

**Exploring the Nature of Putatively Interstitial Traits:  
How do Interstitial Stand-alone Trait Scales Behave on the  
Item Level When Using the Big Five as a Framework?**

Master Thesis June 2021: Master of Psychology



University of  
Southern Denmark

**Authors:** Sofie Emilie Kristensen Dindorp (02.08.1994) & Stefan Crone Larsen (03.03.1987)

**Supervisor:** Steven Ludeke

**University of Southern Denmark:** Department of Psychology

**Adjusted pages:** 62,07

**Characters:** 148.971

## Abstract

**Baggrund:** 'The Big Five' som model er den dominerende måde at konceptualisere personlighed på i psykologien. Modellen fremstilles ofte som en simpel, hierarkisk struktur, som hævder at individer adskiller sig i træk langs fem brede dimensioner. Der er dog evidens for, at mange skalaer som måler personlighedstræk ikke ligger inden for ét Big Five-domæne, men flere. Selvom udbredelsen af disse interstitielle (dvs. mellemliggende) træk kan have betydning for Big Five-modellens påstand om at organisere personlighed effektivt, er de ikke særlig undersøgt, og det er derfor uvist, hvorfor så mange skalaer lader til at befinde sig imellem domæner. Dette studie undersøger dette ved at gennemgå interstitielle skalaers items og disses korrelationer med Big Five-domænerne. **Metode:** Deltagerne ( $N = 1134$ ) var en del af det amerikanske Eugene-Springfield Community Sample. De besvarede en lang række spørgeskemaer, inklusiv en måling af Big Five-domænerne samt adskillige yderligere målinger af personlighedstræk. Korrelationsanalyser blev anvendt på 41 skalaer for at undersøge sammenhængen mellem Big Five-domæner og skalaer der måler enkeltstående personlighedstræk. Skalaer som havde en korrelation på .30 med to eller flere domæner blev herefter undersøgt på item-niveau gennem korrelationsanalyser med Big Five-domænerne. **Resultater:** Resultaterne indikerede, at 11 ud af 41 skalaer var interstitielle. Fire ud af de 11 interstitielle skalaer lader til at være opbygget af sammenhængende items, der i sig selv er interstitielle, mens de resterende skalaers interstitialitet er uklar eller bedst forklares ud fra at forskellige items er korreleret med forskellige domæner snarere end at måle et sammenhængende, interstitielt personlighedstræk. **Konklusion:** Fundene antyder at der ikke er et entydigt billede af, hvordan interstitielle skalaer har afsæt i deres items, idet der både findes eksempler på overbevisende interstitialitet og på uklare item-sammensætninger. Der er derfor forskellige årsager til udbredelsen af interstitielle skalaer, men det understreger også at der er behov for yderligere forskning for at afdække, hvor stor udbredelsen af interstitielle skalaer, der måler et ægte interstitielt personlighedstræk, er.

### **Acknowledgements**

We would like to express our deepest gratitude to our supervisor Steven Ludeke, who has provided us with exceptional guidance and invaluable advice. We are grateful to Lew Goldberg and the Oregon Research Institute for making their data publicly available. We would also like to thank Timothy Bainbridge, Steven Ludeke and Luke Smillie for sharing their paper with us before publication. Also, an extra gratitude to Timothy Bainbridge for providing us with coding for the ESCS data that eased the statistical analyses. Finally, we wish to thank our friends, families and our significant others for indispensable support.

## Table of contents

<b>1. INTRODUCTION.....</b>	<b>5</b>
1.1. READER’S GUIDE .....	7
1.2. THE BIG FIVE.....	7
1.3. BEYOND THE BIG FIVE .....	11
1.4. SUMMARY, RATIONALE, AND RESEARCH QUESTION .....	15
<b>2. METHODS .....</b>	<b>17</b>
2.1. STATISTICS .....	17
2.1.1. Correlations.....	17
2.2. SCALE SELECTION AND EVALUATION .....	17
2.3. PARTICIPANTS .....	19
2.4. PROCEDURE .....	19
2.5. MEASURES .....	19
2.5.1. NEO PI-R.....	20
2.5.2. BAS Drive.....	20
2.5.3. Borderline Personality Inventory .....	20
2.5.4. Cognitive Failures Questionnaire .....	21
2.5.5. Dispositional Optimism .....	21
2.5.6. Impression Management .....	21
2.5.7. Levenson Self-Report Psychopathy Scale.....	22
2.5.8. Narcissism Personality Inventory .....	22
2.5.9. Rosenberg Self-Esteem Scale .....	23
2.5.10. Self-Deception.....	23
2.5.11. Sensitivity to Punishment.....	24
2.5.12. Toronto Alexithymia Scale .....	24
<b>3. RESULTS.....</b>	<b>24</b>
3.1. INITIAL ANALYSES .....	25
3.2. ITEM-CORRELATIONS .....	26
3.2.1. BAS drive .....	27
3.2.2. Borderline Personality Inventory .....	28
3.2.3. Cognitive Failures Questionnaire .....	29
3.2.4. Dispositional Optimism .....	30
3.2.5. Impression Management .....	31
3.2.6. Levenson Self-Report Psychopathy Scale.....	31
3.2.7. Narcissism Personality Inventory (NPI-16) .....	32
3.2.8. Rosenberg Self-Esteem .....	33
3.2.9. Self-Deception.....	34
3.2.10. Sensitivity to Punishment.....	35
3.2.11. Toronto Alexithymia Scale .....	36

<b>4. DISCUSSION .....</b>	<b>37</b>
4.1. MAIN FINDINGS .....	37
4.2. INTERSTITIAL SCALES AND THEIR BIG FIVE CORRELATIONS .....	38
4.3. A COMMENT ON OUR THRESHOLDS .....	40
4.4. HOW ARE INTERSTITIAL TRAIT SCALES CONSTRUCTED ON THE ITEM LEVEL? .....	41
4.5. THE BIGGER PICTURE OF INTERSTITIALITY IN THE BIG FIVE FRAMEWORK .....	47
4.6. FURTHER CONSIDERATIONS .....	50
<b>REFERENCES.....</b>	<b>66</b>

## 1. Introduction

When speaking of personality psychology, there are several ways to approach the understanding of traits and how they influence behavior. Particularly the essential trait approach, which focuses on few traits out of the thousands of options, has come to the forefront of research in personality psychology (Funder, 2013). Frameworks that take this approach, such as The Big Five or the Five Factor Model (Costa & McCrae, 1995; Goldberg, 1992), and HEXACO (Ashton & Lee, 2001), focus on few, broad core traits, composed of multiple, narrower facets, thus making a hierarchical structure of personality. Particularly the Big Five or Five Factor model has won headway and become a big focus in personality research as the prevailing conceptualization of personality (Funder, 2013; Strus et al., 2014). The roots of this model began more than 80 years ago when Allport and Odbert (1936) scoured the dictionary to find all trait-like words and found 17,953. Later, two traditions led to the development of the five broad personality factors - the lexical approach, largely connected to Lewis Goldberg and the term ‘the Big Five’, and the questionnaire approach, which is largely linked to Costa & McCrae’s work and the term ‘the Five-Factor Model’ (Boudreaux & Ozer, 2015; Costa & McCrae, 1995; Goldberg, 1992; Goldberg, 1993). Today, the models are so closely related that the two terms are often used interchangeably (Gurven et al., 2013; Strus et al., 2014). Likewise, in this thesis we will not distinguish between the two except when necessary to describe the differences and will henceforth use the term ‘the Big Five’ collectively.

Proponents of the Big Five argued that this model could provide a robust, scientific framework that could be used to organize the many individual, trait-like differences between people (Goldberg, 1993). Though there has been some disagreement as to the labels of the factors, what exactly each factor is constituted of, and the universality of the factors - particularly the fifth factor, Openness to Experience (henceforth called Openness) - the Big Five remains at the forefront of personality research, and even contributes to other fields outside of psychology (John, 2021).

The Big Five centers around five broad traits, also called domains - Extraversion, Agreeableness, Conscientiousness, Neuroticism, and Openness - which can in turn be divided into facets and nuances (McCrae, 2020). The original idea was that the five domains would be orthogonal, meaning that there should be no correlation between them; getting a high or low score on one should not predict whether one would get a high or low score on any of the others (Funder, 2013; John, 2021). As it turns out, though, the Big Five are not as simple and clean-cut as they may seem. For example, some of the domains are indeed somewhat related

to one another (Digman, 1997). It seems that three of the Big Five domains - Neuroticism, Conscientiousness, and Agreeableness - constitutes a dimension coined *Stability* (or Alpha), while the other two - Extraversion and Openness - constitutes a dimension called *Plasticity* (or Beta) (DeYoung, 2006; Digman, 1997).

Additionally, beyond this, there do seem to be some personality constructs that do not fit too neatly into the tree-like structure that was proposed. The Big Five has received criticism for failing to be comprehensive, in that some traits are seemingly distinct from the Big Five (Hough, 1992; Paunonen & Jackson, 2000) - though some argue that not nearly as many traits as typically thought are distinct from the Big Five (Credé et al., 2016; Saucier & Goldberg, 1998). Some traits - including some seemingly distinct ones - appear to be interstitial, and thus lie between several broader domains rather than fitting into a certain branch of the framework (Costa & McCrae, 1995; Krueger & Markon, 2014; Mooradian et al., 2016). Developers of the NEO inventory assessments of the Big Five, Costa and McCrae (1995), have acknowledged these interstitial traits, as have many other researchers (Credé et al., 2016; Hofstee et al., 1992; Mooradian et al., 2016). In fact, it seems that interstitial traits are actually quite common (Hopwood & Donnellan, 2010). According to De Raad and Barelds (2020) most Big Five trait-variables were found to load substantively on two factors, while less than 10% loaded on only one factor - an observation that drastically compromised the simple-structure concept (De Raad & Barelds, 2020).

Despite this prevalence of interstitial traits, it is hard to find research actually focusing on them. As an example, we have found no papers focusing explicitly on the interstitial traits within a Big Five perspective through our literature search. Part of this hardship may be due to a lack of a common language regarding them, as not everyone calls them interstitial traits. For example, John (2021) calls them “those characteristics that fall in the fuzzy regions *between* the factors” (p. 43, italics in original), which is perhaps slightly long to be a useful label. Others call them compound traits (Credé et al., 2016; Hough et al., 2015) due to them being composed of different domains (Connelly et al., 2018). In this branch of the literature, most attention has been focused on traits such as integrity and customer service orientation, and whether it is possible to synthetically form variables such as these; among them is Hough and Ones (2001), and Credé et al. (2016), who looked into whether certain compound traits could be approximated from composites of Big Five domains. If a compound trait could be indirectly computed this way, useful predictors of real-world outcomes could be obtained exclusively through a thorough measurement of the Big Five (Credé et al., 2016). To our knowledge, however, not much attention beyond this has

been given to interstitial (as they will henceforth be called) traits. Rather, it seems that when interstitiality is mentioned - often in papers attempting to locate individual traits within the Big Five - it is usually a brief comment on the issue which is then quickly set aside to focus on other things (Mooradian et al., 2011; Roberts et al., 2005).

We intend to rectify this in our paper by investigating interstitial traits in their own right. In some cases, an interstitial ‘trait’ could be merely an adjective, that is, a single word describing a characteristic feature, such as the 17,953 words Allport and Odbert (1936) found in the dictionary. Though these are interesting in their own right, this paper will focus on interstitial traits that are measured by a scale (e.g., self-esteem, narcissism). Such scales and their relations to the Big Five were investigated in a recent study by Bainbridge et al. (in press), who found that these so-called stand-alone scales were highly connected to the Big Five, though they were often interstitial. The aim of this study is to explore what makes trait scales interstitial; whether the constructs themselves are interstitial, or whether it is due to a problematic combination of items within the scales.

### **1.1. Reader’s guide**

The following thesis is an empirical study of item-level compositions of interstitial scales in relation to the Big Five domains. Using a publicly available dataset, we wish to shed light on what causes what appear to be interstitiality in trait scales. The current section will briefly illustrate the disposition of the thesis.

This thesis’ structure is based on the IMRaD format. The first section following this includes a presentation of the development, structure, and measurement of the Big Five model. We then review the phenomenon of interstitial traits. The section finishes off with a summary and the study’s rationale and research question. In the second section, we present our methods, with information on statistical analyses, participants, procedure, and measures. The third section contains results from the statistical analyses, including tables and presentation in the text. Finally, the fourth section encompasses the discussion, which provides a summary of the results and then a presentation of our findings in more detail, followed by implications, strengths and limitations, and conclusions.

### **1.2. The Big Five**

Since the promulgation of the Five Factor Model of personality (Costa & McCrae, 1985; McCrae & Costa, 1985a) or Big Five (Goldberg, 1990), the model has won headway and become a big focus in personality research, particularly because it is generally seen as a



comprehensive framework to map personality (John, 2021; McCrae, 2020). The Big Five is often depicted as a simple hierarchical structure, where the five traits - or domains - are usually divided into six facets per domain. These traits capture a lot of variance within personality psychology and the model manages to bring together a variety of very different traits under the five common labels. They are also quite useful in predicting life outcomes, such as popularity, work and academic performance, marital satisfaction, and plain happiness (Funder, 2013).

There is no single developer of the Big Five; numerous researchers have contributed to the creation of the Big Five since the 1930's. The Big Five was originally based on a combination of factor analysis and the lexical approach, which builds on the assumption that most important personality traits are encoded as words in natural languages (John, 2021). The roots of the Big Five go back to 1936, when Allport and Odbert went through the English dictionary and found 17.953 trait terms. They reduced these to 4.504 terms, describing stable personality traits. These 4.504 traits were then used by Cattell (1943) to do a lexical analysis. He reduced the list to 35 clusters of personality traits. Seven years later, Fiske took 22 of these clusters through a factor analysis and discovered that they would fit in a five-factor solution (Fiske, 1949). Tupes and Christal (1961) and Norman (1963) consolidated the five-factor solution, and a long list of replication research has led to the Big Five as we know it today (Botwin & Buss, 1989; Digman & Inouye, 1986; Goldberg, 1981; McCrae & Costa, 1985b; Rammstedt et al., 2010).

Generally, one could say that there are two approaches or traditions within the development of the model: The lexical approach and the questionnaire approach (Boudreaux & Ozer, 2015). Most of the research outlined above belongs to the former. The lexical approach relies on studies based on the lexical hypothesis - that important aspects of human life will eventually be encoded in the language, meaning that if a trait is truly a salient, relevant, and universal trait, people will have named it (more than likely with several words) in all languages (Boudreaux & Ozer, 2015; Funder, 2013; Goldberg, 1993). Thus, it should be possible to identify a common set of constructs that can differentially characterize all humans all over the world (De Raad et al., 1998). The starting point were the thousands of trait descriptive terms provided by Allport and Odbert (1936). Examinations of the factor structure of these terms could then be used to study the structure of personality. This approach led to the model labeled the Big Five by Goldberg (1981). The questionnaire tradition was led by the analysis of questionnaires (McCrae & John, 1992). Eysenck (1947) identified Neuroticism and Extraversion as major dimensions of personality tests. Costa and

McCrae (1976) made a cluster analysis and found a third factor they labeled ‘openness to ideas’. Observing the findings in studies of the lexical tradition, they later added measures of conscientiousness and agreeableness (McCrae & John, 1992). Today, their Five-Factor Model largely represents how we know the five factors: Extraversion, Agreeableness, Conscientiousness, Neuroticism and Openness (to Experience) (Goldberg, 1993; John, 2021). The Big Five model from the lexical tradition is largely similar, though there are some differences. Some are superficial: The fourth factor of the model is labeled Emotional Stability instead, reflecting the opposite pole of Neuroticism, and the third factor is not identical, though very similar, in the two models (Goldberg, 1993). Two other differences are more striking. For one, the first and second factors are systematically rotated, so that the facet Warmth is located in the Extraversion domain in Costa and McCrae’s model, while it belongs in the Agreeableness domain in the lexical model. Another difference lies in the labeling of the fifth factor, coined Openness to Experience in Costa and McCrae’s model, and Intellect or Imagination in the lexical tradition (Goldberg, 1993). Others have suggested using Roman numerals for the domains instead (Hofstee et al., 1997; Saucier & Goldberg, 1996); these are useful because they reflect the relative size of the factors in lexical studies (John, 2021), and they are more theoretically neutral and avoids using labels that can be oversimplified and potentially misleading (Funder, 2013; McCrae & John, 1992). Regardless of the differences between the lexical tradition and the questionnaire tradition, though, the models are mostly similar, and the two terms are often used interchangeably (Visser, 2018).

Analyses of the five factors have been replicated in different languages and cultural groups with remarkably, yet not completely similar results, indicating that they captured somewhat universal dimensions of human personality (De Raad et al., 1998; John, 2021; Saucier & Goldberg, 2001; Saucier et al., 2000). The expression of the domains varies across cultures, with the five factors being well-replicated in Germanic languages, while non-Western languages and cultures tend to be more complex (John, 2021). Still, the results from translating them to different languages have been mostly encouraging so far, with finding the five factors - or at least considerable overlap with them - in most of the languages studied. The fifth factor, Openness, seems to consistently be the factor with the least cultural transmissibility (Funder, 2013; John, 2021).

As for assessment, self-report is by far the most popular basis for measuring personality, and self-report questionnaires remain the most widely used and well-validated tool to assess the five domains, though sometimes peer-report or a combination of the two are used (Funder, 2013; McCrae, 2020). Particularly popular is Costa and McCrae’s NEO

personality inventory (Funder, 2013; McCrae, 2020). Their first version, named NEO-I only contained measures for Neuroticism, Extraversion and Openness - hence the name NEO - and the I for inventory (John, 2021). In 1985 they included all five factors but only measured six facets for the three original factors in their NEO Personality Inventory (NEO-PI) (Costa & McCrae, 1985). In the Revised NEO Personality Inventory - NEO PI-R, they included six facets for all five domains (Costa & McCrae, 1992). In 2005, Costa and McCrae published the NEO PI-3 with revisions of 37 items, making it more useful with younger populations and adults with lower education (McCrae et al., 2005). The NEO PI-R and NEO PI-3 consist of a 240-item questionnaire in both a self-report and an observer-report version. There is also a shorter 60-item NEO Five factor inventory (NEO FFI) version which only measures the five domains, and not the underlying facets (John, 2021). Beyond the NEO inventories, there are several other measuring tools that measure the Big Five. The methods vary from single adjectives in the trait-descriptive adjectives (TDA) (Goldberg, 1992), to long sentence questions used in NEO. But answering 240 or 60 long sentence questions is time consuming, so to address the need for a shorter instrument the Big Five Inventory (BFI) was constructed (John et al., 1991). Like the NEO measurement, it measures the Big Five and 6 underlying facets for each domain - in total 30 facets. There are 8-10 items for each of the five domains, and a total of 44 items. Each facet is measured by only one or two items each. In general, the NEO inventories, the BFI together with the 100-item TDA and the shortened 40-item TDA are the most used measures of the Big Five. The NEO questionnaires are the best-validated but the BFI has been frequently used in research where respondents time is in the essence (John, 2021).

However, the factor structure of the Big Five has been subject to criticism. Some researchers argue that five factors are either too few or too many. One of the most known alternatives to Big Five is the HEXACO model, which consists of 6 factors; Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness and Openness (Ashton & Lee, 2007). In essence HEXACO divides the Big Five domain Agreeableness into an additional factor called Honesty-Humility and rotates some of the factors from their position in the Big Five (Ashton et al., 2014; McCrae, 2020). Even though studies have shown that the Honesty-Humility factor has value, the common argument to defend the Big Five is that it contains both Honesty and Humility since they are roughly the same as straightforwardness and modesty, which places them at a lower level in the Big Five hierarchy (McCrae, 2020). Another alternative to the Big Five model is presented by De Raad and Peabody (2005), who found a more robust support for a three-factor model consisting of

Extraversion, Agreeableness and Conscientiousness, rather than all five factors. Other researchers, however, found that only five factors are consistently replicable (De Raad et al., 1998; John, 2021; Saucier & Goldberg, 1998), though some researchers argue that all of the five factors need redoing (John, 2021). Regardless of the criticisms it has received, the Big Five remains the predominant model of general personality structure (De Raad & Barelds, 2020; John, 2021; McCrae, 2020).

### 1.3. Beyond the Big Five

Despite its name, the Big Five model does not begin and end with the five domains (Wilmot et al., 2016). While the namesakes of the model are very broad and thus are optimal for general use, they fall short in more specific cases, such as predicting a particular life outcome (Ashton et al., 1995). Meanwhile, a given facet of a scale (i.e., a narrower measure) will be more optimal for such specific cases (Anglim et al., 2020; Judge et al., 2002). This phenomenon is often called the bandwidth-fidelity trade-off or the bandwidth-fidelity dilemma, as increased fidelity or increased bandwidth will usually come at a cost to the other (Gleser et al., 1965; Salgado, 2017). So, even though the five domains are important - and have given name to the model - that is not all there is to it.

It is not just the domains and facets within the model that are of interest, though. There have been calls to recognize the complex structure of personality and examine and integrate relationships between the Big Five and other personality traits into the framework. For example, Wilmot et al. (2016) observed how the seemingly independent construct of self-monitoring was actually a bidimensional construct, where the two factors - acquisitive self-monitoring and protective self-monitoring - were located at the level of the metatraits Plasticity (composed of Extraversion and Openness) and Stability (composed of Neuroticism, Agreeableness, and Conscientiousness), respectively. Likewise, Mooradian et al. (2011) examined how dispositional empathy can be mapped onto the Big Five - and they, too, discovered that different dimensions of dispositional empathy are related differently to the five domains.

