenko, 2002)

Please indicate the degree to which you agree with the following statements on a 4-point scale. (1=Strongly disagree, 2=Disagree, 3=Agree, 4=Strongly agree)

Order

1. (R) Being neat is not exactly my strength.
2. Organization is a key component of most things I do.
3. I need a neat environment in order to work well.
4. I become annoyed when things around me are disorganized.
5. (R) For me, being organized is unimportant.
6. (R) Half of the time I do not put things in their proper place.
7. (R) Most of the time my room is in complete disarray.
8. Every item in my room and on my desk has its own designated place.
9. (R) I frequently forget to put things back in their proper place.
10. I hate when people are sloppy

Virtue

11. (R) If I could get away with it, I would not pay taxes.
12. (R) I would lie without hesitation if it serves my purpose.
13. (R) I could be insincere and dishonest if the situation required me to do so.
14. (R) If I find money laying around, I'll keep it to myself
15. If a cashier forgot to charge me for an item I would tell him/her.
16. I would rather get a bad grade than copy someone else's homework and turn it in as my own.
17. It bothers me when people cheat on their taxes.
18. If I accidentally scratched a parked car, I would try to find the owner to pay for the repairs.
19. I firmly believe that under no circumstances it is okay to lie.
20. The people who know me best would say that I am honest.

Traditionalism

21. I have the highest respect for authorities and assist them whenever I can.
22. (R) People respect authority more than they should.
23. Even if I knew how to get around the rules without breaking them, I would not do it.
24. (R) I believe that people should be allowed to take drugs, as long as it doesn't affect others.
25. I support long-established rules and traditions.
26. People who resist authority should be severely punished.
27. (R) When I was in school, I used to break rules quite regularly.
28. In my opinion, all laws should be strictly enforced.
29. (R) In my opinion, censorship slows down progress.
30. When working with others I am the one who makes sure that rules are observed.

Self-Control

31. (R) I often rush into action without thinking about potential consequences.
32. I rarely jump into something without first thinking about it.
33. (R) I am known to make quick, hot-headed decisions.

34. I do not take unnecessary risks.
35. (R) I am easily talked into doing silly things.
36. (R) My friends say I am unpredictable.
37. (R) I get into trouble because I act on impulses rather than on thoughts.
38. I am careful with what I say to others.
39. I dislike being around impulsive people.
40. Even under time pressure, I would rather take my time to think about my answer than to say the first thing that comes to mind.

Responsibility

41. I carry out my obligations to the best of my ability.
42. I often feel responsible for making sure that all group project assignments are completed.
43. I go out of my way to keep my promises.
44. (R) Sometimes it is too much of a bother to do exactly what is promised.
45. I would gladly spend some of my leisure time trying to improve my community.
46. (R) If I am running late to an appointment, I may decide not to go at all.
47. (R) I am usually not the most responsible group member, but I will not shirk on my duties either.
48. If I am running late, I try to call ahead to notify those who are waiting for me.
49. (R) When I make mistakes I often blame others.
50. (R) I have a reputation for being late for almost every meeting or event.

Industriousness

51. I have high standards and work toward them.
52. I go above and beyond of what is required.
53. (R) I do not work as hard as the majority of people around me.
54. (R) I invest little effort into my work.
55. I demand the highest quality in everything I do.
56. I try to be the best at anything I do.
57. I make every effort to do more than what is expected of me.
58. (R) I do what is required, but rarely anything more.
59. (R) Setting goals and achieving them is not very important to me.
60. (R) Getting average grades is enough for me.

Subjective well-being (Diener, Emmons, Larsen, & Griffin, 1985)

Please indicate the degree to which you agree with the following statements on a 5-point scale. (1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly agree)

1. In most ways my life is close to my ideal
2. The conditions of my life are excellent
3. I am satisfied with my life
4. So far I have gotten the important things I want in life
5. If I could live my life over, I would change almost nothing

Organizational Citizenship Behavior (Spector, Bauer, & Fox, 2010)

Recall that in the past year, how often have you done the following things? (1=Never, 2=Once or twice, 3=Once or twice per month, 4= Once or twice per week, 5=Every day)

1. Took time to advise, coach, or mentor a co-worker.

2. Helped co-worker learn new skills or shared job knowledge.
3. Helped new employees get oriented to the job.
4. Lent a compassionate ear when someone had a work problem.
5. Offered suggestions to improve how work is done.
6. Helped a co-worker who had too much to do.
7. Volunteered for extra work assignments.
8. Worked weekends or other days off to complete a project or task.
9. Volunteered to attend meetings or work on committees on own time.
10. Gave up meal and other breaks to complete work.

Counterproductive work behavior (Spector, Bauer, & Fox, 2010)
Recall that in the past year, how often have you done the following things?
(1=Never, 2=Once or twice, 3=Once or twice per month, 4= Once or twice per week, 5=Every day)
1. Purposely wasted your employer's materials/supplies
2. Complained about insignificant things at work
3. Told people outside the job what a lousy place you work for
4. Came to work late without permission
5. Stayed home from work and said you were sick when you weren't
6. Insulted someone about their job performance
7. Made fun of someone's personal life
8. Ignored someone at work
9. Started an argument with someone at work
10. Insulted or made fun of someone at work

Job satisfaction in general (AJIG; Russell, Spitzmuller, Lin, Stanton, Smith, & Ironson, 2004)
Think of your current school or professional work in general. All in all, what is it like most of the time?
(1=Yes, it describes my work, 2=Cannot decide, 3=No, it does not describe my work)
1. Good
2. Undesirable
3. Better than most
4. Disagreeable
5. Make me content
6. Excellent
7. Enjoyable
8. Poor

Telework (Park, & Cho, 2020)
How often do you telework? (never; 1 day per week; 2–3 days per week; 4–5 days per week)

Emotion-Regulation Self-Efficacy (Wang et al., 2011, AMJ)
1. I am able to control my temper and handle difficulties rationally.
2. I am quite capable of controlling my own emotions.
3. I can always calm down quickly when I am very angry.

4. I have good control of my own emotions.

Identity with the Job (adapted from Farmer & van Dyne, 2010)

1. My current job is really important to me
2. My current job is a really important part of whom I am
3. I have strong positive feelings about working on my current job

Overall job performance

(1=Strongly disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly agree)

1. I am very competent at what I do.
2. I get my tasks done very effectively.
3. I perform my responsibilities well.

Overall health

In general, how would you describe your health?

1=Excellent
2=Very good
3=Good
4=Fair
5=Poor

BMI info

How tall are you (in ft and in)?
How much do you weigh (in lbs)?

English communication ability

I am a native English speaker.

1=Yes
2=No, my first language is: _____

Please indicate your level of comfort reading in the English language.

1=Very comfortable
2=Comfortable
3=A little comfortable
4=Neutral
5=A little uncomfortable
6=Uncomfortable
7=Very uncomfortable

Please indicate your level of comfort writing in the English language.

1=Very comfortable
2=Comfortable
3=A little comfortable
4=Neutral
5=A little uncomfortable
6=Uncomfortable
7=Very uncomfortable

GPA info
What is your GPA from your most recently attended academic institute?
What is the maximum GPA a student can get at your most recently attended academic institute?

Organizational status
In your current position at school/work, do you have the right to do the following things?
(1=Yes, 0=No)
1. Hire people
2. Fire people
3. Supervise people
4. Create budgets for the organization
5. Make strategic decisions for the organization

Estimated Hourly Pay: $____ per hour

Estimated Yearly Income: ___

Typical Work Hours

1. How many jobs do you have?
2. How many hours on average do you work per week? Please consider all your jobs if you have multiple ones.

Demographic Questionnaire
1. What is your current employment status?
    1=Employed full time
    2=Employed part time
    3=Unemployed and currently looking for work
    4=Unemployed and not currently looking for work
    5=Student
    6=Retired
    7=Homemaker
    8=Self-employed
    9=Unable to work
2. What is your current household income (pre-tax) in U.S. dollars?
    1=Under $10,000
    2=$10,000 - $19,999
    3=$20,000 - $34,999
    4=$35,000 - $49,999
    5=$50,000 - $74,999
    6=$75,000 - $99,999
    7=$100,000 - $150,000
    8=Over $150,000

3. How many **months** of work experience do you have in total (including internship and volunteering)? ___

4. Please indicate your occupation (choose all that apply):

1=Management, professional, and related
2=Service
3=Sales and office
4=Farming, fishing, and forestry
5=Construction, extraction, and maintenance
6=Production, transportation, and material moving
7=Government
8=Retired
9=Unemployed

5. What is your current marital status?

1=Single
2=Married
3=Living with partner
4=Separated
5=Divorced
6=Widowed

6.What is the highest level of education you have completed?