So, even though The Big Five manages to catch many important subtraits in their model, there are also traits that do not seem to fit so neatly into the equation. While each of the five domains has several facets that it is composed of (e.g., Extraversion is composed of the facets Warmth, Gregariousness, Assertiveness, Activity, Excitement Seeking, and Positive Emotion; Funder, 2013) and encompasses many narrower traits, there are also some traits that do not have a clear affiliation with any one domain. As mentioned earlier, some

aspects of personality may be hard to meaningfully place within this framework entirely; this has been found to be the case with the fantasy facet of dispositional empathy (Mooradian et al., 2011) as well as other personality-related traits such as sensuality, religiosity, thriftiness and many others (Paunonen & Jackson, 2000). Yet other traits - namely, the interstitial ones - may fit in the Big Five framework, only they are not meaningfully related to just one domain but are associated with two or more of the Big Five domains. As mentioned, though, while attempts to clarify what is distinct from the Big Five are widespread (Hough, 1992; Paunonen & Jackson, 2000; Saucier & Goldberg, 1998) research on interstitial traits - at least in relation to the Big Five - are hard to come by.

On the other hand, interstitiality has received some attention in other cases, particularly regarding the lower-order structure of the Personality Inventory for the DSM-5 (PID-5), which is designed to assess the domains and facets of the Alternative Model for Personality Disorders (the AMPD) (Watters et al., 2019). The higher-order structure of the PID-5 is well-established, and in fact generally corresponds well with the Big Five, with at least four of the five factors seemingly corresponding to maladaptive versions of a domain in the Big Five (Negative Affectivity, Detachment, Antagonism, and Disinhibition correspond to Neuroticism, Extraversion, Agreeableness, and Conscientiousness, respectively; Watters et al., 2019). The last factor and domain in each case (psychoticism in the PID-5 and Openness in the Big Five) is less clear - while some of the facets in each appear to relate to one another, it seems that they are best thought of as distinct domains (Chmielewski et al., 2014). Though the higher order structure of the PID-5 seems to be well-established, there has been some inconsistency around the facets of the PID-5 in that some of them appear to be interstitial and that the factor loadings have varied across samples (Watters & Bagby, 2018). This has sparked attempts to clarify the placement of these facets, primarily by Watson, Bagby and their collaborators, and led to the proposition that some facets should be moved from one factor to another (Watters & Bagby, 2018; Watters et al., 2019).

Another model of personality, the interpersonal circumplex, does away with the problem entirely, at least when the interstitial traits are located between only two domains. Gurtman (2009) describes the interpersonal circumplex as a circular space defined by two dimensions usually labeled agency and communion. The implication is that constructs are evenly distributed in this space - contrary to the idea of simple structure models such as the Big Five, which assume few, broad categories in which variables tend to cluster. The two-dimensional space provides a framework in which interpersonal constructs can be mapped onto the model through coordinates (Gurtman, 2009). Thus, the model can be used to study

and differentiate between interpersonal constructs, effectively making an organizing framework as well.

There have been attempts at integrating the Big Five and the circumplex models, called the Abridged Big Five Dimensional Circumplex (AB5C) (Hofstee et al., 1992). By pitting each of the five domains against each other, 10 circumplexes are formed. This makes for a less restrictive model than the simple-structure, hierarchical model that the Big Five is, in which traits at first glance seem to be pure extensions of the domains they are located under. Costa and McCrae (1995) have counter-argued this approach as optimal, citing that facets of the five domains are not of comparable breadth in the AB5C, and that the model fails to discriminate between substantively different facets such as Gregariousness and Positive Emotion. Even so, it at least highlights and provides an easier overview of interstitial variables that lie between two domains (though it does not account for variables that are interstitial between three or more domains) (Hofstee et al., 1992; John, 2021). By combining two of the five NEO PI-R domains into a two-dimensional space and mapping the facets of the two domains onto the circumplex, it becomes apparent that some facets are indeed quite interstitial, even though they belong to only one of the five primary domains; for example, when combining the domains Agreeableness and Extraversion, it is clear that the Extraversion facet labeled Warmth loads positively and almost equally on both domains (John, 2021). In fact, in the lexical tradition, the Warmth facet is indeed located within the Agreeableness domain (Goldberg, 1993). So, even within the Big Five's very own facets, there is interstitiality to be found.

So, even though interstitiality has not been ignored entirely within personality research, not much attention has been devoted to exploring interstitial traits within the Big Five - despite the clear evidence for the prevalence of them, and plenty of authors explicitly acknowledging their existence (Credé et al., 2016; Hofstee et al., 1992; Mooradian et al., 2011). Perhaps this is because it has not been seen as problematic for the framework. Interstitial traits are not necessarily a problem, or even unexpected; in fact, complex models of personality will have traits that are interstitial by nature (Watters & Bagby, 2018). In the case of the PID-5 reviewed above, the interstitial facets were inconsistent across studies, which posed a problem - as it would in any model - due to the implications for the conceptual validity of the AMPD model and the PID-5 measure. So, a clarification of the lower-order structure of the model was needed (Watters & Bagby, 2018). While the Big Five has not had these problems, as the structure of the Big Five has remained consistent across studies (Watters, 2018), the structure of the framework has certainly been studied too. For example,

there has been work on dividing the domains into two aspects (DeYoung et al., 2007), exploring the hierarchical level between facets and domains. Other research has centered on traits above the Big Five, the so-called metatraits (DeYoung, 2006; Digman, 1997; Strus et al., 2014), and some have even argued for the existence of a General Factor of Personality (Musek, 2007). Even though so many levels of the Big Five structure have been studied, it seems that almost no one has cared to look more closely at the interstitial traits in the model, save for the higher-order metatraits.

The latest evidence for the prevalence of interstitial traits comes from a study by Bainbridge et al. (in press). As early as the 1990's, there was considerable consensus that the numerous personality measures available could be described within the Big Five framework, thus providing a map for personality variables, and facilitating meta-analytic cumulation (Ones & Viswesvaran, 1996). Yet despite this early consensus - and despite the relatively common criticism citing the opposite, that the Big Five fails to be comprehensive (Ashton & Lee, 2007; Becker, 1999; Paunonen & Jackson, 2000) - the issue had not been comprehensively evaluated until recently, when Bainbridge et al. (in press) set out to evaluate the model's potential to serve as an organizing framework.

The authors observed that there is an abundance of measures available for assessing narrower traits, which are usually seen as distinct from the Big Five and which are collectively cited much more frequently (Bainbridge et al., in press). The fact that there are so many of these widely used stand-alone scales could indicate that these are considered to lie outside the scope of the Big Five. In this case, it seems that the Big Five is indeed not as all-encompassing a framework of personality as was intended. This observation merited an evaluation of its capacity to be an organizing framework; an evaluation that had not been done. Even though researchers have attempted to place specific traits in the framework, these studies have usually been focused on a single trait and its location in the framework (Mooradian et al., 2011; Roberts et al., 2005), rather than an endeavor to examine how - and if - the model can be used as an organizing framework more generally with numerous trait scales. Upon evaluating this, Bainbridge et al. (in press) decided to look at whether or not the Big Five can actually be used as an organizing framework for personality, including stand-alone scales. They observed that most constructs are highly connected to the Big Five domains, though the connections often operated in a distinct way; namely, interstitially. The stand-alone scales measuring personality trait or trait-like constructs did not simply represent a single of the five broad traits, or even a narrower facet, but rather seemed to blend two or more (Bainbridge et al., in press).

So, despite the clear evidence for and acknowledgement of the prevalence of interstitial traits within the Big Five, not much research has gone into exploring them or their utility to research and practice. This is particularly curious since narrower traits are usually better predictors of specific behaviors and real-life outcomes than broad traits (Anglim et al., 2020; Ashton et al., 1995). If we only utilize the broad traits represented by the Big Five and ignore the narrower traits in-between the five domains, we may unintentionally ignore useful characteristics of human personality and behavior. This may lead us to use high conscientiousness as a marker of job performance, when in fact narrower traits such as integrity or customer service orientation - which are both interstitial - might, at least in specific cases, be better predictors (Credé et al., 2016).

#### **1.4. Summary, rationale, and research question**

During the past 30 or so years, the Big Five model has become the dominating way of conceptualizing personality (Goldberg, 1993). The five domains are highly encompassing of narrower traits in a hierarchical structure and offer much in the way of understanding human personality and behavior, as well as predictive power. Despite its prevalence in the field of personality psychology, however, the model has received criticism as well, mostly regarding exactly its ability to be encompassing. Several authors have put forth traits that seem to be distinct from the Big Five, and Ashton and Lee (2007) have even suggested that a major dimension of personality has been left out in the Big Five framework, instead suggesting a model of six domains, the HEXACO. Still, the Big Five model remains prevalent in personality psychology, and continues to be used for research. It has been used to investigate where other traits are located within the model, and both the lower- and higher-order structure have been subject to examination.

Another thing that has not been investigated as much is the prevalence, importance, and implications of interstitial traits within the Big Five framework. Some research has focused on interstitiality within other models, specifically the Alternative Model for Personality Disorders, while other models such as the interpersonal circumplex attempt to avoid those problems of hierarchical simple-structure models and organize traits in a circular way instead. A combination of circumplex models and the Big Five has been attempted as well; this provides us with 10 circumplexes we can map traits or facets onto and can be a useful tool to illustrate interstitial traits. Even within the Big Five framework itself, there is interstitiality to be found - the Warmth facet of Extraversion is actually equally related to



Agreeableness. However, despite all the evidence of interstitial traits, and plenty of researchers explicitly acknowledging their existence, research on them is sparse.

Recently, more evidence for the prevalence of interstitial traits has come forward. Setting off to evaluate the potential of the Big Five to serve as an organizing framework, Bainbridge et al. (in press) found that the model could indeed be used this way for a majority of stand-alone scales. Most of these scales could be located within the Big Five, though many of them had interstitial relationships with the Big Five.

While some degree of interstitiality is not surprising, as adjective clusters sometimes represent blends of the Big Five domains rather than exclusively centering around the core of one of them, it is nevertheless surprising that so many traits seem to be interstitial. If the Big Five with its five core traits is, as claimed, the most coherent framework for including the numerous adjectives in the language, it would seem intuitive that psychologically meaningful interstitial traits should be the exception rather than the rule. However, the findings by Bainbridge et al. (in press) suggest otherwise; it seems to be common rather than rare that psychologically meaningful traits are interstitial.

So, the question remains: Why is this? A potential explanation of this apparent discrepancy is that many of the putatively interstitial traits might in fact not be as interstitial as they seem; perhaps the scale used for assessing a given trait is merely a confused compilation of items assessing different characteristics rather than a coherent body of items which each assess the same characteristic. In this case, a measure that seems to be a blend of two domains, e.g., Neuroticism and Extraversion, could be composed of 50% of the items measuring Neuroticism and 50% measuring Extraversion, resulting in a construct that seems to be a blend of the two. This would leave that particular scale on shaky ground, as it would seem the scale does not measure a coherent construct after all. Another possibility, though, is that each item in the above example correlates with both Neuroticism and Extraversion, which would indicate that it is indeed a single, coherent construct. If the latter pattern is seen frequently, it would seem to indicate that the scales really do capture narrower, psychologically meaningful traits that are located in the interstices between several domains.

This thesis explores this issue by assessing interstitial trait scales on the item level. Using the Eugene Springfield Community Sample (ESCS), a publicly available dataset, we examine which scales used in this study seem to be interstitial. We then break these scales into their items - and where applicable, facets - to explore why these scales seem to be interstitial, looking at each item in turn and assessing how they relate to the Big Five domains.

We thus seek to answer the following research question: Where does the apparent interstitiality of some scales derive from? Is it that they are collections of individually interstitial items, or are they collections of a diverse set of items which individually represent one or another Big Five domain?

Based on the findings in the study by Bainbridge et al. (in press), we expect that there will be some interstitial scales. Beyond this, there is no particular hypothesis to test and as such we do not have expectations about what the item level correlations will show us.

## 2. Methods

### 2.1. Statistics

All statistical analyses were conducted using IBM SPSS Statistics version 26.

#### 2.1.1. Correlations

The Pearson correlation coefficient  $r$  is a measure of the linear correlation between two variables, i.e., how strong the linear relationship between them is. The value of Pearson's  $r$  lies between -1 and +1. A Pearson's  $r$  closer to either pole (-1 or +1) reflects a stronger linear relationship between the two variables, with a value of (-)1 indicating a perfect correlation; a Pearson's  $r$  close to 1 would indicate a strong, positive correlation, while a Pearson's  $r$  close to -1 would indicate a strong, negative correlation. A Pearson's  $r$  of 0, on the other hand, indicates no correlation between the variables - as long as the assumption of linearity is not violated (Coolican, 2014).

The statistical significance of the correlation is equally important to consider. The significance test, the  $p$  value, describes the probability that the observed result could occur by chance under the null hypothesis. A lower  $p$  value indicates that the result is less likely to occur by chance, and a threshold of 5% (or .05) is usually chosen to determine if a result is statistically significant. The statistical significance largely depends on sample size. In a small sample, the  $r$  value would have to be greater to be statistically significant, while a larger sample can have weaker  $r$  values that are still statistically significant (Coolican, 2014).

Pearson correlations were applied to explore linear relationships between the Big Five domains and other variables: The included scales, as well as their subscales and items.

### 2.2. Scale selection and evaluation

To begin our process of finding scales for comparison with the Big Five domains, we looked through the ESCS technical report (Goldberg, 2008) to find measures of psychological constructs (e.g., self-esteem, optimism). For this, we decided to include

measures of stable traits, but not other psychological constructs (e.g., attitudes). This left us with 41 scales. Using a combination of ESCS syntax files, original articles and a coding sheet provided by our supervisor, we identified the items for each scale, including which ones to reverse score.

After finding all available items, there were times where one or more items were missing from a scale. In those cases, we calculated the total score without them. There were also times where the items had been changed (e.g., the wording had been somewhat altered, or the item had been reversed), or the scoring of the item was different (e.g., the participants in the ESCS scored it on a 5-point Likert scale while the original scale used true/false). These issues have been described separately in the section concerning each scale. For all scales, we excluded participants with missing values. We then scored each scale to get a total score for each where applicable. In some cases, we had to divide the scale into facets as it was not possible to calculate a total score due to the way the measure was designed (e.g., incompatible facets that were not possible to compute into a single total score).

After scoring all scales using total scores, we performed correlations comparing all scales to the five NEO PI-R dimensions. Any scales that had at least a .30 correlation with at least two Big Five domains were interpreted as interstitial and selected for analysis on the item level. Where applicable, the scales were divided into subscales before item-level analysis. The threshold of .30 was chosen as is it a conventional value used to decide, e.g., when a value is worth discussing in the output of a factor analysis (Costello & Osborne, 2005). Though not identical, that is somewhat similar to what we are looking at. Hopwood and Donnellan (2010) even commented in their article specifically on the prevalence of interstitial traits and in that regard noted that “the cross-loadings are typically ‘minor’ from the perspective of EFA studies (e.g., less than .30)” (p. 335). We thus see this as an appropriate, though necessarily arbitrary threshold.

For every scale (and facets of those scales) reaching the .30 threshold with at least two Big Five domains, we computed Cronbach’s Alpha to check the internal consistency of the scale. As Cronbach’s alpha implies internal consistency, but not necessarily homogeneity or unidimensionality of items (Gliem & Gliem, 2003), we also calculated mean inter-item correlations for each interstitial scale and facet, as we were interested specifically in how the items connected with each other. In consideration of the readability only the scales that met our inclusion criteria will be described further in this section. For supplementary information on included scales that was not interpreted as interstitial, see Goldberg (2008).

### 2.3. Participants

The participants were part of the Eugene-Springfield Community Sample, which consists of 30 assessments administered between the summer of 1993 and the fall of 2007. Participants were recruited from a list of homeowners who volunteered to complete their questionnaires for at least five to ten years in return for financial compensation.

The sample originally consisted of 1134 participants. Of these, 46,8% were male and 53,2% were female. The sample consisted mostly of Caucasian participants (98,4%). The age among participants ranged from 18 to 89 years ( $M = 49,67$ ,  $SD = 13,08$ ). Educational background of the participants varied, but 83,9% had at least some college education. Employment status was also varied. For further elaboration on demographics and additional information about the ESCS, see Goldberg (2008).

### 2.4. Procedure

Surveys were distributed over 14 years. Participants were informed that they could decline to answer any question while responding. Each questionnaire was marked with an identification number to ensure participants' anonymity. The researchers obtained ethical approval by the independent ethics committee (IRB) prior to data collection. All data for this study was retrieved from the Harvard Dataverse and is publicly accessible.

The scales used in this study were collected from the following surveys, administered between the summer of 1994 and the fall of 2006: NEO PI-R ( $N = 857$ ), The Behavioral Report Inventory (BRI;  $N = 778$ ), the Personal Attribute Survey (PAS;  $N = 734$ ), A Comprehensive Health Survey (CHS;  $N = 763$ ), Personality, Emotions, and Attitudes (PEA;  $N = 749$ ), Survey of Dispositions and Views (SDV;  $N = 701$ ), Experimental Personality Survey (EPS;  $N = 726$ ), Personal Reactions Survey (PRS;  $N = 736$ ), and Omnibus Personal Attributes Survey (OPAS;  $N = 665$ ).

### 2.5. Measures

The current study was based on questionnaires from the ESCS. Originally, 41 scales from eight different surveys (BRI, PAS, CHS, PEA, SDV, EPS, PRS, and OPAS) were included and compared to the NEO PI-R.

After correlating each of the scales with the five NEO PI-R domains, 11 scales remained that were interstitial, i.e., had a correlation of at least .30 with two or more Big Five domains. These 11 scales came from five different datasets (PAS, CHS, PEA, PRS, and

OPAS). The scales are described in turn below. For information on all 41 scales, we refer to Goldberg (2008).

### **2.5.1. NEO PI-R**

For measuring the Big Five personality traits, the ESCS used the NEO PI-R by Costa and McCrae (1992). The survey consisted of the full 240-item questionnaire, which are used for measuring five domains with 6 facets each: Neuroticism, Extraversion, Agreeableness, Openness, and Conscientiousness. The items were rated on a 5-point Likert scale, from 0 (strongly disagree) to 4 (strongly agree). 104 of the items were reverse scored. The current study obtained acceptable or good internal consistency for all five domains: Neuroticism (48 items,  $\alpha = .85$ ), Extraversion (48 items,  $\alpha = .75$ ), Openness (48 items,  $\alpha = .78$ ), Agreeableness (48 items,  $\alpha = .75$ ), and Conscientiousness (48 items,  $\alpha = .84$ ).

### **2.5.2. BAS Drive**

In the PRS, the ESCS used the four Behavioral Activation System (BAS) and Behavioral Inhibition System (BIS) scales developed by Carver and White (1994) to assess the two motivational systems by the same name developed by Gray (1981); (1982). This assessment includes one BIS scale and three BAS scales: Reward Responsiveness, Drive, and Fun Seeking. Of these, only the BAS Drive scale was interstitial at the .30 level with at least two Big Five domains and are therefore the only one included in this study. This scale includes four items, rated on a 4-point Likert scale (1 = strong agreement, and 4 = strong disagreement), with no neutral response. The PRS included all four items and kept an identical 4-point scale (though switched the numbers; 1 = strong disagreement, 4 = strong agreement). Internal consistency for the BAS Drive scale was acceptable in this study (4 items,  $\alpha = .75$ ), and the mean inter-item correlation was .42.

### **2.5.3. Borderline Personality Inventory**

Borderline personality traits were assessed in the PEA survey using a total of 46 items from the Borderline Personality Inventory (BPI) (Leichsenring, 1999). BPI is a 53-item true-false self-report instrument containing 4 subscales; Identity Diffusion (10 items), Primitive Defenses (8 items), Impaired Reality Testing (5 items), and Fear of Fusion (8 items) (Leichsenring, 1999). In the PEA survey the response format for BPI was converted to a 5-point Likert scale rating from 1 = very inaccurate, to 5 = very accurate. Items that did not fit the Likert scale seem to be the ones that were excluded (e.g. *'I have attempted suicide'*),

though there is no documentation provided by the ESCS about this exclusion. The BPI usually employs a cutoff score using the 20 most discriminatory items (Cut-20), with a recommended cutoff point of  $\text{Cut-20} \geq 10$  (Leichsenring, 1999). We did not employ this cutoff and instead used all 46 included items to calculate a total score. For further analysis, the subscales were computed. The internal consistency for BPI was excellent (46 items,  $\alpha = .90$ ), and the mean inter-item correlation was .18.

#### ***2.5.4. Cognitive Failures Questionnaire***

Cognitive failures were measured in the CHS through the Cognitive Failures Questionnaire (CFQ) developed by Broadbent et al. (1982). The purpose of the CFQ is to assess the frequency of lapses of attention, memory, and cognition in everyday life (Broadbent et al., 1982). The questionnaire consists of 25 items measuring self-reported failures in perception, memory, and motor function during the last six months. Responses are scored on a 5-point scale (from 0 = never, to 4 = very often). All 25 items were included in the CHS and had an identical response format (though ranging from 1 to 5 instead). The internal consistency for the CFQ in this study was good (25 items,  $\alpha = .89$ ), and the mean inter-item correlation was .25.

#### ***2.5.5. Dispositional Optimism***

Dispositional Optimism was measured using the revised Life Orientation Test (LOT-R) by Scheier et al. (1994) as part of the PAS. The LOT-R is used to assess the dispositional level of optimism, and consists of 10 items, of which four are filler items. The PAS included the six test items. The filler items were not included, but the six test items were distributed among other items. Three items are negatively keyed. Respondents score each item on a 5-point Likert scale (from 0 = strongly disagree, to 4 = strongly agree). The scale had similar response wording in the PAS survey (from 1 = very inaccurate; strongly disagree, to 5 = very accurate, strongly agree) and was rated from 1 to 5. In the current study, the internal consistency of the scale was good (6 items,  $\alpha = .82$ ), and the mean inter-item correlation was .44.

#### ***2.5.6. Impression Management***

Impression Management was included in the PAS. Impression Management assesses socially desirable responding where the respondent knows the responses are deliberately inaccurate, more specifically self-presentation tailored to an audience (Paulhus, 1991). The

measure of impression management is a part of the 40-item Balanced Inventory of Desirable Responding (BIDR). 20 of these items measure impression management, 10 of which are negatively keyed. The items are originally scored on a 7-point scale, where points are only given for a response of 6 or 7 - if a participant chooses 1-5, no points are given for that item (Paulhus, 1991). We instead scored it continuously, including all responses. In the PAS, all items were included and scored on a 5-point Likert scale from 1 = very inaccurate; strongly disagree, to 5 = very accurate; strongly agree. The internal consistency of the impression management measure in this study was good (20 items,  $\alpha = .82$ ), and the mean inter-item correlation was .18.

### ***2.5.7. Levenson Self-Report Psychopathy Scale***

The Levenson Self-Report Psychopathy Scale (LSRP) consists of a total of 26 items. The scale was developed for use in research and measures psychopathy on two subscales: Primary Psychopathy (psychopathic emotional affect) and Secondary Psychopathy (psychopathic lifestyle) with 16 items measuring the primary subscale, and 10 items measuring the secondary subscale. The response format is a 5-point Likert scale (Levenson et al., 1995). In the ESCS, the LSRP was a part of the PEA questionnaire where 25 items were used. Items were rated on a 5-point Likert scale (from 1 = very inaccurate, to 5 = very accurate). The item *'My main purpose in life is getting as many goodies as I can'* from the primary facet was missing for reasons unknown. However, later research has implied that occasional low reliability of the LSRP could be due to certain items, as removing them increased the reliability (Gummelt et al., 2012). After calculating correlations between the full scale and the Big Five, the scale was divided into its two facets for further correlations with the Big Five domains. In the current study, internal consistency for all LSRP items was good (25 items,  $\alpha = .80$ ), and the mean inter-item correlation was .15.

### ***2.5.8. Narcissism Personality Inventory***

Narcissism was measured in the OPAS using the 40-item Narcissism Personality Inventory (NPI-40) (Raskin & Terry, 1988). The NPI-40 has yet to reach an agreed psychometric structure, however, and its factor structure has been disputed and reassessed several times (Corry et al., 2008; Kubarych et al., 2004; Rosenthal et al., 2011). The most widely accepted single factor structure seems to be the 16-item version (Ames et al., 2006), which is why it was chosen for the current study. The NPI-16 is considered a measure for subclinical narcissism and consists of 16 of the 40 NPI-40 items. Originally, the scale was

constructed with a forced-choice response format with each item consisting of a pair of statements that participants had to choose between; one statement that is considered narcissistic and one that is considered non-narcissistic. In the OPAS, the response format was converted into a 5-point Likert scale ranging from 1 = very inaccurate to 5 = very accurate, using only the narcissistic statement. The internal consistency for the NPI-16 was good (16 items,  $\alpha = .83$ ), and the mean inter-item correlation was .23.

### ***2.5.9. Rosenberg Self-Esteem Scale***

The Rosenberg Self-Esteem Scale is a unidimensional scale that measures global self-worth by measuring both positive and negative feelings about the self. The scale includes 10 items, five of which are negatively keyed. The items are rated on a 4-point scale from 1 = strongly disagree, to 4 = strongly agree, with no neutral option (Rosenberg, 1979). In the ESCS, this scale was included in the PAS survey, and a 5-point scale was used instead; from 1 = very inaccurate; strongly disagree, to 5 = very accurate; strongly agree, including the option to give a neutral response. All 10 items were included in the PAS. Internal consistency for this scale in the current study was good (10 items,  $\alpha = .87$ ), and the mean inter-item correlation was .43.

### ***2.5.10. Self-Deception***

Self-Deception was measured in the PAS. The Self-Deception measure is a part of the 40-item Balanced Inventory of Desirable Responding (BIDR) along with Impression Management. Self-Deception, also sometimes labeled Self-Deceptive Enhancement, assesses non-deliberate socially desirable responding, i.e. an overly positive self-presentation where the respondent believes the desirable responses are accurate self-reports. 20 out of the 40 BIDR items measure self-deception, and 10 of these are reverse coded. The items are originally scored on a 7-point scale, where points are only given for a response of 6 or 7 - if a participant chooses 1-5, no points are given for that item (Paulhus, 1991). As with the other half of the BIDR inventory, we calculated the full score including all responses. In the PAS, all items were included and were scored on a 5-point Likert scale from 1 = very inaccurate; strongly disagree, to 5 = very accurate; strongly agree. The Self-Deception measure had questionable internal consistency in the ESCS (20 items,  $\alpha = .68$ ), and the mean inter-item correlation was .10.



### ***2.5.11. Sensitivity to Punishment***

The Sensitivity to Punishment scale is a part of The Sensitivity to Punishment and Sensitivity to Reward Questionnaire developed and later revised by Torrubia et al. (2001) to assess BIS and BAS functioning, respectively. Only the Sensitivity to Punishment turned out to be correlated at the .30 level with at least two Big Five domains. The revised Sensitivity to Punishment scale consists of 24 items; 18 from the old version along with six new items. Respondents score these questions by ticking either 'yes' or 'no'. 18 of these items were included in the OPAS. It is unclear why six items were excluded. Instead of a binary response form, the OPAS used a 5-point scale (from 1 = very inaccurate, to 5 = very accurate). The Sensitivity to Punishment scale had good internal consistency in this study (18 items,  $\alpha = .88$ ), and the mean inter-item correlation was .29.

### ***2.5.12. Toronto Alexithymia Scale***

Alexithymia refers to problems identifying and describing emotions and tendency to minimize emotional experience and focus attention externally (Taylor et al., 1985). The twenty-item Toronto Alexithymia Scale (TAS) are based on three facets, difficulty describing emotions, difficulty identifying emotions, and externally oriented thinking. Whereas the difficulty describing feelings facet consists of 5 items, the difficulty identifying feelings facet consists of 7 items, and the externally oriented thinking facet consists of 8 items (Taylor et al., 1985). All items are rated using a 5-point Likert scale, from 1 = strongly disagree, to 5 = strongly agree. Five items are negatively keyed. All 20 items were administered as a part of the PEA, using the original response format. The total score was used to look for interstitially before computing scores for the three facets. Internal consistency for the Toronto Alexithymia Scale was acceptable (20 items,  $\alpha = .78$ ), and the mean inter-item correlation was .16.

## **3. Results**

Results presented in this section will be shown in tables, and the main results will be reported in text as well. The initial analyses of the 41 included scales and their correlations with the Big Five domains will be presented first, followed by individual correlation analyses between interstitial scales and the Big Five.

### 3.1. Initial analyses

Table 1 below shows our initial results, which showed that 11 scales had a correlation of at least .30 with at least two Big Five. Of these, seven scales correlated with two domains (BAS Drive, Borderline Personality Inventory, Cognitive Failures Questionnaire, Dispositional Optimism, Narcissism Personality Inventory, Self-deception, and Sensitivity to Punishment), while four correlated with three domains (Impression Management, Levenson Self- Report Psychopathy Scale, Rosenberg Self-Esteem Scale, and the Toronto Alexithymia Scale). These will be reviewed in turn below. The rest of the 41 scales correlated primarily with one domain, or in a few cases did not correlate substantively with any of the five domains at all. The correlation matrix of all 41 scales compared to the Big Five domains can be viewed in Table 1. This table also shows the intercorrelations between the Big Five domains themselves. Four domains correlated moderately with another domain: Extraversion correlated moderately positively with Openness  $r = .37, p < .001$ , and Neuroticism correlated moderately negatively with Conscientiousness,  $r = -.47, p < .001$ . Neuroticism further had a weak to moderate negative correlation with Extraversion,  $r = -.27, p < .001$ , and with Agreeableness,  $r = -.21, p < .001$ .

**Table 1**  
*Pearson correlation matrix between included scales and Big Five domains.*

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
Neuroticism	1	-.270**	-.022	-.211**	-.466**
Extraversion	-.270**	1	.336**	.048	.177**
Openness to Experience	-.022	.336**	1	.021	-.130**
Agreeableness	-.211**	.048	.021	1	.154**
Conscientiousness	-.466**	.177**	-.130**	.154**	1
Dissociative experiences	.322**	-.003	.177**	-.144**	-.207**
Self-esteem	-.586**	.320**	.157**	.060	.365**
Self-monitoring	.293**	-.298**	-.078*	-.160**	-.238**
Locus of control, internal	-.244**	.144**	.015	-.002	.224**
Locus of control, chance	.302**	-.183**	-.106**	-.074	-.235**
Locus of control, powerful others	.300**	-.245**	-.220**	-.105**	-.074
Private self-consciousness	.296**	.029	.245**	-.121**	-.180**
Public self-consciousness	.430**	.034	-.051	.000	-.116**
Dispositional Optimism	-.562**	.355**	.143**	.170**	.280**
Self-deception	-.559**	.243**	.010	.000	.420**
Impression management	-.378**	.000	-.088*	.445**	.324**
OCD Total	.373**	-.148**	-.239**	-.090*	-.028

Irrational Beliefs total	,125**	,119**	,268**	,041	-,157**
Cognitive failures total	<b>,452**</b>	-,209**	-,050	-,019	<b>-,383**</b>
Need for cognition total	-,139**	,198**	<b>,450**</b>	-,214**	,146**
Health related coping style, distraction	,054	,239**	,297**	,118**	-,044
Health related coping style, palliative	,099*	-,037	,092*	,038	-,041
Health related coping style, instrumental	,028	,058	,140**	,091*	,059
Health related coping style, emotional preoccupation	<b>,320**</b>	-,011	,067	-,067	-,110**
Toronto Alexithymia Scale total	<b>,346**</b>	<b>-,392**</b>	<b>-,329**</b>	-,098*	-,276**
Somatoform Dissociation total	<b>,324**</b>	-,162**	-,088*	-,146**	-,204**
Fantasy Proneness total	,270**	,068	<b>,430**</b>	-,052	-,199**
Hypochondriasis total	<b>,458**</b>	-,213**	-,046	-,186**	-,245**
Magical Beliefs total	,213**	,088'	,222**	-,104**	-,182**
Borderline Personality inventory	<b>,538**</b>	-,125**	,021	-,263**	<b>-,332**</b>
Right-Wing Authoritarianism	,071	-,082*	<b>-,549**</b>	,113**	,135**
Life satisfaction	<b>-,426**</b>	,259**	,013	,145**	,231**
Social Dominance	-,010	-,084*	<b>-,324**</b>	-,291**	,052
Expressions of Spirituality Inventory	,043	,155**	,104**	,220**	-,064
CES-Depression Scale	<b>,478**</b>	-,228**	,020	-,145**	-,254**
Behavioral Inhibition System (BIS)	<b>,531**</b>	-,141**	-,041	,048	-,116**
BAS Reward Responsiveness	,148**	,262**	,125**	,067	,028
BAS Drive	,034	<b>,307**</b>	,172**	<b>-,300**</b>	,069
BAS Fun Seeking	,156**	<b>,346**</b>	,280**	-,125**	-,222**
Machiavellianism	,210**	-,110**	,023	<b>-,358**</b>	-,225**
Narcissism Personality Inventory (NPI16)	-,124**	<b>,381**</b>	,134**	<b>-,331**</b>	,182**
Psychopathy	,130**	,145**	,151**	<b>-,446**</b>	-,193**
Fear Total	<b>,312**</b>	-,149**	-,198**	,097*	-,091*
Sensitivity to Reward total	,120**	,287**	,008	<b>-,391**</b>	-,088*
Sensitivity to Punishment total	<b>,513**</b>	<b>-,372**</b>	-,203**	,089*	-,289**
Levenson Psychopathy Scale	<b>,377**</b>	-,090*	-,102**	<b>-,524**</b>	<b>-,317**</b>

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2. Item-correlations

As for evaluation of the item correlations with the Big Five, a benchmark of .30 (which we used for scale-level assessment) is likely too high to be reasonable for assessment of interstitiality on the item level, given that items have so much less variance than the full scale they are a part of. We are, however, not aware of any established methodology for determining which threshold to use. Reducing this threshold to .20 has some intuitive appeal. Further, we found some informal empirical support for this. In order to provide ourselves

with a benchmark for the item correlations, we collected the three purest scales we found in the initial correlations with the Big Five and compared the scale's correlation with their primary domain to the scale items' correlations with the same domain. A 'pure' scale in this sense would be a scale that is primarily located within a single domain without too high correlations with the other four domains; in essence, the opposite of an interstitial scale. These three were Right-Wing Authoritarianism, which were primarily correlated with Openness, and the Behavioral Inhibition System scale (BIS) and Public Self-consciousness scale, both primarily correlated with Neuroticism. Analyzing these, we found that the average correlation between items from these scales and the corresponding domain (Openness for Right-wing authoritarianism items, Neuroticism for BIS and Public Self-consciousness items) was reduced by just over a third compared to the result found for the total scale in question. Due to this, and given the absence of any alternative justification for a specific criterion level that we are aware of, it seems reasonable to lower the bar for what magnitude of correlation might be interpreted as indicative of a noteworthy relationship - from .30 (the scale-level threshold) to .20 for the item level. For interpretation of scales with subscales or facets, we continued to use a threshold of .30 to evaluate the relationships between subscales and domains; for those subscales exceeding the .30 threshold, we then applied the .20 threshold to the items of those subscales. We use this to present our findings for the items, though we will look at the results in more detail in the discussion and offer a more nuanced view.

### **3.2.1. BAS drive**

The BAS Drive subscale was found to correlate above the threshold with two domains, correlating positively with Extraversion,  $r = .31, p < .001$ , and negatively with Agreeableness,  $r = -.30, p < .001$ . A total number of 654 participants completed both the BAS Drive and the NEO PI-R questionnaire.

All four items showed connections above the threshold with at least one Big Five domain. Two items showed the same relationship as the complete scale with correlations above the threshold with both Extraversion and Agreeableness, while the two others were only associated with Extraversion.

**Table 2**  
Pearson correlation matrix between BAS drive items and Big Five domains.

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
When I want something, I usually go all-out to get it	-,016	<b>,261**</b>	,158**	-,194**	,137**
I go out of my way to get things I want	,054	<b>,212**</b>	,198**	<b>-,262**</b>	-,008
When I go after something, I use a “no holds barred” approach	,071	<b>,250**</b>	,142**	<b>-,269**</b>	-,002
If I see a chance to get something I want, I move on it right away	-,006	<b>,204**</b>	,020	-,181**	,084*

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2.2. Borderline Personality Inventory

The Borderline Personality Inventory was found to correlate above the threshold with two Big Five domains, correlating positively with Neuroticism,  $r = .54, p < .001$ , and negatively with Conscientiousness,  $r = -.38, p < .001$ . Dividing the BPI into its four subscales only the Identity Diffusion correlated above .30 with more than one of the Big Five domains, The remaining three facets did only correlate above the threshold with a maximum of one Big Five domain. The internal validity for the Identity Diffusion subscale was acceptable (9 items,  $\alpha = .75$ .) and the mean inter-item correlation for these nine items was .40. A total number of 857 participants completed both the BPI and the NEO PI-R questionnaire.

**Table 3**  
Pearson correlation matrix between the facets of Borderline Personality inventory and Big Five domains.

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
Borderline Personality inventory	<b>,538**</b>	-,125	,021	-,263	<b>-,332**</b>
BPI Indetify Diffusion	<b>,441**</b>	-,113**	,094*	-,140**	<b>-,314**</b>
BPI Primitive Defenses	<b>,639**</b>	-,227**	-,095*	-,215**	-,281**
BPI Impaired Reality Testing	,222**	-,073	-,008	-,095*	-,145**
BPI Fear of Fusion	<b>,420**</b>	-,286**	-,062	-,271**	-,252**

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

Of the nine included items for the Identity Diffusion subscale, six were associated with at least one domain above the threshold. Four of these showed the same relation to The Big Five as the Identity Diffusion facet, and one was only connected to one of the relevant domains.

**Table 4**  
Pearson correlation matrix between the Identity Diffusion facet items and Big Five domains.

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
Feel that people and things around me are not real	<b>,286**</b>	-,115**	-,083*	<b>-,225**</b>	<b>-,246**</b>
Act or feel in a way that does not fit me	<b>,168**</b>	,073	<b>,329**</b>	,007	-,147**
Have felt the presence of another person when he was not there	<b>,158**</b>	-,056	,048	-,081*	-,137**
Feel like I am "Falling apart"	<b>,267**</b>	,003	,128**	-,136**	-,176**
Find it difficult to tell whether something really happened..	<b>,175**</b>	-,066	-,060	-,011	-,064
Feel a sense of not being real	<b>,440**</b>	-,163**	-,041	-,092*	<b>-,241**</b>
Feel that my body is dissolving or that a part of my body	<b>,276**</b>	-,082*	-,005	-,040	<b>-,230**</b>
Don't know what I really want	<b>,194**</b>	-,077*	,008	-,104**	-,131**
Feel that I'M living in a dream, or see my life before..	<b>,381**</b>	-,175**	,009	-,121**	<b>-,293**</b>

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2.3. Cognitive Failures Questionnaire

The Cognitive Failures Questionnaire was found to be correlated higher than the threshold with two domains, correlating positively with Neuroticism,  $r = .45$ ,  $p < .001$ , and negatively with Conscientiousness,  $r = -.38$ ,  $p < .001$ . A total number of 760 participants completed both the CFQ and the NEO PI-R questionnaire.

Out of 25 total items, 19 correlated above the threshold with a minimum of one Big Five domain. Eight items showed adequate relation to both of the involved Big Five domains as the scale itself while nine items were only related to one of them. The remaining two were related to two Big Five domains, of which only one domain was consistent with the whole scale.

**Table 5**  
Pearson correlation matrix between the Cognitive Failure items and Big Five domains.

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
Read something and find you haven't been thinking about it and must read	<b>,287**</b>	-,137**	-,038	,072	<b>-,259**</b>
Find you forget why you went from one part of the house to the other?	<b>,209**</b>	-,029	-,016	,101**	-,141**
Fail to notice signposts on the road?	<b>,171**</b>	-,088*	-,072	,021	-,177**
Find you confuse right and left when giving directions?	<b>,151**</b>	-,101**	-,058	,090*	-,124**
Bump into people?	<b>,245**</b>	-,063	-,056	-,022	-,163**
Find you forgot whether you've turned off a light or locked the door?	<b>,247**</b>	-,150**	-,073	,006	-,177**
Fail to listen to people's names when you are meeting them?	<b>,151**</b>	-,110**	-,021	-,118**	-,115**
Say something and realize afterwards that it might be insulting?	<b>,263**</b>	-,059	-,061	-,140**	-,142**

Fail to hear people speaking to you when you are doing something else?	,113**	-,052	-,052	-,049	-,060
Lose your temper and regret it?	,333**	,024	,037	-,212**	-,133**
Leave important letters unanswered for days?	,235**	-,135**	,056	-,045	-,369**
Find you forget which way to turn on a road you know well but rarely?	,145**	-,111**	,006	,029	-,203**
Fail to locate what you want in a supermarket (although it's there)?	,160**	-,172**	-,128**	-,014	-,188**
Find yourself suddenly wondering whether you've used a word correctly?	,227**	-,079*	-,128**	,006	-,112**
Have trouble making up your mind?	,387**	-,180**	-,067	,073	-,261**
Find you forget appointments?	,219**	-,036	,037	-,102**	-,309**
Forget where you put something like a newspaper or a book?	,180**	-,068	,040	-,003	-,263**
Find that you accidentally throw away the thing you want and keep what	,231**	-,022	-,031	,044	-,149**
Daydream when you ought to be listening to something?	,268**	-,082*	,066	-,077*	-,311**
Forget people's names?	,143**	-,139**	-,018	-,042	-,131**
Start doing one thing at home and get distracted into doing smtg else	,305**	-,069	,062	,013	-,304**
Find you can't quite remember something although it's on tip of your?	,249**	-,125**	-,008	,059	-,217**
Forget what came to the store to buy?	,262**	-,081*	-,015	-,010	-,324**
Drop things?	,299**	-,154**	,005	,000	-,176**
Find you can't think of anything to say?	,344**	-,406**	-,153**	,044	-,156**

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2.4. Dispositional Optimism

The total score for Dispositional Optimism correlated negatively with Neuroticism,  $r = -.56, p < .001$ , and positively with Extraversion,  $r = .36, p < .001$ . A total number of 663 participants completed both the Dispositional Optimism and the NEO PI-R questionnaire.