1=Less than a high school diploma
2=High school degree or equivalent
3=Some college, no degree
4=Associate degree
5=Bachelor's degree
6=Master's degree
7=Professional degree
8=Doctorate

7. How old are you?
8. Do you reside in the US?
9. Your sex (check one): Female Male Prefer not to answer
10 Your race/ethnicity (check all that apply)

African American/Black □ Asian □ Caucasian □ Hispanic □
Native American/Alaskan Native □ Native Hawaiian/Pacific Islander □
Middle Eastern/North African □ Race Not Listed □

11. How long have you been working? _____years
12. How long have you been working in your current organization? _____years
13. Which state do you currently reside in? _____ (Drop-down menu of all states)
14. Are you currently a student? Yes/no.
14. For the highest education you have completed please estimate your GPA. _____
15. I am satisfied with my job. 1-5.
16. I want to find a new job. 1-5.
17. I am or would be a good remote worker. 1-5.
18. I make enough money. 1-5.
19. I have a lot of friends. 1-5.
20. Are your biological parents currently married? Yes/never married/divorced/unknown.

20. What is the longest romantic relationship you've maintained? ___ years (round accordingly)
21. I prefer serious relationships to casual relationships. 1-5.
22. I have kids. Yes/no.
    (if no): I want to have kids in the future. Yes/no.
    (if yes): How many? __
23. How often do you workout? ___ times per week
24. I have a plan for the future. 1-5.
25. I go to church. 1-5.
26. How many siblings do you have? ___

Overall Feedback

1. For the portion of the survey where you were asked to provide text explanations, please provide your feedback. Was it better or worse than a 4-point or 5-point Likert scale?
2. Do you have any other feedback you'd like to provide to the researchers?

Appendix B - Supplementary Research Question

Introduction

The purpose of this supplementary section is to provide the results of an exploratory analysis and discuss some implications of the results. As discussed earlier, an important aspect of FC is the underlying assumption of Thurstone's Law of Comparative judgement. The local independence assumption in Item Response Theory (IRT) assumes a causal relationship between statement endorsement and trait level (Brown, 2016). However, if the presence of the second item introduces a conditional probability, then scoring FC with IRT would be inappropriate. Understanding the decision-making process underlying FC could provide insight into this relationship and, while not conclusive, could suggest whether FC might, necessarily, violate the local independence assumption. The design for this research lent itself very well to exploring that question since participants are providing a rationale for their decision. The supplementary research question (RQ3) is **what are some reasons participants provide for endorsing one item over another in the FC block?** To answer this, some basic topic modeling will be performed. If the topics are trait relevant, then that would suggest local independence is not being violated. If the topics are more about comparisons, then that would suggest perhaps local independence might be violated, which would warrant a formal study of this question.

Local Independence

Analysis using Item Response Theory (IRT) assumes endorsement of an item (beta) is a direct relationship between the underlying trait level (theta) and item (i.e., local independence). This is assumed for FC blocks as well. However, it could be argued that the probability of preferring one item to the other is actually a joint probability, which would violate the independence assumption.

Analytical Procedures

The materials for this supplementary analysis are the text responses to the forced blocks. Latent Dirichlet Allocation (LDA) analysis is a simple and effective method of identifying topics that make a document. It assumes that a corpus of words (the response text, in this case) is made up of topics, which are made up of words. LDA will divide the corpus of text up into the number of topics requested and show the words or phrases that appear the most in each topic.

Before conducting a Latent Dirichlet Allocation (LDA) analysis, the data was cleaned and processed to ensure that words were comparable and content-irrelevant words (e.g., stop words) were removed. All characters were converted to lower case, all punctuation was removed, and all stop words were removed. The LDA was conducted on the response text using the *topicmodels* package in R (Grün & Hornik 2011). A bigram LDAs was conducted to identify the most used pairs of words.

Since there are six facets of conscientiousness in the present study, $k = 6$ topics were designated to be extracted. Theoretically, since participants are responding to facet-level statements, it makes sense that the topics would divide out by facet (with some overlap since they are correlated). The top 5 phrases from the six topics for the bigrams are displayed below.

Results

Appendix Figure B.1. LDA Topic Modeling Results (Bigrams; *k* = 6 topics)

Discussion of Results

Each of the topics appear to map onto conscientiousness, which suggests that responses were trait relevant. However, the topics don't appear to cleanly map onto any given facet. The facets are all part of conscientiousness and correlated, which could be the cause of them not being distinct. I expected the words in one topic to be all related to honesty/virtue, another to be all related to responsibility/duty, another to be all related to industriousness/hard work, another

to be all related to order/organizational related things, another about self-control/impulsivity, and the last topic to be related to traditionalism/family values. There was a little bit of that, especially for the first phrase in each topic, but also a lot of overlap across topics. This could be that the facets are correlated pretty closely, so grouping them by topic is naturally going to be a little messy.

While not conclusive from this analysis, the results of the LDA suggest that scoring FC with IRT does not violate local independence. Each of these topics contain only words that are trait related. It appears the items themselves are causing the responses provided, and not the presence of a second statement. While this evidence isn't conclusive, it indicates using IRT to score FC assessments is appropriate.

Future Directions

To conclusively answer the question of whether the forced choice format violates the local independence assumption in Item Response Theory, further research would need to be conducted. FC scores, historically, don't struggle with validity concerns. Perhaps the concern as to whether local independence is technically violated isn't practically important, since it works. However, it is valuable to understand the decision-making process, regardless. Future iterations of this line of research will continue to investigate this question.

ur appendix content goes here.