All six items were related above the threshold to at least one Big Five domain. All six items correlated above the threshold with the same two domains as the scale, but only one of these did that exclusively. The remaining five also showed a connection above the threshold with either Conscientiousness or Openness.

**Table 6**

*Pearson correlation matrix between the Dispositional Optimism items and Big Five domains.*

	Openness to				
	Neuroticism	Extraversion	Experience	Agreeableness	Conscientiousness
In uncertain times, I usually expect the best.	-,393**	,267**	,065	,174**	,189**
If something can go wrong for me, it will.	,442**	-,227**	-,096*	-,091*	-,219**
I'm always optimistic about my future.	-,428**	,230**	,025	,096*	,204**
I hardly ever expect things to go my way.	,454**	-,268**	-,118**	-,095*	-,248**
I rarely count on good things happening to me.	,374**	-,305**	-,201**	-,125**	-,171**
Overall, I expect more good things to happen to me than bad.	-,380**	,268**	,138**	,156**	,207**

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2.5. Impression Management

The total score for Impression Management correlated above the threshold with three domains, negatively with Neuroticism,  $r = -.38, p < .001$ , and positively with Agreeableness,  $r = .45, p < .001$ , and Conscientiousness,  $r = .32, p < .001$ . A total number of 664 participants completed both the Impression management and the NEO PI-R questionnaire.

Out of 20 items, 11 correlated above the .20 threshold with at least one Big Five domain. Two items correlated above the threshold with the three relevant domains. Four items correlated with two of the domains and five items correlated with only one of the relevant domains.

**Table 7**

*Pearson correlation matrix between the Impression Management items and Big Five domains.*

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
I sometimes tell lies if I have to.	<b>,256**</b>	-,061	,077*	<b>-,260**</b>	<b>-,236**</b>
I never cover up my mistakes.	<b>-,168**</b>	,064	-,043	,053	,106**
There have been occasions when I have taken advantage of someone.	<b>,204**</b>	-,005	,059	<b>-,369**</b>	<b>-,222**</b>
I never swear.	<b>-,086*</b>	-,062	<b>-,181**</b>	<b>,202**</b>	,124**
I sometimes try to get even rather than forgive and forget.	<b>,233**</b>	-,031	-,041	<b>-,405**</b>	-,044
I always obey laws, even if I'm unlikely to get caught.	<b>-,125*</b>	-,064	<b>-,110**</b>	<b>,236**</b>	,121**
I have said something bad about a friend behind his or her back.	<b>,235**</b>	,032	,000	<b>-,199**</b>	<b>-,134**</b>
When I hear people talking privately, I avoid listening.	<b>-,177**</b>	<b>-,077*</b>	-,055	<b>,313**</b>	,102**
I have received too much change from a salesperson without telling him a	<b>,208**</b>	-,057	-,002	<b>-,264**</b>	<b>-,169**</b>
I always declare everything at customs.	<b>-,188**</b>	,030	,000	<b>,167**</b>	<b>,173**</b>
When I was young I sometimes stole things.	<b>,120**</b>	-,059	-,024	<b>-,181**</b>	<b>-,112**</b>
I have never dropped litter on the street.	<b>-,143**</b>	-,012	,006	<b>,187**</b>	<b>,147**</b>
I sometimes drive faster than the speed limit.	,058	,011	-,004	<b>-,149**</b>	-,056
I never read sexy books or magazines.	<b>-,091*</b>	<b>-,079*</b>	<b>-,119**</b>	<b>,237**</b>	,145**
I have done things that I don't tell other people about.	,121**	<b>-,092*</b>	,031	<b>-,118**</b>	<b>-,078*</b>
I never take things that don't belong to me.	<b>-,130**</b>	-,028	-,022	<b>,156**</b>	,141**
I have taken sick-leave from work or school even though I wasn't really	,181**	,023	<b>,175**</b>	<b>-,144**</b>	<b>-,256**</b>
I have never damaged a library book or store merchandise without reporti	<b>-,112**</b>	,068	-,017	<b>,143**</b>	<b>,147**</b>
I have some pretty awful habits.	<b>,320**</b>	-,059	-,026	<b>-,185**</b>	<b>-,293**</b>
I don't gossip about other people's business.	<b>-,281**</b>	<b>-,102**</b>	-,040	<b>,200**</b>	,141**

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2.6. Levenson Self-Report Psychopathy Scale

The Levenson Psychopathy Scale reached the .30 threshold with three domains, being correlated positively with Neuroticism,  $r = .38, p < .001$ , and negatively with Agreeableness,



$r = -.52, p < .001$ , and Conscientiousness,  $r = .32, p < .001$ . When dividing the scale into its Primary and Secondary subscales, only the Secondary Subscale correlated higher than the threshold with two domains, positively with Neuroticism,  $r = .54, p < .001$ , and negatively with Conscientiousness,  $r = -.48, p < .001$ . The Secondary Subscale had a questionable internal consistency (10 items,  $\alpha = .66$ ), and the mean inter-item correlation for the Secondary subscale was .17. A total number of 665 participants completed both the LSRP and the NEO PI-R questionnaire.

**Table 8**

*Pearson correlation matrix between the Leveson Psychopathy facets and Big Five domains.*

	Openness to				
	Neuroticism	Extraversion	Experience	Agreeableness	Conscientiousness
Levenson Psychopathy scale total	<b>,377**</b>	-,090*	-,102**	<b>-,524**</b>	<b>-,317**</b>
Levenson Psychopathy Primary	<b>,138**</b>	,024	-,067	<b>-,553**</b>	-,102**
Levenson Psychopathy Secondary	<b>,544**</b>	-,204**	-,110**	<b>-,282**</b>	<b>-,476**</b>

\*\**. Correlation is significant at the 0.01 level (2-tailed)*

\**. Correlation is significant at the 0.05 level (2-tailed)*

All 10 items for the Secondary Subscale correlated above the threshold with at least one Big Five domain. Four of these matched the same connection to the Big Five as the Secondary Subscale, while two showed an adequate relation to one of the involved domains. Two items showed a connection to a relevant domain along with a correlation above the threshold with Agreeableness.

**Table 9**

*Pearson correlation matrix between the Secondary facet items and Big Five domains.*

	Openness to				
	Neuroticism	Extraversion	Experience	Agreeableness	Conscientiousness
Think that most of my problems are due..	<b>,319**</b>	-,141**	-,145**	-,098*	-,045
Quickly lose interest in the tasks I start	<b>,297**</b>	-,145**	-,042	-,113**	<b>-,393**</b>
Find myself in the same kinds of trouble, time after time	<b>,405**</b>	-,071	,065	-,152**	<b>-,306**</b>
Don't plan anything very far in advance	<b>,195**</b>	-,192**	-,044	-,002	<b>-,302**</b>
Am often bored	<b>,329**</b>	-,191**	-,180**	-,155**	<b>-,222**</b>
Find that I am able to pursue one goal for a long time	<b>-,230**</b>	,117**	,100**	,057	<b>,321**</b>
Get in shouting matches with other people	<b>,296**</b>	-,005	-,022	<b>-,312**</b>	-,154**
Think that love is overrated	<b>,177**</b>	-,187**	-,166**	<b>-,229**</b>	-,081*
When I get frustrated, I often let off steam by blowing my top	<b>,324**</b>	,039	,001	<b>-,252**</b>	-,126**
Before I do anything, I carefully consider the possible..	-,113**	-,009	-,087*	,049	<b>,386**</b>

\*\**. Correlation is significant at the 0.01 level (2-tailed)*

\**. Correlation is significant at the 0.05 level (2-tailed)*

### 3.2.7. Narcissism Personality Inventory (NPI-16)

The NPI-16 total score correlations reached the threshold with two domains, with positive correlations with Extraversion,  $r = .38, p < .001$ , and negative with Agreeableness,  $r$

= -.33,  $p < .001$ . A total number of 585 participants completed both the NPI-16 and the NEO PI-R questionnaire.

Of the 16 items 12 showed a connection above the threshold with at least one Big Five domain. Four shared the same connection to the Big Five domains as the complete scale. Seven items correlated with just one of the relevant domains, however three of these also correlated above the threshold with a domain not shared by the scale as a whole.

**Table 10**

*Pearson correlation matrix between the NPI-16 items and Big Five domains.*

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
I know I'm good because everybody keeps telling me so	-,015	,097*	-,069	-,012	,086*
I like to be the center of attention	,013	,286**	,117**	-,261**	-,031
I think I am a special person	-,247**	,309**	,108**	-,056	,201**
I like to have authority over other people	-,027	,270**	,022	-,285**	,144**
I find it easy to manipulate people	-,031	,230**	,137**	-,336**	-,025
I insist upon getting the respect that is due me	,065	,134**	,047	-,158**	,064
I will usually show off if I get the chance	,089*	,197**	,117**	-,212**	-,097*
I always know what I am doing	-,242**	-,055	-,142**	,013	,258**
Everybody likes to hear my stories	-,115**	,231**	,080	-,074	,085*
I expect a great deal from other people	,012	,134**	,109**	-,144**	,133**
I really like to be the center of attention	,028	,276**	,093*	-,267**	-,013
People always seem to recognize my authority	-,164**	,252**	,033	-,168**	,214**
I am going to be a great person	-,129**	,300**	,068	-,109**	,177**
I can make anybody believe anything I want them to	-,020	,164**	,111**	-,184**	,005**
I am more capable than other people	-,133**	,115**	,098*	-,294**	,220**
I am an extraordinary person	-,204**	,243**	,195**	-,146**	,185

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2.8. Rosenberg Self-Esteem

The Rosenberg Self-Esteem scale correlated higher than the threshold with three domains, correlating negatively with Neuroticism,  $r = -.59$ ,  $p < .001$ , and positively with Extraversion,  $r = .32$ ,  $p < .001$ , and Conscientiousness,  $r = .37$ ,  $p < .001$ . A total number of 665 participants completed both the Rosenberg Self-Esteem and the NEO PI-R questionnaire.

All 10 items correlated above the threshold with at least one Big Five domain; five with the same three domains as the scale, while four items were related to two of the three involved domains - although one of these also correlated above the threshold with Openness. The last item correlated above the threshold with only one of the involved domains.

**Table 11**

*Pearson correlation matrix between the Rosenberg Self-Esteem items and Big Five domains.*

	Openness to				
	Neuroticism	Extraversion	Experience	Agreeableness	Conscientiousness
I feel that I am a person of worth, at least on an equal basis with othe	<b>-,292**</b>	,145**	,161**	,001	<b>,201**</b>
I feel that I have a number of good qualities.	<b>-,276**</b>	,177**	,193**	,095	,180**
All in all, I am inclined to feel that I am a failure.	<b>,404**</b>	<b>-,264**</b>	-,097*	-,035	<b>-,275**</b>
I am able to do things as well as most other people.	<b>-,241**</b>	,191**	<b>,223**</b>	-,084*	<b>,224**</b>
I feel I do not have much to be proud of.	<b>,289**</b>	<b>-,232**</b>	-,133**	-,027	-,185**
I take a positive attitude toward myself.	<b>-,475**</b>	<b>,295**</b>	,131**	,002	<b>,287**</b>
On the whole, I am satisfied with myself.	<b>-,480**</b>	<b>,285**</b>	,085*	,059	<b>,304**</b>
I wish I could have more respect for myself.	<b>,502**</b>	<b>-,229**</b>	-,070	-,122**	<b>-,318**</b>
I certainly feel useless at times.	<b>,509**</b>	<b>-,241**</b>	-,047	-,059	<b>-,305**</b>
At times I think I am no good at all.	<b>,484**</b>	-,194**	-,059	-,055	<b>-,244**</b>

\*\* Correlation is significant at the 0.01 level (2-tailed)

\* Correlation is significant at the 0.05 level (2-tailed)

### 3.2.9. Self-Deception

Self-Deception correlated above the threshold with two domains, negatively with Neuroticism  $r = -.56, p < .001$  and positively with Conscientiousness  $r = .42, p < .001$ . A total number of 663 participants completed both the Self-Deception and the NEO PI-R questionnaire.

Out of 20 items, 14 correlated above the threshold with at least one Big Five domain, of which five items showed a connection to the same two domains as the scale - of these, two also correlated above the threshold with a third domain. Eight items correlated above the threshold with just one of the relevant domains, of which one also correlated with Extraversion.

**Table 12**

*Pearson correlation matrix between the Self-deception items and Big Five domains.*

	Openness to				
	Neuroticism	Extraversion	Experience	Agreeableness	Conscientiousness
My first impressions of people usually turn out to be right.	-,064	,029	,030	-,029	,138**
It would be hard for me to break any of my bad habits.	<b>,211**</b>	-,100**	-,036	-,129**	-,168**
I don't care to know what other people really think of me.	-,050	-,121**	-,070	-,123**	-,033
I have not always been honest with myself.	<b>,260**</b>	-,070	,037	-,022	<b>-,240**</b>
I always know why I like things.	-,136**	,098*	-,016	,051	<b>,216**</b>
When my emotions are aroused, it biases my thinking.	<b>,201**</b>	-,042	,117**	-,060	-,148**
Once I've make up my mind, other people can seldom change my opinion.	-,085*	-,056	-,107**	-,157**	,170**
I am not a safe driver when I exceed the speed limit.	,066	-,110**	-,015	,182**	-,064
I am fully in control of my own fate.	<b>-,229**</b>	,136**	,078*	-,069	,109**

It's hard for me to shut off a disturbing thought.	<b>,382**</b>	<b>-,115**</b>	<b>-,128**</b>	<b>-,031</b>	<b>-,093**</b>
I never regret my decisions.	<b>-,250**</b>	<b>,132**</b>	<b>-,033</b>	<b>-,035</b>	<b>,134</b>
I sometimes lose out on things because I can't make up my mind soon enou	<b>,348**</b>	<b>-,257**</b>	<b>-,141**</b>	<b>-,008</b>	<b>-,303**</b>
The reason I vote is because my vote can make a difference.	<b>-,187**</b>	<b>,117**</b>	<b>,105**</b>	<b>,165**</b>	<b>,139**</b>
My parents were not always fair when they punished me.	<b>,213**</b>	<b>-,047</b>	<b>,113**</b>	<b>-,056</b>	<b>-,124**</b>
I am a completely rational person.	<b>-,282**</b>	<b>,073</b>	<b>-,191**</b>	<b>,112**</b>	<b>,281**</b>
I rarely appreciate criticism.	<b>,207**</b>	<b>-,181**</b>	<b>-,083*</b>	<b>-,066</b>	<b>-,058</b>
I am very confident of my judgments.	<b>-,430**</b>	<b>,266**</b>	<b>,053</b>	<b>-,066</b>	<b>,385**</b>
I have sometimes doubted my ability as a lover.	<b>,223**</b>	<b>-,227**</b>	<b>-,128**</b>	<b>-,011</b>	<b>-,153**</b>
It's all right with me if some people happen to dislike me.	<b>-,151**</b>	<b>,031</b>	<b>,006</b>	<b>-,117**</b>	<b>,002</b>
I don't always know the reasons why I do the things I do.	<b>,335**</b>	<b>-,052</b>	<b>,071</b>	<b>-,104**</b>	<b>-,330**</b>

\*\**. Correlation is significant at the 0.01 level (2-tailed)*

\**. Correlation is significant at the 0.05 level (2-tailed)*

### 3.2.10. Sensitivity to Punishment

Sensitivity to Punishment reached the threshold with two domains: It correlated positively with Neuroticism,  $r = .51, p < .001$ , and negatively with Extraversion,  $r = -.37, p < .001$ . A total number of 583 participants completed both the Sensitivity to Punishment and the NEO PI-R questionnaire.

17 out of 18 items correlated above the threshold with a minimum of one Big Five domain; 11 items showed an adequate connection with the same two domains as the scale, however five of these also correlated with at least one extra domain. Six items correlated above the threshold with just one of the involved domains, with half of those also correlating above the threshold with another domain.

**Table 13**

*Pearson correlation matrix between the Sensitivity to Punishment items and Big Five domains.*

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
Prefer not to ask for something when I'm not sure I will obtain it	<b>,202**</b>	<b>-,201**</b>	<b>-,056</b>	<b>-,016</b>	<b>-,148**</b>
Am often worried by things that I said or did	<b>,436**</b>	<b>-,082*</b>	<b>-,025</b>	<b>,017</b>	<b>-,211**</b>
Find it difficult to telephone someone I do not know	<b>,242**</b>	<b>-,307**</b>	<b>-,103*</b>	<b>-,050</b>	<b>-,136**</b>
Think a lot before complaining in a restaurant if my meal is not well-prep	<b>-,009</b>	<b>-,083*</b>	<b>-,054</b>	<b>,120**</b>	<b>,026</b>
Often refrain from doing something because of my fear of being embarrassed	<b>,370**</b>	<b>-,284**</b>	<b>-,172**</b>	<b>,065</b>	<b>-,261**</b>
In tasks that I am not prepared for, attach great importance to the possib	<b>,272**</b>	<b>-,188**</b>	<b>-,127**</b>	<b>,066</b>	<b>-,095*</b>
Avoid demonstrating my skills for fear of being embarrassed	<b>,323**</b>	<b>-,228**</b>	<b>-,234**</b>	<b>,096*</b>	<b>-,218**</b>
Am often afraid of new or unexpected situations	<b>,342**</b>	<b>-,220**</b>	<b>-,182**</b>	<b>,049</b>	<b>-,116**</b>

When with a group, have difficulties selecting a good topic to talk about	<b>,246**</b>	<b>-,399**</b>	-,195**	,051	-,149**
Often refrain from doing something I like in order not to be rejected or d	<b>,374**</b>	-,161**	-,135**	,064	-,181**
Am easily discouraged in difficult situations	<b>,383**</b>	<b>-,229**</b>	-,161**	,062	<b>-,285**</b>
Try to avoid speaking in public	<b>,220**</b>	<b>-,335**</b>	-,164**	,161**	-,170**
Am a shy person	<b>,234**</b>	<b>-,440**</b>	-,072	,071	-,154**
Avoid going to unknown places	,167**	<b>-,275**</b>	<b>-,299**</b>	,119**	-,019
Am afraid of more things than others are	<b>,287**</b>	-,176**	-,130**	,018	-,148**
Would find it difficult to ask for a salary increase	<b>,230**</b>	<b>-,202**</b>	-,085	<b>,218**</b>	<b>-,200**</b>
Often think that I could do more things if it was not for my insecurity or	<b>,453**</b>	<b>-,204**</b>	-,057	,025	<b>-,289**</b>
Often worry about things to the extent that my intellectual performance is	<b>,402**</b>	-,082*	-,107**	-,016	<b>-,244**</b>

\*\**. Correlation is significant at the 0.01 level (2-tailed)*

\**. Correlation is significant at the 0.05 level (2-tailed)*

### 3.2.11. Toronto Alexithymia Scale

The Toronto Alexithymia Scale (TAS) was found to be interstitial between three domains. It correlated positively with Neuroticism,  $r = .35, p < .001$ , and negatively with Extraversion,  $r = -.39, p < .001$ , and Openness,  $r = -.53, p < .001$ . Breaking the scale into its three facets, two reached the .30 threshold with one domain only and the third with two domains: Difficulty Identifying Feelings correlated above the threshold with two domains: Positively with Neuroticism,  $r = .51, p < .001$  and negatively with Conscientiousness,  $r = -.33, p < .001$ . The internal validity for the Difficulty identifying feelings was acceptable (7 items  $\alpha = .78$ ) Mean inter-item correlation for the Difficulty Identifying Feelings facet = .34. A total number of 664 participants completed both the TAS and the NEO PI-R questionnaire.

**Table 14**

*Pearson correlation matrix between the Toronto Alexithymia Scale facets and Big Five domains.*

	Neuroticism	Extraversion	Openness to Experience	Agreeableness	Conscientiousness
Alexithymia Scale total	<b>,346**</b>	<b>-,392**</b>	<b>-,329**</b>	-,098*	-,276**
Alexithymia Difficulty Identifying Feelings	<b>,513**</b>	-,222**	-,014	-,106**	<b>-,334**</b>
Alexithymia Difficulty Describing Feelings	,217**	<b>-,425**</b>	-,271**	-,078*	-,151**
Alexithymia Externally-Oriented Thinking	-,034	-,231**	<b>-,472**</b>	-,019	-,069

\*\**. Correlation is significant at the 0.01 level (2-tailed)*

\**. Correlation is significant at the 0.05 level (2-tailed)*

All seven items correlated above the threshold with at least one of the Big Five domains. Three items correlated above the threshold with the same domains as the Difficulty Identifying Feelings facet, while the remaining four were only connected to one of the relevant domains.

**Table 15**

*Pearson correlation matrix between the Difficulty Identifying Feelings facet and Big Five domains.*

	Openness to				
	Neuroticism	Extraversion	Experience	Agreeableness	Conscientiousness
Can be confused about the emotions I'm feeling	<b>.412**</b>	-.107**	.059	-.081*	<b>-.276**</b>
Don't know why I'm angry	<b>.363**</b>	-.171**	-.062	-.145**	<b>-.223**</b>
Have physical sensations that even doctors don't understand	<b>.261**</b>	-.097*	.022	-.093*	-.148**
Don't know what's going on inside me	<b>.298**</b>	-.162**	-.101**	-.053	-.196**
Am often puzzled by sensations in my body	<b>.366**</b>	-.121**	-.028	-.018	-.190**
Have feelings that I can't quite identify	<b>.315**</b>	-.165**	.077*	-.081*	<b>-.259**</b>
When I am upset, I don't know if I am sad, frightened, or angry	<b>.324**</b>	-.187**	-.077*	-.015	-.195**

\*\**. Correlation is significant at the 0.01 level (2-tailed)*

\**. Correlation is significant at the 0.05 level (2-tailed)*

## 4. Discussion

### 4.1. Main findings

To explore our research question, we examined 41 established scales used in a large community sample. We found that many of them were related to the Big Five, and as expected, several were blends of two or more Big Five domains rather than being located within a single domain. Thus, it seems that the prevalence of interstitial trait scales found in the Bainbridge et al. (in press) study was not an isolated instance. We explored where this supposed interstitiality in the scales came from by examining the individual items that constituted them. Investigating the items of these scales, we found no single story that applied to all stand-alone scales. Rather, we found some well-constructed scales built on items that were themselves relatively purely interstitial, while other scales were constituted of quite heterogeneous items that tapped onto several different domains without very much consistency between them, suggesting that some putatively interstitial scales might just reflect heterogeneous items rather than a coherent construct.

During interpretation of these results, we saw that the precise characterization of a single scale or item would sometimes depend on arbitrary numerical thresholds that we adopted for the sake of convenience. An example of this is BAS Drive, which is an interstitial scale where either half of its items are interstitial given our threshold of .20, or all items are if we lowered our threshold by only 10% to .18. Interpretation of our results will therefore build on these thresholds but will necessarily include considerations beyond them as well.

When we did the analyses, we saw that the full Big Five content was represented, but some more than others. Of the 40 scales, 21 had a correlation above .30 with Neuroticism. In

contrast, the other four domains were much less correlated with the stand-alone scales. Comparatively, the domain that was least often correlated notably with scales was Openness with five correlations above .30. The same was true for the 11 interstitial scales. Only one of these was correlated with Openness, which disappeared when dividing the scale into its facets, compared to nine that were correlated with Neuroticism, more often than not in combination with Conscientiousness. This could be due to many interstitial scales simply being located in the interstice between Neuroticism and Conscientiousness, or it could be due to the inter-correlation between Neuroticism and Conscientiousness presenting a skewed picture. Other possibilities include prevalence of Neuroticism-related measures and response bias associated with wording of the items.

Starting with a brief walkthrough of the findings for each scale, we will now walk through how these major themes were reflected in individual results. Afterward, we will consider further points of interest regarding specific scales, evaluate strengths and limitations of the current study, and conclude with a section on implications and possibilities for further research.

#### **4.2. Interstitial scales and their Big Five correlations**

The results indicated 11 scales as interstitial. These will be reviewed briefly here, together with notes on previous findings regarding relationships with the Big Five domains. In general, the findings in this study are largely consistent with earlier research on the same scales.

The BAS Drive scale showed correlations with high Extraversion and low Agreeableness to almost the same degree, consistent with earlier research on the same scale, although usually Agreeableness has been found to be the stronger correlation (Segarra et al., 2014; Smits & Boeck, 2006).

The Cognitive Failures Questionnaire had correlations with high Neuroticism and low Conscientiousness, with the correlation with Neuroticism being stronger. Könen and Karbach (2020) found the same relationships, also with Neuroticism being the stronger correlate.

Dispositional Optimism was found to be interstitial as well, located between low Neuroticism and high Extraversion. Previous research has also focused on links between optimism as a construct and these two domains (Credé et al., 2016), and research on the Dispositional Optimism scale has shown comparable results in terms of correlations with Neuroticism and Extraversion, though sometimes also weaker links to Conscientiousness and Agreeableness (Carver & White, 1994; Sharpe et al., 2011).

Both BIDR scales were indicated as being interstitial, Self-Deception between low Neuroticism and high Conscientiousness, and Impression Management between high Agreeableness, low Neuroticism, and high Conscientiousness. Reid-Seiser and Fritzsche (2001) found similar associations between the scales and the respective Big Five domains, though they found that Self-Deception had a stronger correlation with Extraversion when assessing a group of job applicants compared to a control group.

Two Dark Triad traits were included, psychopathy and narcissism, measured by the Levenson Self-Report Psychopathy scale and the Neuroticism Personality Inventory 16-item version. The Levenson Self-Report Psychopathy scale was correlated with low Agreeableness, high Neuroticism and low Conscientiousness. This scale has previously been examined by Miller et al. (2008) who found largely the same, both in terms of the full scale and its two subscales, although they found slightly stronger correlations. The short-form Neuroticism Personality Inventory we employed was associated with high Extraversion and low Agreeableness. This matches the general findings in terms of the relationships between various versions of the NPI and the Big Five (Ackerman et al., 2011).

The Rosenberg Self-Esteem Scale was, unsurprisingly, included as interstitial; numerous previous studies have found it to be correlated somewhat with all Big Five domains, but most studies have found strong negative correlations with Neuroticism, moderate positive correlations with Extraversion and Conscientiousness, and low positive correlations with Agreeableness and Openness (Pullmann & Allik, 1999; Robins, Hendin, et al., 2001; Robins, Tracy, et al., 2001). This seems largely consistent with the findings of this study, in which the scale was also most strongly correlated with low Neuroticism, and secondarily to Extraversion and Conscientiousness.

The Sensitivity to Punishment scale was associated with high Neuroticism and low Extraversion, consistent with earlier findings by developers of the scale, Torrubia et al. (2001), who found correlations with Eysenckian measures of Neuroticism and Extraversion (quite similar to those domains in the Big Five; Knutson & Bhanji, 2006; Zuckerman et al., 1993). Other researchers have also employed Big Five measures and found similar results (Aluja & Blanch, 2011).

Two scales did not have as much support for their Big Five domain correlations in previous research: The Toronto Alexithymia Scale and the Borderline Personality Inventory. The links between the difficulty identifying feelings facet of the Alexithymia Scale and the Big Five has been investigated by Picardi et al. (2005) - though they used the Big Five Questionnaire with slightly differently labeled domains. While they, too, found that the facet



was associated with low Emotional Stability, i.e., high Neuroticism (in fact, stronger than the association found in this study), they did not find the same as we did for Conscientiousness at all in their analyses. Both Openness and Energy had stronger links to the facet than Conscientiousness did, though none of them were strong. Other research, however, has found somewhat stronger correlations with Conscientiousness, Agreeableness, and Extraversion (Heshmati & Azmoodeh, 2017; Ueno et al., 2014).

As for the Borderline Personality Inventory, we found that this scale was correlated positively with Neuroticism, and negatively with Conscientiousness. Although studies on specific personality profiles and their connections to DSM-IV borderline criteria are abundant (Miller, 2012), we have found no studies examining the relationships between the Borderline Personality Inventory and the Big Five. With other measures, all five domains have been found to predict borderline personality scores, with Neuroticism being by far the strongest predictor (Distel et al., 2009; Trull et al., 2003).

#### **4.3. A comment on our thresholds**

Some of our criteria hinge on semi-arbitrary things. For evaluation on the scale level, we adopted a threshold of .30 as indicative of a noteworthy relationship between a scale and a Big Five domain, i.e., any scale with a correlation .30 or above with two or more domains would be interpreted as interstitial. This threshold was adopted based on the fact that it is a conventional value to use (Costello & Osborne, 2005). Any threshold will be somewhat arbitrary, though, and thus we cannot rule out that some of the scales we excluded might be interstitial. Some excluded scales might in fact be just as interstitial as the included ones and could have been fruitful to include and analyze on the item level. It is perfectly possible that a scale could exhibit strong signs of interstitiality by, e.g., having a correlation of .29 with two or more Big Five domains, yet still be excluded. An example of this is Self-Monitoring, which had a negative correlation ( $r = -.30$ ) with Extraversion, and a positive correlation with Neuroticism, ( $r = .29$ ). Using a different sample, scales such as this one might have been included with a .30 threshold. The same could have been true for this study if our threshold had been slightly different, say, .25 instead of .30 - this would have included Self-Monitoring as well as six other scales (Fantasy Proneness, Life Satisfaction, Social Dominance Orientation, CES-Depression, BAS Fun Seeking, and Sensitivity to Reward; a stricter threshold of .35, in contrast, would have excluded four of our included scales). Lowering the threshold would have included more potentially interstitial trait scales, yet presented new scales close to that threshold; so naturally, interpretation from any threshold will be

somewhat arbitrary as well. Having a more fluent inclusion of scales based on qualitative assessments would leave room for biased opinions and make replications difficult to conduct. While thresholds will always be arbitrary in a sense, we consider setting a threshold to be the better solution for determining inclusion criteria in cases such as this, though we are aware of their limitations.

The same issue presents itself for our evaluation of items - while we adopted a threshold of .20 in order to have something numerical to present from our correlation analyses, this is not an absolute threshold either and may seriously skew interpretations of included scales if taken literally. For this reason, we will consider other aspects in more depth when interpreting item sets of interstitial scales, and this threshold of .20 will not be the be-all and end-all for interpretation of interstitiality on the item level. With that said, we will now move on to present our findings from analyzing interstitial trait scales on the item level.

#### **4.4. How are interstitial trait scales constructed on the item level?**

To answer our research question, we examined the included interstitial scales' item-level correlations with the Big Five domains. We speculated that an interstitial trait scale could be the result of a coherent, interstitial construct at its core, but also could be due to a poorly constructed item pool pulling the total score in different directions. The results of our correlation analyses presented varied findings among interstitial scales and their items. This indicates that interstitial scales exist on a spectrum, with convincingly interstitial scales at one end and unclear scales at the other end. For a trait scale to be convincingly interstitial at its core, we would expect the items themselves to show the same kind of interstitiality as the scale, at least to a certain degree, rather than having some items correlate with one domain and some with the other.

Some scales seem to indeed be interstitial, not only at the scale level, but at the item level as well. A prime example of this is the Rosenberg Self-Esteem Scale, which was interstitial between Neuroticism, Extraversion, and Conscientiousness. Out of 10 items, eight are interstitial with at least two of the three domains that the scale is correlated with; and five of those are interstitial with all three. Another item is correlated above .20 with only Neuroticism; this is not very surprising, though, as Neuroticism had by far the strongest correlation with the full scale. Looking closer at the items, a slightly laxer threshold of .18 or so would include even more correlations with Extraversion and Conscientiousness on the item level, leaving eight items with correlations with all three domains; this seems to suggest that almost all items have relationships with all three domains. This also fits with the

relatively high inter-item correlation at .43. According to Briggs and Cheek (1986), inter-item correlation should ideally be in the .20 to .40 range in order to represent enough variance but still be within one dimension - it really seems, then, that the Rosenberg Self-Esteem Scale is a prime example of a unidimensional interstitial construct located between (low) Neuroticism, Conscientiousness, and Extraversion. Its only problem might in fact be that it could be *too* unidimensional. Briggs and Cheek (1986) does however specify that a scale with inter-item correlations over .50 suggests that the scale likely has overly redundant items or that the construct measured is too specific, meaning that an inter-correlation at .43 is likely still all right.

Another example of a relatively convincing scale is the BAS drive scale. The BAS drive had all four items with primary connections with Extraversion above the threshold, of which two also correlated with Agreeableness. Looking closer, though, the two remaining items came close to our threshold of .20. Yet again, by lowering the threshold by only 10% to .18, all four items would have been considered interstitial, thus making the scale very convincing in its interstitiality. The range of the item correlations with the two domains is quite small, suggesting a fairly even relationship with the domains - the correlations between the four items and Agreeableness indeed all fall rather close to our threshold of .20, with the weakest correlation being -.18, and the strongest being -.27. The BAS Drive scale also has a relatively high inter-item correlation at .42; as with the Rosenberg Self-Esteem Scale, we do not see this as a big issue. One thing we encountered from reviewing both of these scales was that some of their items seem to get some influence from Openness, enhanced when considering a lower threshold; with a threshold of .18, the self-esteem scale sees two items out of ten correlating with Openness, while BAS Drive sees one of its four correlating with Openness. This does not concern us, though. An interstitial construct located between several domains can certainly have elements that lie slightly more toward another domain than the rest of the construct. The important part would appear to be that the elements of the construct (or at least a majority of them) have mostly similar relationships to the domains the full construct seems to be interstitial between, even if some of them also tap other domains. Only if the correlations with another domain seem to be systematic would we have to reevaluate where in the framework the construct is really located.

The story repeats itself for both the Toronto Alexithymia Scale and Dispositional optimism. Both seem to possess relatively convincing interstitiality when reviewing their items. The Difficulty Identifying Feelings facet of Alexithymia has three items that reach the threshold of .20 for both Neuroticism and Conscientiousness, while the remaining four only

reach it with Neuroticism. Yet again, lowering the threshold to .18 would include six of the seven items without any disturbing correlations with other domains. Curiously, this scale was one of few to not have much support for this pattern of trait affiliation in the literature, though there were some mixed findings (e.g., Picardi, 2005; Ueno et al., 2014). It does however seem convincingly interstitial in this study given the current findings, though the evidence for its interstitiality is not as robust as for some of the other scales. As for Dispositional Optimism, all six items correlated above our .20 threshold with both Neuroticism and Extraversion, suggesting a sound interstitial construct. Yet four of these items also correlated with Conscientiousness above the .20 level, with the two remaining items coming close as well. Such a systematic correlation with the domain seems confusing from a perspective where Dispositional Optimism presents as interstitial between only two domains. However, the explanation is again to be found in somewhat arbitrary thresholds. For our scale level analysis, we applied a threshold of .30, which indicated only Neuroticism and Extraversion having a noteworthy relationship with Dispositional Optimism. Looking closer, it is actually no surprise that so many items are related to Conscientiousness, as the full scale itself is also had a correlation on .28 with Conscientiousness. Perhaps, then, our initial screening simply failed regarding this particular scale, and Dispositional Optimism is in truth interstitial between all three domains. In any case, these findings suggest that the interstitiality of some interstitial scales is truly rooted in an interstitial construct.

At the other end of the spectrum, some scales had items that presented a very mixed picture in terms of their correlations with the Big Five domains. One example of this is the Cognitive Failures Questionnaire. The full scale appeared interstitial between Neuroticism and Conscientiousness, which eight of its 25 items reflected. Looking closer at the items, however, only a few items seem likely to be truly conceptually linked to both domains. Rather, it seems that the reason for the interstitial items is that the scale contains some items that are conceptually related to Conscientiousness but not Neuroticism, but which are correlated with Neuroticism anyway due to the valence (positive or negative) of the items influencing neurotic people to respond negatively, i.e., people higher in Neuroticism also rating themselves more negatively on other, unrelated measures (Fossum & Barrett, 2000). An example of such an item could be *'How often did you forget appointments?'*, which is plausibly related to Conscientiousness in that highly Conscientious people are probably less likely to forget appointments. The link to Neuroticism is less convincing conceptually; there is no immediate reason neurotic people should forget appointments more often than emotionally stable people. A more conceivable explanation is that more neurotic people are

more likely to endorse undesirable traits or deny possessing desirable traits (Fossum & Barrett, 2000), and thus more likely to report that they often forget appointments. The scale does not seem to have achieved interstitiality straightforwardly, by assessing items that are themselves interstitial, but instead by compiling psychologically distinct items. Further, it is curious that a scale related to cognition would show such little associations with the cognitive trait, Openness. However, research has found that the Cognitive Failures Questionnaire is not related to cognitive abilities (Könen & Karbach, 2020); rather, it seems that the scale reflects worry about one's own cognition rather than cognition itself (Wilhelm et al., 2010).

Other examples of rather unconvincing interstitiality include the two scales from the Balanced Inventory of Desirable Responding (BIDR): The Impression Management scale, and the Self-deception scale. These two scales both show quite an ambiguous picture in their item correlations. Self-Deception looks particularly scattered. The full scale appears interstitial between Neuroticism and Conscientiousness, yet out of 20 items, only five appear interstitial between the two domains (and sometimes a third domain, Extraversion). Looking closer at item correlations, this seems to be a fair assessment; even considering lowering the threshold to .18 as we have discussed before does not change the picture. The reason for the full scale's apparent interstitiality does not seem to be accounted for by an intrinsic construct in this case, as the items do not consistently reflect the same interstitiality. The range of the correlations is not particularly impressive either as there is quite a large span between the items that are the most and the least correlated with the two domains. This is further supported by Self-deception having a very low mean inter-item correlation at .10; a level at which it is unlikely that a single score can adequately represent the complexity of the items (Briggs & Cheek, 1986). Self-deception also had the only questionable Cronbach's alpha of the 11 scales included for item-level analysis. The other half of the BIDR, Impression Management, does not come out as convincingly interstitial on the item-level either. The scale's relationships with Neuroticism, Agreeableness, and Conscientiousness are not clarified by the item correlations: Only two items represent this interstitiality through correlations with all three domains, and another four items correlating with two of the domains. In the case of Impression Management, the scale actually presents a good Cronbach's alpha - yet it has a less-than-ideal mean inter-item correlation at .18. This is likely to be an example of a case where the number of items in a scale sometimes inflates Cronbach's alpha (Tavakol & Dennick, 2011).

Overall, it appears that the relationships between the two BIDR scales and the Big Five domains stem from few items correlating highly with these domains (though often in

conjunction with another domain), while other items do not show nearly the same connections. In both cases, many items do not display a noteworthy relationship (by our metrics) with any of the five domains at all. If the BIDR is supposed to represent two interstitial constructs, then perhaps the item list should be trimmed or revised to reflect the interstitiality better. Another possibility, though, is that the two BIDR scales are not expressions of coherent intrinsic traits. The BIDR was not developed to assess a coherent, interstitial trait - rather, it is meant to measure the tendency to respond in a specific, biased way, i.e., in a socially desirable way. Self-deception reflects an honest, but overly positive self-presentation, while impression management reflects a self-presentation tailored to pleasing an audience (Paulhus, 1991). So is the tendency to respond in a specific, biased way, i.e., in an impression managing or self-deceptive way, an intrinsic, coherent trait? From the picture presented in this study, this would not seem to be the case (though it has been argued before that it is a single, albeit multidimensional, trait; Furnham 1986). Perhaps the underlying nature of desirable responding is rather a collection of traits pushing people to respond in such a way. This would explain why the correlations at the item-level are so varied in terms of relationships with the Big Five rather than presenting a coherent, interstitial picture. The inter-item correlations of both scales are also below the ideal, and thus appear less coherent than could be desired - indicating that each scale might cover more than just one dimension. Curiously, as the BIDR scales, the Cognitive Failures Questionnaire was developed to assess self-reported cognitive failures, i.e., a form of behavior. Though it could be assumed that cognitive failures stem from an underlying personality trait, particularly if the scale in reality measures worry about cognitive failures, it is nevertheless interesting that the most unclear scales were developed to assess behavior in one way or the other. Perhaps this provides a clue as to why they displayed so questionable interstitiality; if they did not set out to assess a coherent trait, the lack of consistency in items might not be so surprising after all.

Another point of interest is what seems to be the common occurrence that less-convincingly interstitial scales in this study have been disputed in one way or the other. For instance, the BIDR questionnaire has been criticized for both reliability (Li & Bagger, 2007) and factor structure (Gignac, 2013). Similarly, several other included scales have been criticized as well, often citing issues with their factor structures (e.g., Levenson Self-Report Psychopathy Scale, Garofalo et al., 2019; Cognitive Failures Questionnaire, Wilhelm et al., 2010; and Narcissism Personality Inventory, Rosenthal et al., 2011). We will consider these in greater depth in the section on further considerations regarding the scales. In contrast, the

four scales that appear more sound in their interstitiality have generally received less criticism in comparison and have been found to be sound scales, e.g., BAS Drive (Leone et al., 2001), Rosenberg Self-Esteem (Huang & Dong, 2011), Toronto Alexithymia Scale, (Schroeders et al., 2021) Dispositional optimism as measured by the revised Life Orientation Test, (Chiesi et al., 2013).

Of course, some scales are trickier to categorize at either end of the spectrum; they fall somewhere in the middle. One example of this is the Sensitivity to punishment scale. 11 of its 18 items were interstitial between Neuroticism and (low) Extraversion, even if sometimes with additional domains present as well. One domain that was frequently correlated with the scale was Conscientiousness. A closer look at the full scale's correlations shows us that it does indeed have a correlation of  $-.29$  with Conscientiousness, thus falling just short of our initial assessment as to whether it had a noteworthy relationship with the domain. This could just be an error in our screening; excluding it as an interstitial domain might have simply been a mistake. Whether sensitivity to punishment is truly an interstitial construct - and which domains it is interstitial between - seems to be unanswered at this time.