Appendix C - R Code Used For Analyses

```
#calculating reliability for the TAPAS, CCS scale, and NLP-R
library("psych")
alpha(TAPAS_SS_Data, check.keys=TRUE)
alpha(CCS, check.keys=TRUE)

##########################################################
##########################################################
#Getting embeddings
##########################################################
##########################################################

#text library
library(text)

#read data in
data = read.csv("/Users/xavierheron/Desktop/PhD Program/Lab/NLP & FC Project/Master's
Project/R Scripts for Analysis/Data Used/202402029Cleaned_All.xlsx",
              header = T)
#subset for response items
sub=X202402029Cleaned_All[c(25,27,29,31,33,35,37,39,41,43,45,47,49,51,53,55,57,59,61,63,6
5,67,69,71,73,75,77,
          79,81,83,85,87,89,91,93,95,97,99,101,103,105,107,109,111,113)]
#subset for response items and FC pairs
sub2=X202402029Cleaned_All[c(24:113)]

#word embeddings for just response
Masters<- textEmbed(
  sub,
  model = "bert-base-uncased",
  aggregation_from_layers_to_tokens = "concatenate",
  aggregation_from_tokens_to_texts = "mean",
  keep_token_embeddings = FALSE
)

# Save the word embeddings to avoid having to embed the
#text again.
saveRDS(Masters, "Masters.rds")

#word embeddings for responses and statements
Masters.statements.responses<- textEmbed(
  sub2,
  model = "bert-base-uncased",
  aggregation_from_layers_to_tokens = "concatenate",
  aggregation_from_tokens_to_texts = "mean",
```

```
  keep_token_embeddings = FALSE
)
saveRDS(Masters.statements.responses,"  .rds")



# Extract word type embeddings and text embeddings
MastersTypes <- textEmbed(
  texts = sub,
  aggregation_from_layers_to_tokens = "concatenate",
  aggregation_from_tokens_to_texts = "mean",
  aggregation_from_tokens_to_word_types = "mean",
  keep_token_embeddings = FALSE)
saveRDS(MastersTypes, "MastersTypes.rds")


library("text")

CCS.traits.facets=parameter_estimates_all[c(2,4,6,8,10,12,14)]
CCS.IRT.CTT=data.frame(CCS.traits.facets[c(1:7)],X202402029Cleaned_All$SC0,X20240202
9Cleaned_All$SC1, X202402029Cleaned_All$SC2,
X202402029Cleaned_All$SC3,X202402029Cleaned_All$SC4,X202402029Cleaned_All$SC5,X
202402029Cleaned_All$SC6)

#texttrainlist
Master_Model <- textTrainLists(
  Masters_embeddings$texts,
  CCS.traits.facets)
saveRDS(Master_Model,"Master_Model.rds")

#texttrainlist
Master_Model4.0 <- textTrainLists(
  Masters.statements.responses$texts,
  CCS.IRT.CTT)
saveRDS(Master_Model4.0,"Masters.statements.responses.rds")



#train on conscientiousness (CCS IRT Scores)
NLP_Conscientiousness_CCS <- textTrain(
  x = Masters_embeddings$texts,
  y = CCS.traits.facets$`Consc. Theta`,
  model_description = "Predicted Conscientiousenss, n=466, bert uncased."
)

#save the model
```

```
saveRDS(NLP_Conscientiousness_CCS,"NLP_Conscientiousness_CCS.rds")

#train on Order
NLP_Order_CCS <- textTrain(
  x = Masters_embeddings$texts,
  y = CCS.traits.facets$`Order Theta`,
  model_description = "Predicted Order, n=466, bert uncased."
)

#save the model
saveRDS(NLP_Order_CCS,"NLP_Order_CCS.rds")

#train on Virtue
NLP_Virtue_CCS <- textTrain(
  x = Masters_embeddings$texts,
  y = CCS.traits.facets$`Virtue Theta`,
  model_description = "Predicted Virtue, n=466, bert uncased."
)

#save the model
saveRDS(NLP_Virtue_CCS,"NLP_Virtue_CCS.rds")

#train on Traditionalism
NLP_Traditionalism_CCS <- textTrain(
  x = Masters_embeddings$texts,
  y = CCS.traits.facets$`Trad Theta`,
  model_description = "Predicted Traditionalism, n=466, bert uncased."
)

#save the model
saveRDS(NLP_Traditionalism_CCS,"NLP_Traditionalism_CCS.rds")

#train on Self-Control
NLP_Self-Control_CCS <- textTrain(
  x = Masters_embeddings$texts,
  y = CCS.traits.facets$`SC Theta`,
  model_description = "Predicted Self-Control, n=466, bert uncased."
)

#save the model
saveRDS(NLP_Self-Control_CCS,"NLP_Self-Control_CCS.rds")

#train on Responsibility
NLP_Responsibility_CCS <- textTrain(
  x = Masters_embeddings$texts,
  y = CCS.traits.facets$`Resp Theta`,
```

```
  model_description = "Predicted Responsibility, n=466, bert uncased."
)

#save the model
saveRDS(NLP_Responsibility_CCS,"NLP_Responsibility_CCS.rds")

#train on Industriousness
NLP_Industriousness_CCS <- textTrain(
  x = Masters_embeddings$texts,
  y = CCS.traits.facets$`Ind Theta`,
  model_description = "Predicted Industriousness, n=466, bert uncased."
)

#save the model
saveRDS(NLP_Industriousness_CCS,"NLP_Industriousness_CCS.rds")

##########################################################
##########################################################
#IRT Item Parameters for TAPAS
##########################################################
######### ################################################

library(dplyr)
library(GGUM)

#data file is "TAPAS SS Data.xlsx"
TAPAS=TAPAS_SS_Data
#ensuring all facets and trait are integers
TAPAS.in=apply(TAPAS, 2, as.integer)
Order.raw=TAPAS[c("C_ord_16","C_ord_7","C_ord_14","C_ord_8","C_ord_1","C_ord_15","C
_ord_5","C_ord_17",
              "C_ord_6","C_ord_13","C_ord_9","C_ord_18","C_ord_11","C_ord_3")]
Order.in=apply(Order.raw,2,as.integer)
Virtue.raw=TAPAS[c("C_vir_2","C_vir_9","C_vir_17","C_vir_8","C_vir_19","C_vir_12","C_v
ir_11","C_vir_13",
                "C_vir_14","C_vir_16","C_vir_18","C_vir_5")]
Virtue.in=apply(Virtue.raw,2,as.integer)
Trad.raw=TAPAS[c("C_trd_9","C_trd_4","C_trd_8","C_trd_13","C_trd_6","C_trd_11","C_trd_
12","C_trd_20",

"C_trd_18","C_trd_10","C_trd_19","C_trd_15","C_trd_7","C_trd_17","C_trd_1","C_trd_2",
              "C_trd_5")]
Trad.in=apply(Trad.raw,2,as.integer)
SC.raw=TAPAS[c("C_sec_5","C_sec_11","C_sec_19","C_sec_13","C_sec_8","C_sec_7","C_se
c_14","C_sec_6",
             "C_sec_9","C_sec_10","C_sec_15","C_sec_20")]
```

```
SC.in=apply(SC.raw,2,as.integer)
Resp.raw=TAPAS[c("C_res_6","C_res_4","C_res_7","C_res_9","C_res_13","C_res_11","C_res
_14","C_res_3",

"C_res_19","C_res_18","C_res_8","C_res_10","C_res_5","C_res_20","C_res_12","C_res_16",
            "C_res_1")]
Resp.in=apply(Resp.raw,2,as.integer)
Ind.raw=TAPAS[c("C_ind_9","C_ind_7","C_ind_8","C_ind_12","C_ind_14","C_ind_17","C_in
d_19","C_ind_13",

"C_ind_4","C_ind_5","C_ind_20","C_ind_1","C_ind_11","C_ind_10","C_ind_16","C_ind_15",
           "C_ind_3","C_ind_18")]
Ind.in=apply(Ind.raw,2,as.integer)

#getting parameters for each trait
fit1 <- GGUM(TAPAS.in, 4)
TAPAS_Con = Theta.EAP(fit1)
fit2 <- GGUM(Ind.in, 4)
TAPAS_Ind = Theta.EAP(fit2)
fit3 <- GGUM(SC.in, 4)
TAPAS_SC = Theta.EAP(fit3)
fit4 <- GGUM(Trad.in, 4)
TAPAS_Trad = Theta.EAP(fit4)
fit5 <- GGUM(Resp.in, 4)
TAPAS_Resp = Theta.EAP(fit5)
fit6 <- GGUM(Order.in, 4)
TAPAS_Order = Theta.EAP(fit6)
fit7 <- GGUM(Virtue.in, 4)
TAPAS_Virtue = Theta.EAP(fit7)


#Trim to just the essentials and export to .csv file
TAPAS_All_IRT_Scores <- bind_cols(
  list(TAPAS_Ind,
       TAPAS_SC,
       TAPAS_Trad,
       TAPAS_Resp,
       TAPAS_Order,
       TAPAS_Virtue,
       TAPAS_Con))
TAPAS_Traits_Facets = TAPAS_All_IRT_Scores[c(1:3,5,6,8,9,11,12,14,15,17,18,20,21)]
names(TAPAS_Traits_Facets) = c("Person","Ind Theta","Ind SE","SC Theta","SC SE","Trad
Theta","Trad SE",
                               "Resp Theta","Resp SE","Order Theta","Order SE","Virtue Theta","Virtue
SE",
                               "Consc. Theta","Consc. SE")
```

```
#write file to train model on
write.csv(x=TAPAS_Traits_Facets,file="TAPAS_Traits_Facets.csv")

#getting alphas and deltas for fcscoring
TAPAS_Consc_Alph_Delts = data.frame(fit1$alpha, fit1$delta)
write.csv(x=TAPAS_Consc_Alph_Delts,file="TAPAS_Consc_Alph_Delts.csv")


#fcscoring
#GTUM
install.packages("edstan")
install.packages("rstan")
install.packages("Rcpp")
devtools::install_github("Naidantu/fcscoring")
devtools::install_github("stan-dev/rstan/blob/develop/rstan/rstan/R/")
install.packages("rstantools")
install.packages("stanc_builder")

remove.packages(c("StanHeaders", "rstan"))
stan-dev/rstan/blob/develop/rstan/rstan/R/
  install.packages("rstan", repos = c("https://mc-stan.org/r-packages/", getOption("repos")))


library("edstan")
library("rstan")
library("Rcpp")
library("fcscoring")
library("rstantools")

options(mc.cores = parallel::detectCores())
rstan_options(auto_write = TRUE)

## Step 1: Read data
# 1.1 Response data in wide format for a 60-statement forced choice test measuring 5 trait from
50 respondents
gtum.Data <- fcpairchoice[c(2:46)]

# 1.2 A two-column matrix mapping each statement to each trait.
ind <- data.frame(statement1 = c(1,3,3,2,2,  1,      5,      4,      5,      1,      4,      3,
        6,      6,      3,      4,      3,      5,      4,
                        3,      2,      5,      5,      4,      5,      3,      1,      3,      1,
        1,      3,      2,      2,      5,      5,      4,
                        1,      1,      1,      6,      1,      1,      1,      1,      4),
            statement2 = c(2,4,      4,      1,      3,      5,      6,      2,      2,      6,
        6,      4,      2,      1,      5,      6,      5,      4,
```

```
                5,      1,      3,      3,      4,      3,      4,      5,      2,      1,
4,      6,      6,      3,      6,      1,      4,      2,
                6,      5,      4,      4,      3,      2,      3,      6,      3))
```