Another, slightly more confusing example is the Levenson Psychopathy scale. While four of its 10 items appear interstitial between Neuroticism and Conscientiousness, and one additional item could be interpreted as interstitial as it barely missed our threshold of  $.20$ , there are also several items that do not display this interstitiality at all. Some items are simply related to only one of the two of the two domains; another two items are correlated with Agreeableness above all other domains. This might suggest these items do not belong in this subscale and would be more appropriately located in the Primary subscale, which had a primary correlation with Agreeableness. Another possibility might be that the Secondary subscale is actually interstitial between three domains (Neuroticism, Conscientiousness, and Agreeableness), and our initial analysis only missed the Agreeableness dimension due to it being just below our threshold of  $.30$ . The possibility that this construct could be interstitial between all three domains - at least as it assessed here - does not seem particularly likely, though, as no items shared a correlation with both Conscientiousness and Agreeableness. It seems, beyond whether the scale is interstitial, an evaluation of its items might be in place - but perhaps there is a truly interstitial construct to be found underneath.

To summarize, there is no particular pattern as to how putatively interstitial scales behave on the item level. It does seem that at least some scales are made up of coherent, interstitial collections of items, meaning that those scales are likely assessing constructs that are interstitial in nature. On the other hand, there are also scales that certainly seem to be

confusing compilations of items, where the full scale appears interstitial due to few items pulling in each direction rather than all items sharing correlations with two or more domains. In yet other cases, it is difficult to distinguish; some scales are neither convincing enough to deem them truly interstitial nor confusing enough at the item level to deny it. Next, we will consider how interstitial scales were distributed across the Big Five framework.

#### **4.5. The bigger picture of interstitiality in the Big Five framework**

Some domains were seen to be correlated with the stand-alone scales much more often than others; Neuroticism was by far the domain most frequently related to the scales while Openness was the least. This was also the case for interstitial scales, where nine of the 11 scales and subscales that turned out interstitial had a correlation above .30 with Neuroticism. Of these, six were primarily correlated with Neuroticism. In contrast, none of the final interstitial scales and subscales were correlated at or above .30 with Openness. The latter is the fifth factor of the Big Five framework and is both the least universal domain (John, 2021), and, more importantly, the domain that typically accounts for the least variance in participants' responses to other personality measures (Peabody & Goldberg, 1989), so it is not surprising that this would be the least important factor in relation to stand-alone scales. In contrast, the importance of Neuroticism could be considered rather surprising as it is typically considered only the fourth most important domain in terms of explaining variance (Peabody & Goldberg, 1989). However, these findings are entirely consistent with the findings of Bainbridge et al. (in press), whose study also showed an over-representation of Neuroticism compared to other domains and an underrepresentation of Openness. The relative importance of Neuroticism might suggest that it is a more important domain in terms of internal psychological constructs than we give it credit for. Perhaps Neuroticism has been generally underestimated from the beginning because a domain such as Extraversion (the first, most widespread factor) is a personality domain that summarizes traits of interpersonal nature (John, 2021) - other people will usually easily spot that a person is very extraverted, and thus many more words might exist for this trait compared to a trait such as Neuroticism, which might be more hidden within a person. While fewer words in the dictionary pertain to it, it might still be as important as or even more important than Extraversion when it comes to influence on other personality traits, as suggested by the results here. Paunonen and Jackson (2000) raised this issue as well, suggesting that the number of words for a trait are not necessarily indicative of its importance, though with regards to dimensions of behavior that were far less widespread than Neuroticism. The argument can conceivably be extended to this



instance, though - domains with more words clustering around them are not necessarily more important than those with fewer. This does not explain why lexical research has found it to be only the fourth most important domain in terms of accounting for variance, though. So perhaps there are other explanations for the importance of Neuroticism in stand-alone trait scales.

A related explanation for the apparent overrepresentation of Neuroticism is that many scales that researchers have found important to develop and use are simply tied to Neuroticism. Neuroticism has great predictive power in relation to negative life outcomes - for example, it is greatly associated with many mental disorders - while the inverse, emotional stability, is associated with a range of positive outcomes, such as subjective well-being (Jeronimus et al., 2016; Vittersø, 2001). In a similar context, Credé et al. (2016) argued that many narrow traits in their study may overlap with Neuroticism and Extraversion because many measures were designed to predict success in real-life settings, and these domains are adaptive in many settings requiring social interaction. A related thing may be going on here: The implications of Neuroticism for life outcomes may spur more researchers to develop scales that measure constructs related to this domain than constructs related to the other four. This is especially true for fields that work with topics primarily related to Neuroticism, such as psychiatry and clinical psychology, which are very prominent fields within psychology. As observed by Bainbridge et al. (in press), 145,411 papers were published in these two fields from 2015 to 2017, compared to 207,716 papers for all other fields of psychology combined. The prominence of these fields could lead scales that are correlated with Neuroticism to be more widespread, and thus more likely to be included in a sample such as the ESCS. In the same vein, the other domains would likely be underrepresented. Another possibility that Bainbridge et al. (in press) also presented is that neurotic people may rate themselves more negatively on scales that are otherwise unrelated to Neuroticism, thus exaggerating the importance of the domain (Fossum & Barrett, 2000). This bias, which Bainbridge et al. (in press) labeled *neurotic-content contamination*, could be due to how items are worded in the stand-alone scales. In some cases, for instance regarding some items from the Cognitive Failures Questionnaire as outlined in a previous section, this might certainly have happened; looking qualitatively at the items, though, we do not believe this is a widespread issue in this sample.

More often than not, correlation with Neuroticism was coupled with a correlation with Conscientiousness as well; this was the case for seven of the scales and subscales. So, as it turns out, well over half - in fact, almost two thirds - of the interstitial scales in this study

are correlated with both Neuroticism and Conscientiousness (and sometimes a third domain). If this sample is assumed to be representative of interstitial trait scales as a whole, this would suggest that more than half of interstitial scales lie in the interstice between Neuroticism and Conscientiousness. This could have implications for the Big Five as a model - if so many important constructs are located between Neuroticism and Conscientiousness, that might imply that there is quite an important dimension of personality located there that the Big Five does not account very well for. Realistically, though, it is extremely unlikely that this sample is representative of all stand-alone scales. Though 41 scales are a fairly sizable sample from which to extract interstitial scales, there are a myriad of other scales available. In our study, 11 out of 41 scales indicated interstitiality - just over a quarter. If this is generally true for psychological measures, that means there are many interstitial scales out there, and these 11 are just a small selection. From that standpoint, it is unlikely that such a small selection of interstitial scales would be representative, and the abundance of traits located between Neuroticism and Conscientiousness could simply be a fluke. A more likely explanation for what we see, though, is the correlation between Neuroticism and Conscientiousness themselves, which is rather high, although not unheard of before (Strus et al., 2014). It would be interesting to see results in a study in which intercorrelations are controlled for.

Curiously, of the four scales that we interpret as relatively convincingly interstitial, three are related to Extraversion. Three are also related to Neuroticism, though this is less surprising given the overall prevalence of associations to Neuroticism. Extraversion looked comparatively less important than Neuroticism, although not unimportant - of the full scales and facets that were interstitial, five were correlated with Extraversion, compared to nine for Neuroticism. It is somewhat surprising that despite having just over half of the amount of relationships Neuroticism had with interstitial scales, there were equally many convincing scales that were associated with both. Extraversion-related scales thus seem to have a higher success rate in terms of being convincingly interstitial. Perhaps this could mean that many truly interstitial constructs are located between Extraversion and other domains - or it could mean that Extraversion-related interstitial constructs are easier to assess than interstitial constructs related to other domains.

An interesting endeavor concerned with the distribution of interstitial scales among domains would be to examine interstitial scales in the HEXACO framework as well as the Big Five. As some of the domains in HEXACO are rotated relative to their Big Five counterparts, the Big Five domain Neuroticism is split between Emotionality and Agreeableness, as well as a bit of Extraversion in HEXACO (Ashton et al., 2014; Ludeke et

al., 2019). Perhaps using this framework, we could gain a different understanding of the abundance of traits we found that were related to Neuroticism. The HEXACO domains also generally intercorrelate less than the Big Five domains (Moshagen et al., 2019; Van der Linden et al., 2010), so the results would be less affected by intercorrelations in the broader sense. We might even see a different picture of interstitial traits overall, as distribution of interstitial scales between domains could change. As Big Five Extraversion, Conscientiousness, and Openness generally correspond fairly well to their HEXACO counterparts (Ashton et al., 2014), any interstitial scales located between these three domains would likely appear in a similar fashion using HEXACO as a framework. On the other hand, interstitial scales correlating with Big Five Neuroticism or Agreeableness (which, in this study, is all of them) could look very different in a HEXACO perspective. We might find that some scales that seem interstitial in the Big Five lie relatively purely within one HEXACO domain. The reverse could also be the case and some scales that seem to be relatively purely located within a Big Five domain could appear interstitial in an analysis using HEXACO.

#### **4.6. Further considerations**

We noticed several issues of theoretical interest beyond the item-level behavior of each scale. First of all, we noticed that in some cases, constructs were assessed with more than one scale, and one scale appeared interstitial while the other did not. One example of this is the Sensitivity to Punishment scale, included in this study due to its apparent interstitiality between Neuroticism and (low) Extraversion. Sensitivity to Punishment was developed to assess BIS functioning (Torrubia et al., 2001), so it is curious that the BIS scale in the ESCS developed by Carver and White (1994) was overwhelmingly correlated with Neuroticism and did not appear interstitial at all - after all, they claim to measure (at least roughly) the same construct. A relatively pure representation of Neuroticism corresponds well with the work of several researchers, who view BIS and BAS as similar to Eysenck's Neuroticism and Extraversion, respectively (Davidson, 2003; Knutson & Bhanji, 2006; Kosslyn et al., 2002). Gray himself did initially place BIS and BAS at rotated degrees (i.e. high BAS lies between high Neuroticism and high Extraversion, and high BIS lies between high Neuroticism and low Extraversion), but later reconsidered this and located BIS closer to Neuroticism and BAS closer to Extraversion (Pickering et al., 1999). Eysenck's Neuroticism and Extraversion are quite similar to their Big Five namesakes, although not exactly the same (Knutson & Bhanji, 2006; Zuckerman et al., 1993), so we would expect that the BIS and BAS would probably also lie relatively purely within these domains in the Big Five - as the BIS scale indeed does.

However, while the Carver and White (1994) scales were generally well received, some also criticized precisely their location in the Eysenckian space, i.e., in the case of the BIS scale, it was missing the link to Extraversion (Heubeck et al., 1998; Torrubia et al., 2001; Zuckerman et al., 1999), which suggests that there has been some inconsistency about how BIS functioning should be conceptualized. So, it appears that if the Sensitivity to Punishment scale does assess a truly interstitial construct, we cannot be completely sure it is assessing BIS functioning as it claims.

Another construct assessed by two different scales in the ESCS was psychopathy, which was assessed using the Levenson Self-Report Psychopathy Scale (LSRP), included in this study, and the Self-Report Psychopathy Scale (SRP; Williams et al., 2003), which did not give indications of being interstitial in our initial analysis. The SRP seems to be heavily related to Agreeableness just as the LSRP but does not show the same relationships with the other domains. Again, it is somewhat surprising that two scales claiming to measure the same construct are not more alike, particularly since they were both modelled after the two-factor structure in the Psychopathy Checklist (Miller et al., 2008). A reason for this could be that the two scales do not measure exactly the same sub elements of psychopathy after all; when the LSRP was taken apart, only the Secondary subscale appeared interstitial. A measure not assessing secondary psychopathic traits to the same degree might then appear less interstitial. Though the SRP has items that qualitatively look similar to those included in the Secondary Subscale of the LSRP (e.g., *'I am an impulsive person'*, *'I always plan out my weekly activities'*), there are not as many of them, which could explain why the SRP only reached the .30 threshold for Agreeableness and no other Big Five domain. Comparing the SRP to the LSRP, the correlations of the SRP actually look more alike to the Primary Subscale of the LSRP, which also correlated at the .30 threshold with only Agreeableness. Perhaps, then, the LSRP simply considers an additional element of psychopathy (or perhaps this interstitial element of the LSRP is not truly a part of the construct we know as psychopathy).

Besides these seemingly confusing cases, we also noted during this process that some scales have characteristics that could affect how their interstitiality presents. As mentioned earlier, there is a tendency for the scales that are less convincing in their interstitiality to also often be scales that have been disputed in one way or another. For some of these scales, there have been developments in the understanding of their constructs. The factor structure of several of our included interstitial scales have been under debate. An example of this is the Narcissism Personality Inventory (NPI). Ackerman et al. (2011) proposed a three-factor structure consisting of the dimensions Leadership/Authority, Grandiose Exhibitionism and

Entitlement/Exploitativeness, which utilizes 25 of the 40 items. Others have found stronger support for a two-factor model (Corry et al., 2008), but many researchers have also found other solutions from the two factors up to seven (Emmons, 1987; Raskin & Terry, 1988). Thus, the factor structure of NPI seems confusing at best. Furthermore, these solutions were generally built on the forced-choice NPI (Boldero et al., 2015). As the ESCS used a Likert scale response format, we cannot be sure that a factor structure extracted from a forced-choice format translates to this study (Boldero et al., 2015). We therefore took the consequence of this empirical ambiguity, and decided to use the single factor, shortened 16-item version of the NPI (Ames et al., 2006) instead (though an analysis with a validated factor-structure would certainly be interesting as well).

Another scale with a disputed factor structure is the Levenson Self-Report Psychopathy scale. Our initial analysis of this scale suggested that the scale was interstitial between three domains (Neuroticism, Agreeableness, and Conscientiousness), but after dividing the scale into its subscales, only the Secondary subscale appeared interstitial (between Neuroticism and Conscientiousness) at our threshold. From this, it would seem the scale is divided into facets in a reasonable way - the interstitiality of the full scale is thus explained by the Primary subscale measuring a part of the construct that pertains to low Agreeableness, while the Secondary subscale is interstitial in itself and measures a part of the construct that pertains to Neuroticism and Conscientiousness. However, maybe this division is not so reasonable after all. The LSRP was originally developed to distinguish between primary and secondary psychopathic traits, first proposed by Karpman (1948) - a distinction between deliberate, callous antisocial behavior and neurotic, impulsive traits leading to antisocial behavior (Levenson et al., 1995). Levenson et al. (1995) believed these two types were compatible with the two factors of the measure the LSRP was modelled after - the original Psychopathy Checklist (PCL), which is considered the standard measuring device of psychopathy (Williams et al., 2003; Williams & Paulhus, 2004). Between these two factors, one factor assesses interpersonal and affective features, while the other assesses antisocial and lifestyle features, though there has been some discussion as to whether the LSRP factors really capture the same as the PCL factors (Miller et al., 2008). However, since then it has become apparent that the two-factor structure of the PCL does not hold up; Rather, a four-factor oblique structure of psychopathy has been identified in the PCL-R and related measures (Williams et al., 2007). The factor structure of the LSRP has been studied as well, as increasing evidence suggested that the two-factor structure was inadequate (Brinkley et al., 2008; Garofalo et al., 2019). There have been different results from this; though some have

found a four-factor structure closely resembling Hare's four factors of the PCL-R, a three-factor structure is more widely accepted at this time (Brinkley et al., 2008; Pzederska et al., 2020). Due to the lack of consensus, we decided to continue analysis with the original two factors. Regardless of which factor structure might be the consensus in the future, it would be interesting to examine how the LSRP would behave if divided into three or four subscales instead of the two it currently is. Perhaps it would give a clearer picture of the interstitiality in the scale. It might also improve the inter-item correlation of the subscales, as the mean inter-item correlation of the Second Subscale is currently on the low side at .166 (Briggs & Cheek, 1986); presumably, if different subscales are more accurately representative of underlying factors, those would yield more ideal inter-item correlations, which would likely also come with a higher Cronbach's alpha.

The Cognitive Failures Questionnaire also appears on shaky ground. The authors of the original paper did not find it possible to extract stable factors or dimensions other than the general factor (Broadbent et al., 1982). However, the idea that these cognitive slips are connected to just one underlying trait has been challenged both theoretically and empirically, and several researchers have constructed different factors for the CFQ (Wilhelm et al., 2010). Wilhelm et al. (2010) proposed a three-factor structure with the factors Clumsiness, Retrieval, and Intention Forgotten, though this only contains 12 of the original 25 items. Other solutions range from two to seven factors (Larson et al., 1997; Matthews et al., 1990), of which the most widely used is perhaps a four-factor structure with the dimensions Memory, Distractibility, Blunders, and (Memory for) Names (Wallace, 2004; Wallace et al., 2002). The lack of consensus led us to use the single-factor structure proposed originally, for which there has also been found some support (Bruce et al., 2007). In future research it could be relevant to see whether the 12-item, three-factor structure by Wilhelm et al. (2010) provides a clearer picture of the interstitiality of the CFQ.

These developments regarding the included interstitial scales mean that the results found in this study regarding these specific scales will not necessarily be applicable in the future. The same could potentially be true for the other scales as well. With new revisions or developments in the understanding of the scales, they might be structured completely differently - and so the interstitiality might change drastically or even disappear entirely. Perhaps some of the scales that look interstitial are simply contaminated with items tapping unrelated constructs, and a new revision will see them fit into a single domain. On the other hand, new revisions with better core items might yield more consistent interstitial results on

the item level forming a more convincing scale as a whole that measures a truly interstitial construct.

#### **4.7. Implications and future research**

Implications for this study may not be obvious. Above all, it contributes to the field of knowledge regarding the scope of the Big Five.

The Big Five is supposed to capture most of the interesting trait variance and package it well - that is, it is assumed that the five domains is the most sensible way to categorize personality in the broad sense (John, 2021). Bainbridge et al. (in press) found support for the Big Five's ability to function as an organizing framework, as 71,2% of scales in their study could be located within the Big Five. Looking at the full collection of scales extracted from the ESCS in this study, the current findings seem to be in line with that - 34 of the initial 41 scales have at least a moderate correlation, i.e., a correlation of .30 or above, with at least one Big Five domain. It seems, then, that the Big Five performs fairly well at connecting to the scales, as over 80% of scales attained at least one moderate correlation - despite not being selected based on an expectation of being related to the Big Five, and presumably not being developed with their space in the Big Five framework in mind.

The Big Five is not just a comprehensive framework for describing variance of traits, though. Proponents also make the claim that exactly these five domains are the optimal way to organize personality (De Raad et al., 1998; McCrae & John, 1992). From the results of this study, it would appear that 11 out of the 34 scales that are at least moderately correlated with the Big Five share this association with at least two domains. This does not challenge that the Big Five captures variance - quite the contrary, these 11 scales seem to be very related to the Big Five as they correlate at least moderately with a minimum of two domains. It can however challenge whether the Big Five manages to divide personality into optimal categories.

The Big Five is frequently depicted as a simple, tree-like hierarchical structure, with domains branching into aspects, then facets, then nuances (McCrae, 2020). The binary interpretation of the presence of interstitial traits within this model would be that either all interstitial trait scales are bad measures that do not measure unidimensional traits, or that the Big Five simply does not cut it as an organizing framework due to it not having attained the simple structure. The former explanation would of course require all interstitial scales to have bad item sets pulling the full scale in the direction of different domains, while convincingly interstitial scales with coherent item sets would support the latter explanation. As seen,

interstitial scales behaved differently from each other. Given the premises of a strictly simple, hierarchical model, that might be taken as support that the Big Five fails to deliver. However, the simple hierarchical structure is not truly how most researchers see the Big Five. Plenty of researchers have expressly acknowledged the existence of interstitial traits (Credé et al., 2016; Mooradian et al., 2016), and some presentations of the Big Five model, such as the Abridged Big Five Dimensional Circumplex (Goldberg, 1992; Hofstee et al., 1992) mentioned in the introduction, explicitly emphasizes the complexity of the Big Five. This two-dimensional model combined of two given Big Five domains provide a better understanding of interstitial traits located between those two domains. This might be a satisfactory way to depict the Big Five, although it would have some limitations of its own regarding its discriminatory power between similarly interstitial traits. Another issue with this model concerns interstitial constructs that are located between three domains. In this study, both the Rosenberg Self-Esteem scale and the Impression Management scale are associated with three domains. It would be intuitive to think that constructs that correlate with three domains might have some issues, which is indeed the case with Impression management. However, the Rosenberg Self-esteem scale seems to have sound interstitiality rooted within its items, so at least in some cases, a two-dimensional model falls short. In order to accommodate three-way interstitial scales, the circumplex would have to be three-dimensional - and while theoretically possible, it would then not be able to deliver the same kind of simple graphic presentation that a two-dimensional circumplex provides. Conceivably, interstitial traits located between four or even all five domains might also exist, which this model would certainly have trouble depicting, at least while retaining the ability to provide an overview.

Another potential explanation for interstitial traits is the concept of metatraits. Several researchers have noted that the Big Five are connected to each other in systematic ways: Agreeableness, Conscientiousness, and (reversed) Neuroticism form the metatrait Stability (or Alpha), while Extraversion and Openness form the metatrait Plasticity (or Beta) (DeYoung, 2006; Digman, 1997; Strus et al., 2014). If all interstitial traits are related to domains within the same metatrait, that could explain the connection to more than one domain - the interstitial traits could really just be located at a higher level of the hierarchy. For instance, the observation that many traits are interstitial between Neuroticism and Conscientiousness would be entirely reasonable according to this explanation; indeed, the Impression Management scale could be a prime example of a trait affiliated with high Stability if it were not otherwise so unclear. But we do not only find interstitial scales within



metatrait domains. Several convincing interstitial traits are located between domains related to the Stability metatrait and those related to the Plasticity metatrait. For example, the Rosenberg Self-esteem scale is associated with both Neuroticism and Conscientiousness, which are Stability domains, but also with Extraversion, a Plasticity domain. So, it appears that neither the metatraits nor the circumplex structure seem to provide a structural solution that can incorporate interstitial traits satisfactorily in a hierarchical model. This emphasizes the proposition made by Bainbridge et al. (in press) that the Big Five is more accurately described as having a heterarchical or quasi-hierarchical structure with some facets and traits located at interstitial locations between two or more domains. Research on interstitial traits may serve as a reminder that the classic representation of the Big Five is not telling the whole story.

So, if the Big Five is best thought of not as a simple hierarchy, but as a heterarchy, are interstitial traits no threat to the model at all? Well, it depends. While interstitial traits in and of themselves are not necessarily a challenge, the prevalence of them could be. Despite having a complex structure, the Big Five still claims to possess the superior organizing system for personality traits. Preferably, such a model should be able to relatively cleanly categorize most traits, and so a sizable majority of personality constructs should be readily located within the five categories rather than having a lot of them located between domains. Yet of the 34 included scales that were at least moderately related to the Big Five, 11 scales - almost a third - was interpreted as interstitial by our criterion; similarly, Bainbridge et al. (in press) found that many of their included scales appeared interstitial. These findings would appear to suggest that many trait-like constructs are in fact located between domains, potentially threatening the idea that these five categories represent the optimal organization of personality. There are conceivable counter arguments as to why these findings might not be a problem, though.

Particularly the question of the soundness of such interstitial scales merits discussion. As mentioned, only four of the 11 interstitial scales expressed very convincing interstitiality when examined on the item level. Now, four out of 34 is much less and seems much less challenging for the Big Five. Just over a tenth of scales being interstitial would not as readily imply that the Big Five failed to divide personality in a meaningful way. However, this also comes with a caveat - we do not know how the other scales in the sample would fare if subjected to an item-level analysis. Four out of eleven scales being convincing after an item-level analysis does not sound impressive, but before we know whether four out of eleven scales is a high success rate or not when it comes to whether the scale indicates a uniquely

bad or good scale concept, we would first need to also subject the remaining, non-interstitial scales, to scrutiny. Until then, we cannot know whether so many interstitial scales are badly combined on the item level because it is difficult to assess interstitial constructs - or if that is just the baseline amount of scales that are pretty badly constructed. Another possibility is of course that comparatively fewer non-interstitial scales are constructed well compared to the interstitial scales.

One suggestion by Bainbridge et al. (in press) was that it would be prudent that researchers attempt to locate new and existing scales within and peripheral to the Big Five, as this can help to reduce redundancy and provide a better understanding of trait scales' locations in the personality space. During this study we found a number of papers examining relationships between scales and Big Five domains, so this is being done to some degree already; yet it is incredibly rare that studies take note of interstitiality. If more researchers try to organize stand-alone scales more within a Big Five perspective and take note of interstitiality, this could also help lay a foundation where we may get a clearer picture of the prevalence of interstitial traits.

Bainbridge et al. (in press) also stated that if a scale has measurement, validity or reliability concerns then alternative assessments should be created. As Bainbridge et al. (in press) showed that the average scale can be reasonably said to be located within the Big Five, scale developers can also use that in refining their ideas of what they are attempting to do when developing scales, thus likely developing more sound scales. At one point in scale development, developers will sometimes include a Big Five inventory (though usually to demonstrate that their scale has incremental validity beyond the Big Five domains; Bainbridge et al., in press). If developers follow the recommendation of Bainbridge et al. (in press) and locate their scales within the personality space by using Big Five measures, it would also be prudent to examine item correlations with the Big Five rather than only looking at internal consistency. For example, if five items reflect Extraversion and only one reflect Neuroticism, then perhaps that item should be evaluated more closely, and developers should make a deliberate choice about whether to include it or not. Evaluating the Big Five profile of individual items can be part of explaining why their construct makes sense and thus be part of the refinement of scales as it can guide intuitions of whether it has a coherent trait space.

The current study has demonstrated that it does indeed make sense to assess the item-level correlations of interstitial scales with the Big Five in order to gain understanding of the root of their interstitiality. It also serves as a reminder that the Big Five is not a simple

structure, and that psychologically meaningful constructs can be found in the interstices between domains. While, as mentioned, many researchers do know that the structure of the Big Five is not that of a simple hierarchy, the depiction of the model as such can cause problems. A representation of the Big Five as five buckets which each hold collections of narrower traits risks creating the impression that traits should fit into one of five categories in order to be located within the Big Five. Thus, a trait that is not strongly associated with either of the five domains may be misinterpreted as being distinct from the Big Five, which may be entirely inaccurate; as Bainbridge et al. (in press) demonstrated, a trait can have no strong relationship with any one domain, yet have much of its variance explained by the Big Five. Indeed, some interstitial traits seem to be so related to the Big Five that an individual's standing on them can be approximated from composites of Big Five scores, yet they may sometimes be interpreted as distinct (Credé et al., 2016). The assumption that interstitial traits are distinct from the Big Five also means that the framework's comprehensiveness and utility may be underestimated, because researchers may not attempt to locate these traits in the Big Five framework. For instance, Carver and Scheier (2014) noted that in the case of the optimism construct and its relationships with the Big Five, researchers have, as a practical matter, "largely ignored the issue of overlap with other traits" (p. 294). The consequence of this can be that traits that lie relatively purely within one domain will be more readily located in the Big Five by researchers, while the location of interstitial traits in the framework are more likely to be ignored, and so the prevalence of interstitial scales may be underestimated. In the utmost consequence, this could have implications for the understanding of the Big Five as an organizing framework. Consider this: If the simple-structure, hierarchical depiction leads researchers to mostly locate relatively pure extensions of the five domains within the framework, it would seem that the Big Five is effective at categorizing traits in a mostly simple way, as most traits located within it would fall into five neat categories. At the same time, though, it would seem that the Big Five is not as comprehensive as wished for, as many traits would be seen as distinct. The truth may be exactly the opposite: The Big Five may be highly comprehensive, yet not be so simply hierarchical as it appears. Locating interstitial traits in the framework would provide a more nuanced view of the Big Five and the structure of personality, and underscore the need to think of the Big Five as a heterarchical or quasi-hierarchical framework. The current research has demonstrated the existence of coherent, interstitial trait scales, yet we still have no idea exactly how prevalent sound, interstitial constructs really are.

Future research can expand on the findings presented in this study in numerous ways. Of course, locating trait scales in the Big Five is a crucial step in gaining understanding of the prevalence of interstitial scales. Furthermore, research may center on investigating whether interstitial traits are evenly distributed among the interstices between the five dimensions, or if some interstices are more saturated. For examining this, one should probably not use scales, at least not exclusively; as reviewed, there are several reasons why this could be likely to skew results. Relying on the lexical hypothesis that important traits will be encoded in language and using trait adjectives could be an option instead. In the case that researchers find that many interstitial traits are located in specific interstices, perhaps that merits a reevaluation of whether these dimensions should be acknowledged more explicitly. Big Five measures do not explicitly assess facets between domains, although they do have facets that appear interstitial. But perhaps if one interstice proves particularly important, we ought to consider this particular dimension of personality in more depth. For instance, if the interstice between Extraversion and Agreeableness proves to be particularly important, one might assess an ‘interpersonal facet’. Another interesting slice of this would be to see which part of interstices between domains are most saturated; for instance, are interstitial traits more abundant between high Extraversion and high Agreeableness than between the low poles of those domains?

As for future item-level analyses of interstitial trait scales, any research should focus on using well-validated measures of both the Big Five and stand-alone scales. Furthermore, as mentioned, we cannot be sure whether four scales being convincingly interstitial is many or few; for that, more information is needed. To achieve a better understanding of this, examinations of how single-domain scales behave on the item level would be useful. This would make possible a more telling evaluation of the prevalence of interstitial scales, i.e., how many scales out of a selection of sound scales are interstitial. Future research could also contribute to the field of interstitial scales by examining an increased population of scales and by employing different criteria when selecting scales for item-level analyses. One limitation of this study is that the included scales were restricted to those included in the ESCS, which were likely not selected for their representativeness of the broader selection of scales. Systematic criteria for selection could provide a better understanding of what proportion of scales are interstitial, and how scales generally behave on the item level. One option for selection could be to follow the example of Bainbridge et al. (in press), and select scales based on citation count. That would also ensure that the included measures are widely used.

Other options could be to include all scales created in a given time period or selecting random scales.

Another interesting perspective would be to compare scales to both the Big Five and the HEXACO and see how they differ. For instance, one might investigate the same scales included in this study in a HEXACO perspective, or, ideally, investigating both at the same time. It is possible that scales are generally more purely located within single domains in one model or the other; HEXACO may divide constructs more cleanly than the Big Five or vice versa. Such a finding could suggest that one model organizes personality in a more optimal way.

There is also some potential in researching to what extent neurotic-content contamination causes interstitiality by adding a relationship to Neuroticism where, conceptually, there is none, i.e., is this scale interstitial due to the trait itself being a partly neurotic trait, or due to response bias stemming from the wording of items? We saw a potential example of this contamination with the Cognitive Failures Questionnaire, though we do not know the extent to which this happens.

#### **4.8. Strengths and limitations of the current study**

This study comes with some noteworthy strengths. Regarding sampling, we used the ESCS to assess stand-alone scales and their relationships with the Big Five domains. The ESCS is a community sample, and thus will generally be more representative of the general population than another sample consisting of a clinical population or college students. This is not the most important aspect, though, as while generalizability is usually a big part of the methodological strength of a study, it has less to say here, at least theoretically; one would assume that while mean scores of different scales and domains could change if assessed in a different population, the correlation between the scales and domains would most likely not. An exception to this is samples with larger risks of bias which could skew the results, e.g., a population in which neurotic-content contamination would occur more easily. Such a population might, for example, be a clinical population which would probably have a greater likelihood of scoring higher on Neuroticism on average (Jeronimus et al., 2016). In any case, a community sample will rarely be a negative thing if we are looking at general phenomena rather than something very population specific. In this study, we can be reasonably sure that the relationships we found will be relatively consistent across populations.

One strength pertaining to the use of the ESCS that has a relatively larger influence on our study is the sample size. This means that the data will most likely be

reasonably precise. In a smaller sample, extreme outliers could distort the data and show correlations that would not hold true in a larger sample. With an  $N$  of at least 567 in any correlation, we believe we are most likely free of this problem and can trust that we have a small margin of error and fairly accurate values (Guadagnoli & Velicer, 1988). The ESCS was also collected over a number of years, meaning it probably consisted of an attentive group of sincere responders.

Another benefit from investigating interstitial trait scales this way is that we have had minimal influence on the scales that have been selected. By using the ESCS, we have relied on other researchers to select scales that they found to be valid measures of psychological constructs rather than choosing them ourselves. As such, our pool of stand-alone constructs was determined independent of our hypotheses and was not biased by us unintentionally picking the best or worst interstitial measures. We did not include all measures in the ESCS, however, as we did select scales from the ESCS qualitatively based on what seemed sensible to analyze. The ESCS includes many measures, not all of which we deemed appropriate for this study; for selection, we decided to include scales that measured stable traits rather than measures of, say, attitudes, trait adjectives or leisure activities. Still, we were blind to the content of the measures we included at the time of selection. Using the ESCS also has the benefit that we can be fairly sure that the scales we have analyzed are reasonably well-established and frequently used, as they have been selected for inclusion in this sample.

Furthermore, these scales are unabbreviated measures of psychological constructs, save for a few cases where some items were missing (e.g., 7 items were missing from the Borderline Personality Inventory, reviewed below, though usually fewer items were missing in these cases). In a large sample such as the ESCS, it would not be unheard of to have used abbreviated measures in order to save time and money. This would have provided us with less-than-ideal stand-alone scales. As it is now, we have had gold-standard psychological measures available meaning that we do not have to consider whether our findings will translate from an abbreviated measure to a full scale.

On the flip side, a limitation of this study is that while the ESCS used unabbreviated measures, some scales were somewhat changed from the validated measure in different ways, though the severity of the changes are varied. Some scales were changed in minor ways. Of the scales included as interstitial, an example of this is the Rosenberg Self-Esteem scale. Originally a 4-point Likert scale with no neutral option, it was changed to a 5-point Likert scale with a neutral option. Other scales were changed more fundamentally;

several scales (the Borderline Personality Inventory, the Narcissism Personality Inventory, and Sensitivity to punishment) were changed from binary (e.g., true/false) or forced-choice (choosing between two different statements) response formats to a 5-point Likert scale; it is unknown to us how these changed response formats might have affected our results. Previous research on different response formats has found that response format can affect responses in some cases, though it seems unclear how much; similarly, it seems to vary whether a binary or a graded response format is superior (Boldero et al., 2015; Dolnicar & Grün, 2007; Papadatou-Pastou et al., 2013; Sharp et al., 2009). In one case, the NPI, a changed response format has been investigated by Boldero et al. (2015), who found that the forced-choice version might underestimate narcissism, though another study by Miller et al. (2018) found only minor differences from switching to a Likert scale. In any case, the matter seems to be unresolved for now, and we do not know how changed response formats might have affected correlations with the Big Five.

In one of the scales with changed response formats, the BPI, some items were also excluded. Looking qualitatively at these items, it seems that these items did not fit the new response format. While it makes sense to exclude these items, it could also leave out important information - and it could certainly bias which results we see when comparing the scale to the Big Five. The BPI is also meant to have a cutoff (Leichsenring, 1999), while we have used the scale in a continuous fashion rather than dividing total scores into a 'borderline' group and a 'non-borderline' group. Doing this differently might have yielded different results. The same is true for the BIDR scales. For this study, we did not score it as it was originally intended to be. While only scores of 6 or 7 (out of the 7-point Likert scale) are supposed to be counted, we again used it in a continuous way, assuming that a higher score equals more prominence of that particular trait. Again, scoring this scale as the authors intended it to be scored might have given us very different results.

Also pertaining to the scales are the representativeness of them. While selection of scales from the ESCS was largely unaffected by bias on our part due to them already being selected by someone else, they may have been selected originally with some kind of bias, though we cannot know. The scales accessible to us were simply the ones Goldberg (2008) decided to assess, and thus we do not know whether our scales are representative of the personality trait scale space. For instance, the included scales could be more or less related to the Big Five compared to the broader selection of trait measures. It could also be that the selection of scales skews results toward one domain; this could potentially explain why Neuroticism was so overrepresented while Openness was so underrepresented. We find this

unlikely, though, since Bainbridge et al. (in press) saw very similar patterns in their study with a different selection method, i.e. selection based on citation counts.

Other limitations include that the ESCS, while being a community sample, did have quite a homogeneous composition of participants; for instance, 98,4% were Caucasian and 83,9% had at least some college education. The demographics are not surprising considering the region from which the sample was collected (U.S. Census Bureau, 2019), though the fact that participants were recruited from a list of home-owners would be likely to have skewed the sample on some demographic parameters. Still, as we mentioned earlier, the representativeness of the sample vis-à-vis the broader population is not necessarily an issue for this particular study, though it is hard to know for sure. Since the Big Five is most well-validated in Western cultures (Funder, 2013), it could be that a sample from another part of the world would have shown different results. Particularly connections between scales and Openness would conceivably be affected differently in other cultures, as it is the factor that is consistently found to be the least universal of the five (Funder, 2013), though even in this sample from the USA it seems underrepresented.

An issue that is somewhat more likely to affect the results is that the data was collected from 1993 to 2006, and some of the included scales have since been revised. This includes the NEO PI-R that now exists in a NEO PI-3 with revision of 37 items, which improved the readability and therefore the usage with younger populations, and adults with lower levels of education (McCrae et al., 2005). A number of the included interstitial scales, as mentioned previously, have also since had issues regarding their factor structure, which have led to propositions of altered factor structures for these scales. Using the latest up-to-date scales and employing consensus-based factor structures could have changed results and would maximize validity. A further point of interest with regards to the time of the data collection is the gap between collections of data, which was up to 12 years. While personality does change during the lifetime, people usually maintain their distinctive patterns of behavior, i.e. even though mean scores of Neuroticism might have changed, an individual in the ESCS that was highly neurotic compared to the rest of the sample at the beginning of data collection will likely also be highly neurotic compared to the rest of the sample at the end (Funder, 2013). As all Big Five domains were assessed at the same time, we thus assume that relationships between the Big Five and personality variables are fairly stable.

As we are aware of no studies of this kind investigating the items of interstitial scales, we have had nothing to compare to. Though not a weakness per se, this has proven to be a challenge as consensus on many things can be hard to come by when there is no beaten path



to follow. It follows naturally that we have likely done some things differently from how another researcher might have seen fit to do. As such, we might have made some interpretations that others would not have made, and which might be challenged in the future. It is also possible that we left something on the table because we did not consider another test we could have done to achieve more valid conclusions.

#### **4.9. Conclusion**

The emergence of the Big Five as the dominating model of personality has found overwhelming consensus in personality psychology. Considering that the Big Five is supposed to represent the most coherent, sensible packaging of personality, 11 out of 41 originally included scales - or just over a quarter - of included scales being interstitial seems like a lot at first glance. From the research leading to the Big Five, it would seem to be a given that traits (and trait scales) would cluster around the five domains, and personality constructs located between the domains should be rare. However, up close it does seem that not all interstitial scales are truly interstitial at their core; only four of the original 11 scales appeared to have items that convincingly reflected the interstitiality found from the full scale in a convincing way, thus reflecting an interstitial construct. The study demonstrates that putatively interstitial trait scales can certainly stem from what appears to be a coherent, interstitial construct underneath, though they can also result from incohesive item sets. Of course, the interpretations made in this study regarding unclear interstitial scales are not to say that these constructs are not indeed interstitial; only that, in this study, the interstitiality of the full scale does not seem to be rooted in core items that all assess an interstitial construct. It is entirely possible that a revised version might yield far more convincing interstitiality for some scales that were questionable in this study.

While the four relatively convincingly interstitial scales had domains in common, they were not all located within the same interstice. Had this been the case, it could be indicative that an important dimension of personality is hiding in an interstice, yet that does not seem to be the case from these findings. This seems to support the Big Five as a sensible, comprehensive framework that manages to cover most of the variance of personality within its five domains, although there are some rare exceptions where traits are located between domains. Rare interstitial traits do not in and of themselves challenge the Big Five as a framework, so long as they are not abundant. It is curious that Neuroticism and Extraversion are involved in so many of the convincingly interstitial scales. For Neuroticism, this may not mean so much as Neuroticism was generally a dominant domain for trait affiliation; 20 scales

correlated above .30 with Neuroticism, of which nine were interstitial. For Extraversion, this is somewhat more surprising. While it was the domain that was second most related to scales, only eight scales had a correlation above .30 with it. Perhaps there are more interstitial constructs that are tangential to Extraversion, or perhaps it is simply easier to make meaningful interstitial scales that are related to Extraversion; this, we do not know. Of course, conclusive results regarding the prevalence of interstitial scales are difficult to obtain as this is the first study to focus explicitly on them, and findings should be considered with caution. Still, we believe this study has illuminated the complex subject that is interstitial scales, underscoring the need for further research.