1.3 A two-column matrix containing initial values for the statement parameters alpha and delta block by block.
1 for alphas is recommended and -1/-2 or 1/2 for deltas are recommended depending on the signs of the statements.
#file used is "TAPAS_Consc_Alph_Delts.csv"
ParInits = matrix(c(alpha_deltas_by_item$alpha,alpha_deltas_by_item$delta),ncol=2)

Step 2: Fit the model1
mod9 <- gtum(gtum.Data=gtum.Data, ind=ind, block=2, ParIn its=ParInits, iter=5000)

Step 3: Extract the estimated results
# 13.1 Extract the theta estimates
theta <- extract(x=mod9, pars='theta')
# Turn the theta estimates into p*trait matrix where p equals sample size and trait equals the number of latent traits
theta <- theta[,1]
# nrow=trait
theta <- matrix(theta, nrow=6)
theta <- t(theta)
# theta estimates in p*trait matrix format
#first 10 respondents' thetas
theta[1:10,]
# 3.2 Extract the tau estimates
tau <- extract(x=mod9, pars='tau')
tau <- tau[,1]
tau
#3.3 Extract the estimates of the correlations among dimensions
cor <- extract(x=mod9, pars='cor')

Step 4: Plottings
# 4.1 Obtain the density plots for alphas of the first four statements
bayesplot(x=mod9, pars=paste0("alpha[",1:4,"]"), plot='density', inc_warmup=FALSE)
# 4.2 Obtain the trace plots for alphas of the first four statements
bayesplot(x=mod9, pars=paste0("alpha[",1:4,"]"), plot='trace', inc_warmup=FALSE)

saveRDS(mod9,file="mod9.rds")
write.csv(theta,file="FC.GTUM.Theta9.csv")
write.csv(cor,file="cor.among.dimensions.gtum9.csv")
write.csv(tau,file="tau.estimates.gtum9.csv")

```
#############################################################
#############################################################
#Correlation between predicted and actual score - NLP Model#
#############################################################
#############################################################

#############################################################
#############################################################
#Correlation between predicted and actual score - NLP Model#
#############################################################
#############################################################

####
#extracting predicted score
conscientiousness.predicted =
data.frame(Master_Model$predictions$`Response#1_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#2_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#3_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#4_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#5_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#6_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#7_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#8_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#9_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#10_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#11_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#12_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#13_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#14_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#15_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#16_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#17_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#18_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#19_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#20_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#21_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#22_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#23_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#24_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#25_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#26_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#27_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#28_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#29_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#30_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#31_Consc_CCS_Theta_pred`,
```

```
                        Master_Model$predictions$`Response#32_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#33_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#34_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#35_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#36_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#37_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#38_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#39_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#40_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#41_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#42_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#43_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#44_Consc_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#45_Consc_CCS_Theta_pred`
)
###
write.csv(conscientiousness.predicted, file="conscientiousness.predicted.csv")
#downloaded csv file, summed by row, added in actual score for correlation
cor(x=conscientiousness_predicted_actual$`Total Predicted
Score`,y=conscientiousness_predicted_actual$`Actual Score`)

#repeat for industriousness
industriousness.predicted =
data.frame(Master_Model$predictions$`Response#1_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#2_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#3_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#4_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#5_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#6_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#7_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#8_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#9_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#10_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#11_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#12_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#13_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#14_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#15_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#16_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#17_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#18_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#19_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#20_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#21_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#22_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#23_Ind_CCS_Theta_pred`,
```

```
                        Master_Model$predictions$`Response#24_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#25_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#26_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#27_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#28_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#29_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#30_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#31_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#32_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#33_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#34_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#35_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#36_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#37_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#38_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#39_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#40_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#41_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#42_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#43_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#44_Ind_CCS_Theta_pred`,
                        Master_Model$predictions$`Response#45_Ind_CCS_Theta_pred`
)
write.csv(industriousness.predicted, file="industriousness.predicted.csv")
cor(x=industriousness_predicted_actual$`Predicted
Score`,y=industriousness_predicted_actual$`Actual Score`)

cor(x=industriousness_predicted_actual$`Actual Score`, y= X202402029Cleaned_All$SC6)
cor(x=industriousness_predicted_actual$`Predicted Score`,y=X202402029Cleaned_All$SC6)

############################################################
############################################################
#Work to generating the NLP trait scores#
############################################################
############################################################


#Getting the conscientiousness and facet scores by predicting the variable it was trained on
ensemble_conscientiousness3 = lm(parameter_estimates_for_matrix$CCS_Conscientiousness
~a$'2'+
                        a$'4'+a$'6'+a$'8'+a$'10'+
                        a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                        a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                        a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                        a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                        a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
```

```
                        a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                        a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                        a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_conscientiousness3)
ensemble_conscientiousness_predicted_values3 <- predict(ensemble_conscientiousness3)
cor(ensemble_conscientiousness_predicted_values3,parameter_estimates_for_matrix$CCS_Con
scientiousness)
NLP_R_Conscientious = ensemble_conscientiousness_predicted_values3


#order (1/6)
a=b[c(1:90)]
ensemble_order = lm(parameter_estimates_for_matrix$CCS_Order ~a$'2'+
                        a$'4'+a$'6'+a$'8'+a$'10'+
                        a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                        a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                        a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                        a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                        a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                        a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                        a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                        a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_order)
ensemble_order_predicted_values <- predict(ensemble_order)
cor(ensemble_order_predicted_values,parameter_estimates_for_matrix$CCS_Order)
NLP_R_Order = ensemble_order_predicted_values


#Virtue (2/6)
a=b[c(91:180)]
new_column_names <- as.character(1:90)
colnames(a) <- new_column_names
#confirming right columns selected
a[1,1]
c[1,91]
NLP_virtue = lm(parameter_estimates_for_matrix$CCS_Virtue ~
                a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(NLP_virtue)
```

```
NLP_virtue_predicted_values <- predict(NLP_virtue)
cor(NLP_virtue_predicted_values,parameter_estimates_for_matrix$CCS_Virtue)
NLP_R_Virtue = NLP_virtue_predicted_values


#trad (3/6)
a=b[c(181:270)]
colnames(a) <- new_column_names
#confirming right columns selected
a[1,1]
c[1,181]
NLP_trad = lm(parameter_estimates_for_matrix$CCS_Traditionalism ~
                a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(NLP_trad)
NLP_trad_predicted_values <- predict(NLP_trad)
cor(NLP_trad_predicted_values,parameter_estimates_for_matrix$CCS_Traditionalism)
NLP_R_Tradition = NLP_trad_predicted_values


#self-control (4/6)
a=b[c(271:360)]
colnames(a) <- new_column_names
#confirming right columns selected
a[1,1]
c[1,271]
NLP_sc = lm(parameter_estimates_for_matrix$`CCS_Self-Control` ~
              a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
              a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
              a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
              a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
              a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
              a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
              a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
              a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
              a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(NLP_sc)
NLP_sc_predicted_values <- predict(NLP_sc)
cor(NLP_sc_predicted_values,parameter_estimates_for_matrix$`CCS_Self-Control`)
```

```
NLP_R_SControl = NLP_sc_predicted_values


#responsibility (5/6)
a=b[c(361:450)]
colnames(a) <- new_column_names
#confirming right columns selected
a[1,1]
c[1,361]
NLP_resp = lm(parameter_estimates_for_matrix$CCS_Responsibility ~
            a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
            a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
            a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
            a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
            a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
            a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
            a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
            a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
            a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(NLP_resp)
NLP_resp_predicted_values <- predict(NLP_resp)
cor(NLP_resp_predicted_values,parameter_estimates_for_matrix$CCS_Responsibility)
NLP_R_Responsibility = NLP_resp_predicted_values


#industry (6/6)
a=b[c(451:540)]
colnames(a) <- new_column_names
#confirming right columns selected
a[1,1]
c[1,451]
NLP_ind = lm(parameter_estimates_for_matrix$CCS_Industriousness ~
            a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
            a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
            a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
            a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
            a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
            a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
            a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
            a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
            a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(NLP_ind)
NLP_ind_predicted_values <- predict(NLP_ind)
cor(NLP_ind_predicted_values,parameter_estimates_for_matrix$CCS_Industriousness)
NLP_R_Industry = NLP_ind_predicted_values
```

```
#getting the ensemble scores for all by adding the FC text data and then the FC score
#Order (1/6)
a=d[c(1:90)]
colnames(a) <- new_column_names
#confirming right columns selected
a[1,2]
d[1,2]
#modeling the statement text
ensemble_order_fc_text = lm(parameter_estimates_for_matrix$CCS_Order ~
          a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
          a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
          a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
          a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
          a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
          a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
          a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
          a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
          a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_order_fc_text)
ensemble_order_fc_text_predicted_values <- predict(ensemble_order_fc_text)
cor(ensemble_order_fc_text_predicted_values,parameter_estimates_for_matrix$CCS_Order)
#0.6315937
#adding in the FC score as a predictor
ORDER_ensemble_model = lm(parameter_estimates_for_matrix$CCS_Order ~
                    ensemble_order_fc_text_predicted_values +
parameter_estimates_for_matrix$`GTUM Order` + NLP_R_Order)
summary(ORDER_ensemble_model)
ORDER_ensemble = predict(ORDER_ensemble_model)
cor(ORDER_ensemble,parameter_estimates_for_matrix$CCS_Order)
#0.701788


#Virtue (2/6)
a=d[c(91:180)]
colnames(a) <- new_column_names
#NLP model for the response text
ensemble_virtue_fc_text = lm(parameter_estimates_for_matrix$CCS_Virtue ~
                      a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                      a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                      a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                      a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                      a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                      a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                      a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                      a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
```

```
                    a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_virtue_fc_text)
ensemble_virtue_fc_text_predicted_values <- predict(ensemble_virtue_fc_text)
cor(ensemble_virtue_fc_text_predicted_values,parameter_estimates_for_matrix$CCS_Virtue)
#0.6013629
#adding in the FC score as a predictor
VIRTUE_ensemble_model = lm(parameter_estimates_for_matrix$CCS_Virtue ~
                    ensemble_virtue_fc_text_predicted_values +
parameter_estimates_for_matrix$`GTUM Virtue` + NLP_R_Virtue)
summary(VIRTUE_ensemble_model)
VIRTUE_ensemble = predict(VIRTUE_ensemble_model)
cor(VIRTUE_ensemble,parameter_estimates_for_matrix$CCS_Virtue)
#0.6477114


#Tradition (3/6)
a=d[c(181:270)]
colnames(a) <- new_column_names
#NLP model for the response text
ensemble_tradition_fc_text = lm(parameter_estimates_for_matrix$CCS_Traditionalism ~
                    a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                    a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                    a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                    a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                    a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                    a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                    a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                    a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                    a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_tradition_fc_text)
ensemble_tradition_fc_text_predicted_values <- predict(ensemble_tradition_fc_text)
cor(ensemble_tradition_fc_text_predicted_values,parameter_estimates_for_matrix$CCS_Traditi
onalism)
#0.5595471
#adding in the FC score as a predictor
TRADITION_ensemble_model = lm(parameter_estimates_for_matrix$CCS_Traditionalism ~
                    ensemble_tradition_fc_text_predicted_values +
parameter_estimates_for_matrix$`GTUM Traditionalism` + NLP_R_Tradition)
summary(TRADITION_ensemble_model)
TRADITION_ensemble = predict(TRADITION_ensemble_model)
cor(TRADITION_ensemble,parameter_estimates_for_matrix$CCS_Traditionalism)
#0.6812751


#Self-Control (4/6)
a=d[c(271:360)]
```

```
colnames(a) <- new_column_names
#NLP model for the response text
ensemble_sc_fc_text = lm(parameter_estimates_for_matrix$`CCS_Self-Control` ~
                           a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                           a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                           a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                           a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                           a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                           a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                           a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                           a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                           a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_sc_fc_text)
ensemble_sc_fc_text_predicted_values <- predict(ensemble_sc_fc_text)
cor(ensemble_sc_fc_text_predicted_values,parameter_estimates_for_matrix$`CCS_Self-
Control`)
#0.6090949
#adding in the FC score as a predictor
SC_ensemble_model = lm(parameter_estimates_for_matrix$`CCS_Self-Control` ~
                         ensemble_sc_fc_text_predicted_values +
parameter_estimates_for_matrix$`GTUM Self-Control`  + NLP_R_SControl)
summary(SC_ensemble_model)
SC_ensemble = predict(SC_ensemble_model)
cor(SC_ensemble,parameter_estimates_for_matrix$`CCS_Self-Control`)
#0.6787462


#Responsibility (5/6)
a=d[c(361:450)]
colnames(a) <- new_column_names
#NLP model for the response text
ensemble_resp_fc_text = lm(parameter_estimates_for_matrix$CCS_Responsibility ~
                         a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                         a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                         a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                         a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                         a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                         a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                         a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                         a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                         a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_resp_fc_text)
ensemble_resp_fc_text_predicted_values <- predict(ensemble_resp_fc_text)
cor(ensemble_resp_fc_text_predicted_values,parameter_estimates_for_matrix$CCS_Responsibil
ity)
#0.4992073
```

```
#adding in the FC score as a predictor
resp_ensemble_model = lm(parameter_estimates_for_matrix$CCS_Responsibility ~
                  ensemble_resp_fc_text_predicted_values +
parameter_estimates_for_matrix$`GTUM Responsibility` + NLP_R_Responsibility)
summary(resp_ensemble_model)
RESPONSIBILITY_ensemble = predict(resp_ensemble_model)
cor(RESPONSIBILITY_ensemble,parameter_estimates_for_matrix$CCS_Responsibility)
#0.6099472


#Industriousness (6/6)
a=d[c(451:540)]
colnames(a) <- new_column_names
#NLP model for the response text
ensemble_ind_fc_text = lm(parameter_estimates_for_matrix$CCS_Industriousness ~
                        a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                        a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
                        a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                        a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                        a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                        a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                        a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                        a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                        a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_ind_fc_text)
ensemble_ind_fc_text_predicted_values <- predict(ensemble_ind_fc_text)
cor(ensemble_ind_fc_text_predicted_values,parameter_estimates_for_matrix$CCS_Industriousn
ess)
#0.5139634
#adding in the FC score as a predictor
ind_ensemble_model = lm(parameter_estimates_for_matrix$CCS_Industriousness ~
                    ensemble_ind_fc_text_predicted_values +
parameter_estimates_for_matrix$`GTUM Industriousness` + NLP_R_Industry)
summary(ind_ensemble_model)
INDUSTRY_ensemble = predict(ind_ensemble_model)
cor(INDUSTRY_ensemble,parameter_estimates_for_matrix$CCS_Industriousness)
#0.6152031


#Conscientiousness
a=d[c(541:630)]
colnames(a) <- new_column_names
#NLP model for the response text
ensemble_con_fc_text = lm(parameter_estimates_for_matrix$CCS_Conscientiousness ~
                        a$'2'+a$'4'+a$'6'+a$'8'+a$'10'+
                        a$'12'+a$'14'+a$'16'+a$'18'+a$'20'+
```

```
                    a$'22'+a$'24'+a$'26'+a$'28'+a$'30'+
                    a$'32'+a$'34'+a$'36'+a$'38'+a$'40'+
                    a$'42'+a$'44'+a$'46'+a$'48'+a$'50'+
                    a$'52'+a$'54'+a$'56'+a$'58'+a$'60'+
                    a$'62'+a$'64'+a$'66'+a$'68'+a$'70'+
                    a$'72'+a$'74'+a$'76'+a$'78'+a$'80'+
                    a$'82'+a$'84'+a$'86'+a$'88'+a$'90')
summary(ensemble_con_fc_text)
ensemble_con_fc_text_predicted_values <- predict(ensemble_con_fc_text)
cor(ensemble_con_fc_text_predicted_values,parameter_estimates_for_matrix$CCS_Conscientio
usness)
#0.5276434
#adding in the FC score as a predictor
con_ensemble_model = lm(parameter_estimates_for_matrix$CCS_Conscientiousness~
                  ensemble_con_fc_text_predicted_values +
parameter_estimates_for_matrix$`GTUM Conscientiousness` + NLP_R_Conscientious)
summary(con_ensemble_model)
CONSCIENTIOUSNESS_ensemble = predict(con_ensemble_model)
cor(CONSCIENTIOUSNESS_ensemble,parameter_estimates_for_matrix$CCS_Conscientiousn
ess)
#0.6569908

#writing a file for each method for easier matrix viewing
ensemble=data.frame(CONSCIENTIOUSNESS_ensemble,
                ORDER_ensemble,
                VIRTUE_ensemble,
                TRADITION_ensemble,
                SC_ensemble,
                RESPONSIBILITY_ensemble,
                INDUSTRY_ensemble)
NLP=data.frame(NLP_R_Conscientious,
            NLP_R_Order,
            NLP_R_Virtue,
            NLP_R_Tradition,
            NLP_R_SControl,
            NLP_R_Responsibility,
            NLP_R_Industry)
CCS=data.frame(parameter_estimates_for_matrix$CCS_Conscientiousness,
            parameter_estimates_for_matrix$CCS_Order,
            parameter_estimates_for_matrix$CCS_Virtue,
            parameter_estimates_for_matrix$CCS_Traditionalism,
            parameter_estimates_for_matrix$`CCS_Self-Control`,
            parameter_estimates_for_matrix$CCS_Responsibility,
            parameter_estimates_for_matrix$CCS_Industriousness)
FC=data.frame(parameter_estimates_for_matrix$`GTUM Conscientiousness`,
           parameter_estimates_for_matrix$`GTUM Order`,
```

```
        parameter_estimates_for_matrix$`GTUM Virtue`,
        parameter_estimates_for_matrix$`GTUM Traditionalism`,
        parameter_estimates_for_matrix$`GTUM Self-Control`,
        parameter_estimates_for_matrix$`GTUM Responsibility`,
        parameter_estimates_for_matrix$`GTUM Industriousness`)
All_estimates_final = data.frame(ensemble,NLP,CCS,FC)
write.csv(All_estimates_final,"all estimates FINAL.csv")

#hierarchical regression and criterion-related validity for MTMM table
data =X202402029Cleaned_All

#hierarchical regression for GPA
step1.GPA = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(step1.GPA)
step2.GPA = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
+

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(step2.GPA)
step3.GPA = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
+

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess. +
        Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(step3.GPA)
step4.GPA = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
+

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess. +
        Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious +
        Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(step4.GPA)

#hierarchical regression for SWB
```

```
step1.SWB = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(step1.SWB)
step2.SWB = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
+

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(step2.SWB)
step3.SWB = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
+

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess. +
        Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(step3.SWB)
step4.SWB = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
+

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess. +
        Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious +
        Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(step4.SWB)


#adding to the MTMM table
ens.gpa.con = lm(Outcome_variables_and_demographics$AdjGPA ~
          Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(ens.gpa.con)
ens.gpa.ind = lm(Outcome_variables_and_demographics$AdjGPA ~
          Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(ens.gpa.ind)
ens.gpa.ord = lm(Outcome_variables_and_demographics$AdjGPA ~
          Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(ens.gpa.ord)
ens.gpa.sc = lm(Outcome_variables_and_demographics$AdjGPA ~
          Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
```

```
summary(ens.gpa.sc)
ens.gpa.resp = lm(Outcome_variables_and_demographics$AdjGPA ~
            Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(ens.gpa.resp)
ens.gpa.vir = lm(Outcome_variables_and_demographics$AdjGPA ~
             Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(ens.gpa.vir)
ens.gpa.trd = lm(Outcome_variables_and_demographics$AdjGPA ~
             Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(ens.gpa.trd)
#fc
FC.gpa.con = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.gpa.con)
FC.gpa.ind = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.gpa.ind)
FC.gpa.ord = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.gpa.ord)
FC.gpa.sc = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.gpa.sc)
FC.gpa.resp = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.gpa.resp)
FC.gpa.vir = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.gpa.vir)
FC.gpa.trd = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.gpa.trd)
#nlp
NLP.gpa.con = lm(Outcome_variables_and_demographics$AdjGPA ~
             Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(NLP.gpa.con)
NLP.gpa.ind = lm(Outcome_variables_and_demographics$AdjGPA ~
```

```
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.gpa.ind)
NLP.gpa.ord = lm(Outcome_variables_and_demographics$AdjGPA ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.gpa.ord)
NLP.gpa.sc = lm(Outcome_variables_and_demographics$AdjGPA ~
           Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.gpa.sc)
NLP.gpa.resp = lm(Outcome_variables_and_demographics$AdjGPA ~
             Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
summary(NLP.gpa.resp)
NLP.gpa.vir = lm(Outcome_variables_and_demographics$AdjGPA ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.gpa.vir)
NLP.gpa.trd = lm(Outcome_variables_and_demographics$AdjGPA ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.gpa.trd)
#CCS
CCS.gpa.con = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(CCS.gpa.con)
CCS.gpa.ind = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.gpa.ind)
CCS.gpa.ord = lm(Outcome_variables_and_demographics$AdjGPA ~
            Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.gpa.ord)
CCS.gpa.sc = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.gpa.sc)
CCS.gpa.resp = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.gpa.resp)
CCS.gpa.vir = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.gpa.vir)
CCS.gpa.trd = lm(Outcome_variables_and_demographics$AdjGPA ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.gpa.trd)
```

```
#SWB
ens.swb.con = lm(Outcome_variables_and_demographics$SWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(ens.swb.con)
ens.swb.ind = lm(Outcome_variables_and_demographics$SWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(ens.swb.ind)
ens.swb.ord = lm(Outcome_variables_and_demographics$SWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(ens.swb.ord)
ens.swb.sc = lm(Outcome_variables_and_demographics$SWB ~
                Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
summary(ens.swb.sc)
ens.swb.resp = lm(Outcome_variables_and_demographics$SWB ~
                  Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(ens.swb.resp)
ens.swb.vir = lm(Outcome_variables_and_demographics$SWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(ens.swb.vir)
ens.swb.trd = lm(Outcome_variables_and_demographics$SWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(ens.swb.trd)
#fc
FC.swb.con = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.swb.con)
FC.swb.ind = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.swb.ind)
FC.swb.ord = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.swb.ord)
FC.swb.sc = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.swb.sc)
FC.swb.resp = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.swb.resp)
FC.swb.vir = lm(Outcome_variables_and_demographics$SWB ~
```

```
Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.swb.vir)
FC.swb.trd = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.swb.trd)
#nlp
NLP.swb.con = lm(Outcome_variables_and_demographics$SWB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(NLP.swb.con)
NLP.swb.ind = lm(Outcome_variables_and_demographics$SWB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.swb.ind)
NLP.swb.ord = lm(Outcome_variables_and_demographics$SWB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.swb.ord)
NLP.swb.sc = lm(Outcome_variables_and_demographics$SWB ~
                  Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.swb.sc)
NLP.swb.resp = lm(Outcome_variables_and_demographics$SWB ~
                    Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
summary(NLP.swb.resp)
NLP.swb.vir = lm(Outcome_variables_and_demographics$SWB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.swb.vir)
NLP.swb.trd = lm(Outcome_variables_and_demographics$SWB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.swb.trd)
#CCS
CCS.swb.con = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(CCS.swb.con)
CCS.swb.ind = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.swb.ind)
CCS.swb.ord = lm(Outcome_variables_and_demographics$SWB ~
                   Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.swb.ord)
CCS.swb.sc = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.swb.sc)
```

```
CCS.swb.resp = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.swb.resp)
CCS.swb.vir = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.swb.vir)
CCS.swb.trd = lm(Outcome_variables_and_demographics$SWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.swb.trd)
#OCB
ocb.swb.con = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(ocb.swb.con)
ocb.swb.ind = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(ocb.swb.ind)
ocb.swb.ord = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(ocb.swb.ord)
ocb.swb.sc = lm(numeric_202402029Cleaned_All$OCB ~
                  Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
summary(ocb.swb.sc)
ocb.swb.resp = lm(numeric_202402029Cleaned_All$OCB ~
                    Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(ocb.swb.resp)
ocb.swb.vir = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(ocb.swb.vir)
ocb.swb.trd = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(ocb.swb.trd)
#fc
FC.swb.con = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.swb.con)
FC.swb.ind = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.swb.ind)
FC.swb.ord = lm(numeric_202402029Cleaned_All$OCB ~
```

```
Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.swb.ord)
FC.swb.sc = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.swb.sc)
FC.swb.resp = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.swb.resp)
FC.swb.vir = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.swb.vir)
FC.swb.trd = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.swb.trd)
#nlp
NLP.swb.con = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(NLP.swb.con)
NLP.swb.ind = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.swb.ind)
NLP.swb.ord = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.swb.ord)
NLP.swb.sc = lm(numeric_202402029Cleaned_All$OCB ~
                  Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.swb.sc)
NLP.swb.resp = lm(numeric_202402029Cleaned_All$OCB ~
                    Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
summary(NLP.swb.resp)
NLP.swb.vir = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.swb.vir)
NLP.swb.trd = lm(numeric_202402029Cleaned_All$OCB ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.swb.trd)
#CCS
CCS.swb.con = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
```

```
summary(CCS.swb.con)
CCS.swb.ind = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.swb.ind)
CCS.swb.ord = lm(numeric_202402029Cleaned_All$OCB ~
            Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.swb.ord)
CCS.swb.sc = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.swb.sc)
CCS.swb.resp = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.swb.resp)
CCS.swb.vir = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.swb.vir)
CCS.swb.trd = lm(numeric_202402029Cleaned_All$OCB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.swb.trd)
#CWB
cwb.swb.con = lm(numeric_202402029Cleaned_All$CWB ~
            Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(cwb.swb.con)
cwb.swb.ind = lm(numeric_202402029Cleaned_All$CWB ~
            Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(cwb.swb.ind)
cwb.swb.ord = lm(numeric_202402029Cleaned_All$CWB ~
            Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(cwb.swb.ord)
cwb.swb.sc = lm(numeric_202402029Cleaned_All$CWB ~
           Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
summary(cwb.swb.sc)
cwb.swb.resp = lm(numeric_202402029Cleaned_All$CWB ~
             Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(cwb.swb.resp)
cwb.swb.vir = lm(numeric_202402029Cleaned_All$CWB ~
            Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(cwb.swb.vir)
cwb.swb.trd = lm(numeric_202402029Cleaned_All$CWB ~
            Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(cwb.swb.trd)
```

```
#fc
FC.swb.con = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.swb.con)
FC.swb.ind = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.swb.ind)
FC.swb.ord = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.swb.ord)
FC.swb.sc = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.swb.sc)
FC.swb.resp = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.swb.resp)
FC.swb.vir = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.swb.vir)
FC.swb.trd = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.swb.trd)
#nlp
NLP.swb.con = lm(numeric_202402029Cleaned_All$CWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(NLP.swb.con)
NLP.swb.ind = lm(numeric_202402029Cleaned_All$CWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.swb.ind)
NLP.swb.ord = lm(numeric_202402029Cleaned_All$CWB ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.swb.ord)
NLP.swb.sc = lm(numeric_202402029Cleaned_All$CWB ~
                Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.swb.sc)
NLP.swb.resp = lm(numeric_202402029Cleaned_All$CWB ~
                  Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
```

```
summary(NLP.swb.resp)
NLP.swb.vir = lm(numeric_202402029Cleaned_All$CWB ~
        Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.swb.vir)
NLP.swb.trd = lm(numeric_202402029Cleaned_All$CWB ~
        Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.swb.trd)
#CCS
CCS.swb.con = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(CCS.swb.con)
CCS.swb.ind = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.swb.ind)
CCS.swb.ord = lm(numeric_202402029Cleaned_All$CWB ~
        Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.swb.ord)
CCS.swb.sc = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.swb.sc)
CCS.swb.resp = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.swb.resp)
CCS.swb.vir = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.swb.vir)
CCS.swb.trd = lm(numeric_202402029Cleaned_All$CWB ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.swb.trd)
#OJP
ojp.swb.con = lm(numeric_202402029Cleaned_All$OJP ~
        Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(ojp.swb.con)
ojp.swb.ind = lm(numeric_202402029Cleaned_All$OJP ~
        Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(ojp.swb.ind)
ojp.swb.ord = lm(numeric_202402029Cleaned_All$OJP ~
        Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(ojp.swb.ord)
```

```
ojp.swb.sc = lm(numeric_202402029Cleaned_All$OJP ~
                Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
summary(ojp.swb.sc)
ojp.swb.resp = lm(numeric_202402029Cleaned_All$OJP ~
                  Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(ojp.swb.resp)
ojp.swb.vir = lm(numeric_202402029Cleaned_All$OJP ~
                 Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(ojp.swb.vir)
ojp.swb.trd = lm(numeric_202402029Cleaned_All$OJP ~
                 Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(ojp.swb.trd)
#fc
FC.swb.con = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.swb.con)
FC.swb.ind = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.swb.ind)
FC.swb.ord = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.swb.ord)
FC.swb.sc = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.swb.sc)
FC.swb.resp = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.swb.resp)
FC.swb.vir = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.swb.vir)
FC.swb.trd = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.swb.trd)
#nlp
NLP.swb.con = lm(numeric_202402029Cleaned_All$OJP ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
```

```
summary(NLP.swb.con)
NLP.swb.ind = lm(numeric_202402029Cleaned_All$OJP ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.swb.ind)
NLP.swb.ord = lm(numeric_202402029Cleaned_All$OJP ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.swb.ord)
NLP.swb.sc = lm(numeric_202402029Cleaned_All$OJP ~
           Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.swb.sc)
NLP.swb.resp = lm(numeric_202402029Cleaned_All$OJP ~
             Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
summary(NLP.swb.resp)
NLP.swb.vir = lm(numeric_202402029Cleaned_All$OJP ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.swb.vir)
NLP.swb.trd = lm(numeric_202402029Cleaned_All$OJP ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.swb.trd)
#CCS
CCS.swb.con = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(CCS.swb.con)
CCS.swb.ind = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.swb.ind)
CCS.swb.ord = lm(numeric_202402029Cleaned_All$OJP ~
            Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.swb.ord)
CCS.swb.sc = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.swb.sc)
CCS.swb.resp = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.swb.resp)
CCS.swb.vir = lm(numeric_202402029Cleaned_All$OJP ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.swb.vir)
CCS.swb.trd = lm(numeric_202402029Cleaned_All$OJP ~
```

```
Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.swb.trd)
#IWJ
iwj.swb.con = lm(numeric_202402029Cleaned_All$IWJ ~
                 Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(iwj.swb.con)
iwj.swb.ind = lm(numeric_202402029Cleaned_All$IWJ ~
                 Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(iwj.swb.ind)
iwj.swb.ord = lm(numeric_202402029Cleaned_All$IWJ ~
                 Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(iwj.swb.ord)
iwj.swb.sc = lm(numeric_202402029Cleaned_All$IWJ ~
                Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
summary(iwj.swb.sc)
iwj.swb.resp = lm(numeric_202402029Cleaned_All$IWJ ~
                  Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(iwj.swb.resp)
iwj.swb.vir = lm(numeric_202402029Cleaned_All$IWJ ~
                 Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(iwj.swb.vir)
iwj.swb.trd = lm(numeric_202402029Cleaned_All$IWJ ~
                 Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(iwj.swb.trd)
#fc
FC.swb.con = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.swb.con)
FC.swb.ind = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.swb.ind)
FC.swb.ord = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.swb.ord)
FC.swb.sc = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.swb.sc)
FC.swb.resp = lm(numeric_202402029Cleaned_All$IWJ ~
```

```
Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.swb.resp)
FC.swb.vir = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.swb.vir)
FC.swb.trd = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.swb.trd)
#nlp
NLP.swb.con = lm(numeric_202402029Cleaned_All$IWJ ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(NLP.swb.con)
NLP.swb.ind = lm(numeric_202402029Cleaned_All$IWJ ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.swb.ind)
NLP.swb.ord = lm(numeric_202402029Cleaned_All$IWJ ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.swb.ord)
NLP.swb.sc = lm(numeric_202402029Cleaned_All$IWJ ~
                  Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.swb.sc)
NLP.swb.resp = lm(numeric_202402029Cleaned_All$IWJ ~
                    Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
summary(NLP.swb.resp)
NLP.swb.vir = lm(numeric_202402029Cleaned_All$IWJ ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.swb.vir)
NLP.swb.trd = lm(numeric_202402029Cleaned_All$IWJ ~
                   Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.swb.trd)
#CCS
CCS.swb.con = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(CCS.swb.con)
CCS.swb.ind = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.swb.ind)
CCS.swb.ord = lm(numeric_202402029Cleaned_All$IWJ ~
                   Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.swb.ord)
```

```
CCS.swb.sc = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.swb.sc)
CCS.swb.resp = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.swb.resp)
CCS.swb.vir = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.swb.vir)
CCS.swb.trd = lm(numeric_202402029Cleaned_All$IWJ ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.swb.trd)
#ERSE
erse.swb.con = lm(numeric_202402029Cleaned_All$ERSE ~
                  Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(erse.swb.con)
erse.swb.ind = lm(numeric_202402029Cleaned_All$ERSE ~
                  Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(erse.swb.ind)
erse.swb.ord = lm(numeric_202402029Cleaned_All$ERSE ~
                  Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(erse.swb.ord)
erse.swb.sc = lm(numeric_202402029Cleaned_All$ERSE ~
                 Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
summary(erse.swb.sc)
erse.swb.resp = lm(numeric_202402029Cleaned_All$ERSE ~
                   Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(erse.swb.resp)
erse.swb.vir = lm(numeric_202402029Cleaned_All$ERSE ~
                  Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(erse.swb.vir)
erse.swb.trd = lm(numeric_202402029Cleaned_All$ERSE ~
                  Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(erse.swb.trd)
#fc
FC.swb.con = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.swb.con)
FC.swb.ind = lm(numeric_202402029Cleaned_All$ERSE ~
```

```
Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.swb.ind)
FC.swb.ord = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.swb.ord)
FC.swb.sc = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.swb.sc)
FC.swb.resp = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.swb.resp)
FC.swb.vir = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.swb.vir)
FC.swb.trd = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.swb.trd)
#nlp
NLP.swb.con = lm(numeric_202402029Cleaned_All$ERSE ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(NLP.swb.con)
NLP.swb.ind = lm(numeric_202402029Cleaned_All$ERSE ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.swb.ind)
NLP.swb.ord = lm(numeric_202402029Cleaned_All$ERSE ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.swb.ord)
NLP.swb.sc = lm(numeric_202402029Cleaned_All$ERSE ~
                Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.swb.sc)
NLP.swb.resp = lm(numeric_202402029Cleaned_All$ERSE ~
                  Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
summary(NLP.swb.resp)
NLP.swb.vir = lm(numeric_202402029Cleaned_All$ERSE ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.swb.vir)
NLP.swb.trd = lm(numeric_202402029Cleaned_All$ERSE ~
                 Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.swb.trd)
```

```
#CCS
CCS.swb.con = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(CCS.swb.con)
CCS.swb.ind = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.swb.ind)
CCS.swb.ord = lm(numeric_202402029Cleaned_All$ERSE ~
            Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.swb.ord)
CCS.swb.sc = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.swb.sc)
CCS.swb.resp = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.swb.resp)
CCS.swb.vir = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.swb.vir)
CCS.swb.trd = lm(numeric_202402029Cleaned_All$ERSE ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.swb.trd)
#AJIG
ajig.swb.con = lm(numeric_202402029Cleaned_All$AJIG ~
            Scores_for_Ensemble_NLP_FC_CCS$CONSCIENTIOUSNESS_ensemble)
summary(ajig.swb.con)
ajig.swb.ind = lm(numeric_202402029Cleaned_All$AJIG ~
            Scores_for_Ensemble_NLP_FC_CCS$INDUSTRY_ensemble)
summary(ajig.swb.ind)
ajig.swb.ord = lm(numeric_202402029Cleaned_All$AJIG ~
            Scores_for_Ensemble_NLP_FC_CCS$ORDER_ensemble)
summary(ajig.swb.ord)
ajig.swb.sc = lm(numeric_202402029Cleaned_All$AJIG ~
           Scores_for_Ensemble_NLP_FC_CCS$SC_ensemble)
summary(ajig.swb.sc)
ajig.swb.resp = lm(numeric_202402029Cleaned_All$AJIG ~
             Scores_for_Ensemble_NLP_FC_CCS$RESPONSIBILITY_ensemble)
summary(ajig.swb.resp)
ajig.swb.vir = lm(numeric_202402029Cleaned_All$AJIG ~
```

```
            Scores_for_Ensemble_NLP_FC_CCS$VIRTUE_ensemble)
summary(ajig.swb.vir)
ajig.swb.trd = lm(numeric_202402029Cleaned_All$AJIG ~
            Scores_for_Ensemble_NLP_FC_CCS$TRADITION_ensemble)
summary(ajig.swb.trd)
#fc
FC.swb.con = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Conscientiousn
ess.)
summary(FC.swb.con)
FC.swb.ind = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Industriousness.
)
summary(FC.swb.ind)
FC.swb.ord = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Order.)
summary(FC.swb.ord)
FC.swb.sc = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Self.Control.)
summary(FC.swb.sc)
FC.swb.resp = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Responsibility.)
summary(FC.swb.resp)
FC.swb.vir = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Virtue.)
summary(FC.swb.vir)
FC.swb.trd = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..GTUM.Traditionalism.)
summary(FC.swb.trd)
#nlp
NLP.swb.con = lm(numeric_202402029Cleaned_All$AJIG ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Conscientious)
summary(NLP.swb.con)
NLP.swb.ind = lm(numeric_202402029Cleaned_All$AJIG ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Industry)
summary(NLP.swb.ind)
NLP.swb.ord = lm(numeric_202402029Cleaned_All$AJIG ~
            Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Order)
summary(NLP.swb.ord)
```

```
NLP.swb.sc = lm(numeric_202402029Cleaned_All$AJIG ~
              Scores_for_Ensemble_NLP_FC_CCS$NLP_R_SControl)
summary(NLP.swb.sc)
NLP.swb.resp = lm(numeric_202402029Cleaned_All$AJIG ~
               Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Responsibility)
summary(NLP.swb.resp)
NLP.swb.vir = lm(numeric_202402029Cleaned_All$AJIG ~
              Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Virtue)
summary(NLP.swb.vir)
NLP.swb.trd = lm(numeric_202402029Cleaned_All$AJIG ~
              Scores_for_Ensemble_NLP_FC_CCS$NLP_R_Tradition)
summary(NLP.swb.trd)
#CCS
CCS.swb.con = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Conscientiousness
)
summary(CCS.swb.con)
CCS.swb.ind = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Industriousness)
summary(CCS.swb.ind)
CCS.swb.ord = lm(numeric_202402029Cleaned_All$AJIG ~
              Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Order)
summary(CCS.swb.ord)
CCS.swb.sc = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix..CCS_Self.Control.)
summary(CCS.swb.sc)
CCS.swb.resp = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Responsibility)
summary(CCS.swb.resp)
CCS.swb.vir = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Virtue)
summary(CCS.swb.vir)
CCS.swb.trd = lm(numeric_202402029Cleaned_All$AJIG ~

Scores_for_Ensemble_NLP_FC_CCS$parameter_estimates_for_matrix.CCS_Traditionalism)
summary(CCS.swb.trd)

#topic modeling for supplementary RQ
library(tm)
library(topicmodels)
library(tidytext)
```

```
library(slam)
library(ggplot2)
library(dplyr)
#looking at only response columns
text =
X202402029Cleaned_All[c(25,27,29,31,33,35,37,39,41,43,45,47,49,51,53,55,57,59,61,63,65,67,
69,71,
                        73,75,77,79,81,83,85,87,89,91,93,95,97,99,101,103,105,107,109,111,113)]
#Data cleaning and preprocessing
corpus <- VCorpus(VectorSource(text))
corpus <- tm_map(corpus, content_transformer(tolower))
corpus <- tm_map(corpus, removePunctuation)
corpus <- tm_map(corpus, removeNumbers)
corpus <- tm_map(corpus, removeWords, stopwords("english"))

# Create document-term matrices for unigrams, bigrams, and trigrams
dtm_unigram <- DocumentTermMatrix(corpus)
dtm_bigram <- DocumentTermMatrix(corpus,
                        control = list(tokenize = function(x) unlist(lapply(ngrams(words(x), 2),
                                                                            paste, collapse = " "))))
dtm_trigram <- DocumentTermMatrix(corpus,
                         control = list(tokenize = function(x) unlist(lapply(ngrams(words(x), 3),
                                                                             paste, collapse = " "))))

#LDA models - k=6 for 6 facets
lda_unigram <- LDA(dtm_unigram, k = 6)
topics_unigram <- topics(lda_unigram)

lda_bigram <- LDA(dtm_bigram, k = 6)
topics_bigram <- topics(lda_bigram)

lda_trigram <- LDA(dtm_trigram, k = 6)
topics_trigram <- topics(lda_trigram)


#get topic terms
terms(lda_unigram, 10)
terms(lda_bigram, 10)
terms(lda_trigram, 10)

#uni, bi, + tri grams
combine_ngrams <- function(x) {
  unigrams <- unlist(strsplit(as.character(x), " "))
  bigrams <- unlist(lapply(ngrams(words(x), 2), paste, collapse = " "))
  trigrams <- unlist(lapply(ngrams(words(x), 3), paste, collapse = " "))
  return(c(unigrams, bigrams, trigrams))
```

```
}
dtm_combined <- DocumentTermMatrix(corpus, control = list(tokenize = combine_ngrams))

#LDA model on the combined DTM
lda_combined <- LDA(dtm_combined, k = 6)
topics_combined <- topics(lda_combined)
terms(lda_combined, 20)


#### visualization ####
#lda_bigram
topics <- tidy(lda_bigram, matrix = "beta")

top_terms <- topics %>%
  group_by(topic) %>%
  arrange(desc(beta)) %>%
  slice_head(n = 5) %>%
  ungroup() %>%
  arrange(topic, -beta)

top_terms %>%
  group_by(topic) %>%
  mutate(term = reorder_within(term, beta, topic)) %>%
  ggplot(aes(term, beta, fill = factor(topic))) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~ topic, scales = "free_y") +
  coord_flip() +
  scale_x_reordered()
```