## References

- Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What does the narcissistic personality inventory really measure? *Assessment, 18*(1), 67-87.
- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological monographs, 47*(1), i-171. <https://doi.org/10.1037/h0093360>
- Aluja, A., & Blanch, A. (2011). Neuropsychological behavioral inhibition system (BIS) and behavioral approach system (BAS) assessment: A shortened sensitivity to punishment and sensitivity to reward questionnaire version (SPSRQ-20). *Journal of personality assessment, 93*(6), 628-636.
- Ames, D. R., Rose, P., & Anderson, C. P. (2006). The NPI-16 as a short measure of narcissism. *Journal of research in personality, 40*(4), 440-450.
- Anglim, J., Horwood, S., Smillie, L. D., Marrero, R. J., & Wood, J. K. (2020). Predicting psychological and subjective well-being from personality: A meta-analysis. *Psychological Bulletin, 146*(4), 279.
- Ashton, M. C., Jackson, D. N., Paunonen, S. V., Helmes, E., & Rothstein, M. G. (1995). The Criterion Validity of Broad Factor Scales versus Specific Facet Scales. *Journal of research in personality, 29*(4), 432-442. <https://doi.org/10.1006/jrpe.1995.1025>
- Ashton, M. C., & Lee, K. (2001). A theoretical basis for the major dimensions of personality. *European Journal of Personality, 15*(5), 327-353.
- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and social psychology review, 11*(2), 150-166.
- Ashton, M. C., Lee, K., & De Vries, R. E. (2014). The HEXACO Honesty-Humility, Agreeableness, and Emotionality factors: A review of research and theory. *Personality and social psychology review, 18*(2), 139-152.
- Bainbridge, T. F., Ludeke, S. G., & Smillie, L. D. (in press). Evaluating the Big Five as an Organizing Framework for Commonly Used Psychological Trait Scales. *Journal of personality and social psychology*.
- Becker, P. (1999). Beyond the Big Five. *Personality and individual differences, 26*(3), 511-530. [https://doi.org/10.1016/S0191-8869\(98\)00168-8](https://doi.org/10.1016/S0191-8869(98)00168-8)
- Boldero, J. M., Bell, R. C., & Davies, R. C. (2015). The structure of the narcissistic personality inventory with binary and rating scale items. *Journal of personality assessment, 97*(6), 626-637.
- Botwin, M. D., & Buss, D. M. (1989). Structure of Act-Report Data: Is the Five-Factor Model of Personality Recaptured? *Journal of personality and social psychology, 56*(6), 988-1001. <https://doi.org/10.1037/0022-3514.56.6.988>

- Boudreaux, M., & Ozer, D. (2015). Five Factor Model of Personality, Assessment of. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2 ed., Vol. 9, pp. 230-235). Elsevier.
- Briggs, S. R., & Cheek, J. M. (1986). The role of factor analysis in the development and evaluation of personality scales. *Journal of personality*, *54*(1), 106-148.
- Brinkley, C. A., Diamond, P. M., Magaletta, P. R., & Heigel, C. P. (2008). Cross-Validation of Levenson's Psychopathy Scale in a Sample of Federal Female Inmates. *Assessment (Odessa, Fla.)*, *15*(4), 464-482. <https://doi.org/10.1177/1073191108319043>
- Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British journal of clinical psychology*, *21*(1), 1-16.
- Bruce, A. S., Ray, W. J., & Carlson, R. A. (2007). Understanding cognitive failures: what's dissociation got to do with it? *The American journal of psychology*, 553-563.
- Carver, C. S., & Scheier, M. F. (2014). Dispositional optimism. *Trends in cognitive sciences*, *18*(6), 293-299.
- Carver, C. S., & White, T. L. (1994). Behavioral Inhibition, Behavioral Activation, and Affective Responses to Impending Reward and Punishment: The BIS/BAS Scales. *Journal of personality and social psychology*, *67*(2), 319-333. <https://doi.org/10.1037/0022-3514.67.2.319>
- Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The journal of abnormal and social psychology*, *38*(4), 476.
- Chiesi, F., Galli, S., Primi, C., Innocenti Borgi, P., & Bonacchi, A. (2013). The Accuracy of the Life Orientation Test-Revised (LOT-R) in Measuring Dispositional Optimism: Evidence From Item Response Theory Analyses. *Journal of personality assessment*, *95*(5), 523-529. <https://doi.org/10.1080/00223891.2013.781029>
- Chmielewski, M., Bagby, R. M., Markon, K., Ring, A. J., & Ryder, A. G. (2014). Openness to experience, intellect, schizotypal personality disorder, and psychoticism: Resolving the controversy. *Journal of personality disorders*, *28*(4), 483-499.
- Connelly, B., Ones, D., & Hülshager, U. (2018). Personality in industrial, work and organizational psychology: theory, measurement and application. In D. S. Ones, N. Anderson, & C. Viswesvaran (Eds.), *The sage handbook of industrial, work and organizational psychology* (pp. 320-365). SAGE Publications Ltd., <https://doi.org/https://www-doi-org.proxy1-bib.sdu.dk/10.4135/9781473914940.n13>
- Coolican, H. (2014). Correlation and Regression. In H. Coolican (Ed.), *Research Methods and Statistics in Psychology* (6 ed., pp. 520-569). Psychology Press.
- Corry, N., Merritt, R. D., Mrug, S., & Pamp, B. (2008). The Factor Structure of the Narcissistic Personality Inventory. *Journal of personality assessment*, *90*(6), 593-600. <https://doi.org/10.1080/00223890802388590>

- Costa, J., Paul T., & McCrae, R. R. (1995). Domains and Facets: Hierarchical Personality Assessment Using the Revised NEO Personality Inventory. *Journal of personality assessment*, 64(1), 21-50. [https://doi.org/10.1207/s15327752jpa6401\\_2](https://doi.org/10.1207/s15327752jpa6401_2)
- Costa, J. P. T., & McCrae, R. R. (1976). Age differences in personality structure: a cluster analytic approach. *Journal of gerontology (Kirkwood)*, 31(5), 564-570. <https://doi.org/10.1093/geronj/31.5.564>
- Costa, P. T., & McCrae, R. R. (1985). *The NEO personality inventory : manual, form S and form R*. Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). *Neo personality inventory-revised (NEO PI-R)*. Psychological Assessment Resources Odessa, FL.
- Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical assessment, research, and evaluation*, 10(1), 7.
- Credé, M., Harms, P. D., Blacksmith, N., & Wood, D. (2016). Assessing the Utility of Compound Trait Estimates of Narrow Personality Traits. *Journal of personality assessment*, 98(5), 503-513. <https://doi.org/10.1080/00223891.2016.1170023>
- Davidson, R. J. (2003). Affective neuroscience and psychophysiology: Toward a synthesis. *Psychophysiology*, 40(5), 655-665.
- De Raad, B., & Barelds, D. P. H. (2020). Models of Personality Structure. In P. J. Corr & G. Matthews (Eds.), *The Cambridge handbook of personality psychology* (Second ed., pp. 115-128). Cambridge University Press.
- De Raad, B., & Peabody, D. (2005). Cross-culturally recurrent personality factors: analyses of three factors. *European Journal of Personality*, 19(6), 451-474. <https://doi.org/10.1002/per.550>
- De Raad, B., Perugini, M., Hrebickova, M., & Szarota, P. (1998). Lingua franca of personality - Taxonomies and structures based on the psycholexical approach. *Journal of cross-cultural psychology*, 29(1), 212-232. <https://doi.org/10.1177/0022022198291011>
- DeYoung, C. G. (2006). Higher-Order Factors of the Big Five in a Multi-Informant Sample. *Journal of personality and social psychology*, 91(6), 1138-1151. <https://doi.org/10.1037/0022-3514.91.6.1138>
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of personality and social psychology*, 93(5), 880.
- Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of personality and social psychology*, 73(6), 1246.
- Digman, J. M., & Inouye, J. (1986). Further Specification of the Five Robust Factors of Personality. *Journal of personality and social psychology*, 50(1), 116-123. <https://doi.org/10.1037/0022-3514.50.1.116>

- Distel, M. A., Trull, T. J., Willemsen, G., Vink, J. M., Derom, C. A., Lynskey, M., Martin, N. G., & Boomsma, D. I. (2009). The Five-Factor Model of Personality and Borderline Personality Disorder: A Genetic Analysis of Comorbidity. *Biological psychiatry (1969)*, *66*(12), 1131-1138. <https://doi.org/10.1016/j.biopsych.2009.07.017>
- Dolnicar, S., & Grün, B. (2007). How constrained a response: A comparison of binary, ordinal and metric answer formats. *Journal of Retailing and Consumer Services*, *14*(2), 108-122.
- Emmons, R. A. (1987). Narcissism: Theory and Measurement. *Journal of personality and social psychology*, *52*(1), 11-17. <https://doi.org/10.1037/0022-3514.52.1.11>
- Eysenck, H. J. (1947). *Dimensions of personality*. Routledge & Kegan Paul.
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *Journal of abnormal and social psychology*, *44*(3), 329-344. <https://doi.org/10.1037/h0057198>
- Fossum, T. A., & Barrett, L. F. (2000). Distinguishing evaluation from description in the personality-emotion relationship. *Personality and Social Psychology Bulletin*, *26*(6), 669-678.
- Funder, D. C. (2013). Using Personality Traits to Understand Behavior. In D. C. Funder (Ed.), *The Personality Puzzle* (6 ed., pp. 196-243). W. W. Norton & Company.
- Furnham, A. (1986). Response bias, social desirability and dissimulation. *Personality and individual differences*, *7*(3), 385-400.
- Garofalo, C., Noteborn, M. G., Sellbom, M., & Bogaerts, S. (2019). Factor structure and construct validity of the Levenson Self-Report Psychopathy Scale (LSRP): A replication and extension in Dutch nonclinical participants. *Journal of personality assessment*, *101*(5), 481-492.
- Gignac, G. E. (2013). Modeling the balanced inventory of desirable responding: Evidence in favor of a revised model of socially desirable responding. *Journal of personality assessment*, *95*(6), 645-656.
- Gleser, G. C., Cronbach, L. J., & Rajaratnam, N. (1965). Generalizability of scores influenced by multiple sources of variance. *Psychometrika*, *30*(4), 395-418. <https://doi.org/10.1007/BF02289531>
- Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's alpha reliability coefficient for Likert-type scales.
- Goldberg, L. R. (1981). Language and Individual Differences; The Wearch for Universals in Personality Lexicons. *Personality and Social Psychology*, *2*, 141-165.
- Goldberg, L. R. (1990). An alternative description of personality: the big-five factor structure. *Journal of personality and social psychology*, *59*(6), 1216-1229. <https://doi.org/10.1037//0022-3514.59.6.1216>

- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *The American psychologist*, 48(1), 26-34. <https://doi.org/10.1037//0003-066X.48.1.26>
- Goldberg, L. R. (2008). The Eugene-Springfield Community Sample: Information Available from the Research Participants. *Oregon Research Institute, Ori Technical Report Vol. 48 No. 1*.
- Gray, J. A. (1981). A Critique of Eysenck's Theory of Personality. In H. J. Eysenck (Ed.), *A Model for Personality* (1 ed., pp. 246-276). Springer, Berlin, Heidelberg. <https://doi.org/https://doi-org.proxy1-bib.sdu.dk/10.1007/978-3-642-67783-0>
- Gray, J. A. (1982). The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system. *Oxford University Press*.
- Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265.
- Gummelt, H. D., Anestis, J. C., & Carbonell, J. L. (2012). Examining the Levenson Self Report Psychopathy Scale using a Graded Response Model. *Personality and individual differences*, 53(8), 1002-1006. <https://doi.org/10.1016/j.paid.2012.07.014>
- Gurtman, M. B. (2009). Exploring Personality with the Interpersonal Circumplex. *Social and personality psychology compass*, 3(4), 601-619. <https://doi.org/10.1111/j.1751-9004.2009.00172.x>
- Gurven, M., von Rueden, C., Massenkoff, M., Kaplan, H., & Lero Vie, M. (2013). How Universal Is the Big Five? Testing the Five-Factor Model of Personality Variation Among Forager-Farmers in the Bolivian Amazon. *Journal of personality and social psychology*, 104(2), 354-370. <https://doi.org/10.1037/a0030841>
- Heshmati, R., & Azmoodeh, S. (2017). Study of alexithymia trait based on Big-Five Personality Dimensions. *Mediterranean Journal of Clinical Psychology*, 5(3).
- Heubeck, B. G., Wilkinson, R. B., & Cologon, J. (1998). A second look at Carver and White's (1994) BIS/BAS scales. *Personality and individual differences*, 25(4), 785-800. [https://doi.org/10.1016/S0191-8869\(98\)00124-X](https://doi.org/10.1016/S0191-8869(98)00124-X)
- Hofstee, W. K., De Raad, B., & Goldberg, L. R. (1992). Integration of the big five and circumplex approaches to trait structure. *Journal of personality and social psychology*, 63(1), 146.
- Hofstee, W. K. B., Kiers, H. A. L., De Raad, B., Goldberg, L. R., & Ostendorf, F. (1997). A Comparison of Big-Five structures of personality traits in Dutch, English, and German. *European Journal of Personality*, 11(1), 15-31. [https://doi.org/10.1002/\(SICI\)1099-0984\(199703\)11:1<15::AID-PER273>3.0.CO;2-8](https://doi.org/10.1002/(SICI)1099-0984(199703)11:1<15::AID-PER273>3.0.CO;2-8)



- Hopwood, C. J., & Donnellan, M. B. (2010). How Should the Internal Structure of Personality Inventories Be Evaluated? *Personality and social psychology review*, 14(3), 332-346. <https://doi.org/10.1177/1088868310361240>
- Hough, L. M. (1992). The 'Big Five' Personality Variables--Construct Confusion: Description Versus Prediction. *Human performance*, 5(1-2), 139-155. <https://doi.org/10.1080/08959285.1992.9667929>
- Hough, L. M., & Ones, D. S. (2001). The structure, measurement, validity, & use of personality variables in industrial, work, and organizational psychology. In D. S. O. N. Anderson, H. K. Sinangil, & C Viswesvaran & (Eds.) (Eds.), *Handbook of industrial, work, and organizational psychology* (Vol. 1, pp. 233–277). Sage.
- Hough, L. M., Oswald, F. L., & Ock, J. (2015). Beyond the Big Five: New Directions for Personality Research and Practice in Organizations. *Annual review of organizational psychology and organizational behavior*, 2(1), 183-209. <https://doi.org/10.1146/annurev-orgpsych-032414-111441>
- Huang, C., & Dong, N. (2011). Factor Structures of the Rosenberg Self-Esteem Scale: A Meta-Analysis of Pattern Matrices. *European journal of psychological assessment : official organ of the European Association of Psychological Assessment*, 28(2), 132-138. <https://doi.org/10.1027/1015-5759/a000101>
- Jeronimus, B., Kotov, R., Riese, H., & Ormel, J. (2016). Neuroticism's prospective association with mental disorders halves after adjustment for baseline symptoms and psychiatric history, but the adjusted association hardly decays with time: a meta-analysis on 59 longitudinal/prospective studies with 443 313 participants. *Psychological medicine*, 46(14), 2883.
- John, O. P. (2021). History, Measurement, and Conceptual Elaboration of the Big-Five Trait Taxonomy: The Paradigm Matures. In O. P. John & R. W. Robins (Eds.), *Handbook of personality: theory and research* (Fourth ed., pp. 35-82). The Guilford Press.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). Big five inventory. *Journal of personality and social psychology*.
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: a qualitative and quantitative review. *Journal of applied psychology*, 87(4), 765.
- Karpman, B. (1948). The myth of the psychopathic personality. *American Journal of Psychiatry*, 104(9), 523-534.
- Knutson, B., & Bhanji, J. (2006). Neural Substrates for Emotional Traits?: The Case of Extraversion. In T. Canli (Ed.), *Biology of personality and individual differences* (pp. 116-132). Guilford Press.
- Kosslyn, S. M., Cacioppo, J. T., Davidson, R. J., Hugdahl, K., Lovallo, W. R., Spiegel, D., & Rose, R. (2002). Bridging psychology and biology: the analysis of individuals in groups. *American psychologist*, 57(5), 341.



- Krueger, R. F., & Markon, K. E. (2014). The Role of the DSM-5 Personality Trait Model in Moving Toward a Quantitative and Empirically Based Approach to Classifying Personality and Psychopathology. *Annual review of clinical psychology*, 10(1), 477-501. <https://doi.org/10.1146/annurev-clinpsy-032813-153732>
- Kubarych, T. S., Deary, I. J., & Austin, E. J. (2004). The Narcissistic Personality Inventory: factor structure in a non-clinical sample. *Personality and individual differences*, 36(4), 857-872. [https://doi.org/10.1016/S0191-8869\(03\)00158-2](https://doi.org/10.1016/S0191-8869(03)00158-2)
- Könen, T., & Karbach, J. (2020). Self-Reported Cognitive Failures in Everyday Life: A Closer Look at Their Relation to Personality and Cognitive Performance. *Assessment (Odessa, Fla.)*, 27(5), 982-995. <https://doi.org/10.1177/1073191118786800>
- Larson, G. E., Alderton, D. L., Neideffer, M., & Underhill, E. (1997). Further evidence on dimensionality and correlates of the Cognitive Failures Questionnaire. *The British journal of psychology*, 88(1), 29-38. <https://doi.org/10.1111/j.2044-8295.1997.tb02618.x>
- Leichsenring, F. (1999). Development and First Results of the Borderline Personality Inventory: A Self-Report Instrument for Assessing Borderline Personality Organization. *Journal of personality assessment*, 73(1), 45-63. <https://doi.org/10.1207/S15327752JPA730104>
- Leone, L., Perugini, M., Bagozzi, R. P., Pierro, A., & Mannetti, L. (2001). Construct validity and generalizability of the Carver-White behavioural inhibition system/behavioural activation system scales. *European Journal of Personality*, 15(5), 373-390. <https://doi.org/10.1002/per.415>
- Levenson, M. R., Kiehl, K. A., & Fitzpatrick, C. M. (1995). Assessing Psychopathic Attributes in a Noninstitutionalized Population. *Journal of personality and social psychology*, 68(1), 151-158. <https://doi.org/10.1037/0022-3514.68.1.151>
- Li, A., & Bagger, J. (2007). The Balanced inventory of desirable responding (BIDR) a reliability generalization study. *Educational and psychological measurement*, 67(3), 525-544.
- Ludeke, S. G., Bainbridge, T. F., Liu, J., Zhao, K., Smillie, L. D., & Zettler, I. (2019). Using the Big Five Aspect Scales to translate between the HEXACO and Big Five personality models. *Journal of personality*, 87(5), 1025-1038.
- Matthews, G., Coyle, K., & Craig, A. (1990). Multiple factors of cognitive failure and their relationships with stress vulnerability. *Journal of psychopathology and behavioral assessment*, 12(1), 49-65. <https://doi.org/10.1007/BF00960453>
- McCrae, R. R. (2020). The Five-Factor Model of Personality: Consensus and Controversy. In G. Matthews & P. J. Corr (Eds.), *The Cambridge Handbook of Personality Psychology* (2 ed., pp. 129-141). Cambridge University Press. <https://doi.org/DOI:10.1017/9781108264822.013>

- McCrae, R. R., & Costa, P. T. (1985a). Comparison of EPI and psychoticism scales with measures of the five-factor model of personality. *Personality and individual differences*, 6(5), 587-597. [https://doi.org/10.1016/0191-8869\(85\)90008-X](https://doi.org/10.1016/0191-8869(85)90008-X)
- McCrae, R. R., & Costa, P. T. (1985b). Updating Norman's "Adequate Taxonomy": Intelligence and Personality Dimensions in Natural Language and in Questionnaires. *Journal of personality and social psychology*, 49(3), 710-721. <https://doi.org/10.1037/0022-3514.49.3.710>
- McCrae, R. R., & John, O. P. (1992). An Introduction to the Five-Factor Model and Its Applications. *Journal of personality*, 60(2), 175-215. <https://doi.org/10.1111/j.1467-6494.1992.tb00970.x>
- McCrae, R. R., Martin, T. A., & Costa, P. T. (2005). Age Trends and Age Norms for the NEO Personality Inventory-3 in Adolescents and Adults. *Assessment (Odessa, Fla.)*, 12(4), 363-373. <https://doi.org/10.1177/1073191105279724>
- Miller, J. D. (2012). Five-Factor Model Personality Disorder Prototypes: A Review of Their Development, Validity, and Comparison to Alternative Approaches: Five-Factor Model Personality Disorder Prototypes. *Journal of personality*, 80(6), 1565-1591. <https://doi.org/10.1111/j.1467-6494.2012.00773.x>
- Miller, J. D., Gaughan, E. T., & Pryor, L. R. (2008). The Levenson Self-Report Psychopathy Scale: An Examination of the Personality Traits and Disorders Associated With the LSRP Factors. *Assessment (Odessa, Fla.)*, 15(4), 450-463. <https://doi.org/10.1177/1073191108316888>
- Miller, J. D., Gentile, B., Carter, N. T., Crowe, M., Hoffman, B. J., & Campbell, W. K. (2018). A comparison of the nomological networks associated with forced-choice and likert formats of the narcissistic personality inventory. *Journal of personality assessment*, 100(3), 259-267.
- Mooradian, T., Matzler, K., Uzelac, B., & Bauer, F. (2016). Perspiration and inspiration: Grit and innovativeness as antecedents of entrepreneurial success. *Journal of Economic Psychology*, 56, 232-243.
- Mooradian, T. A., Davis, M., & Matzler, K. (2011). Dispositional Empathy and the Hierarchical Structure of Personality. *The American journal of psychology*, 124(1), 99-109. <https://doi.org/10.5406/amerjpsyc.124.1.0099>
- Moshagen, M., Thielmann, I., Hilbig, B. E., & Zettler, I. (2019). Meta-Analytic Investigations of the HEXACO Personality Inventory(-Revised): Reliability Generalization, Self-Observer Agreement, Intercorrelations, and Relations to Demographic Variables. *Zeitschrift für Psychologie*, 227(3), 186-194. <https://doi.org/10.1027/2151-2604/a000377>
- Musek, J. (2007). A general factor of personality: Evidence for the Big One in the five-factor model. *Journal of research in personality*, 41(6), 1213-1233. <https://doi.org/10.1016/j.jrp.2007.02.003>

- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of abnormal and social psychology*, 66(6), 574-583. <https://doi.org/10.1037/h0040291>
- Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of organizational behavior*, 17(6), 609-626. [https://doi.org/10.1002/\(SICI\)1099-1379\(199611\)17:6<609::AID-JOB1828>3.0.CO;2-K](https://doi.org/10.1002/(SICI)1099-1379(199611)17:6<609::AID-JOB1828>3.0.CO;2-K)
- Papadatou-Pastou, M., Martin, M., & Munafò, M. R. (2013). Measuring hand preference: a comparison among different response formats using a selected sample. *Laterality: Asymmetries of Body, Brain and Cognition*, 18(1), 68-107.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of Personality and Social Psychological Attitudes* (pp. 17-59). Academic Press.
- Paunonen, S. V., & Jackson, D. N. (2000). What is beyond the big five? Plenty! *Journal of personality*, 68(5), 821-835.
- Peabody, D., & Goldberg, L. R. (1989). Some determinants of factor structures from personality-trait descriptors. *Journal of personality and social psychology*, 57(3), 552.
- Picardi, A., Toni, A., & Caroppo, E. (2005). Stability of Alexithymia and Its Relationships with the 'Big Five' Factors, Temperament, Character, and Attachment Style. *Psychotherapy and psychosomatics*, 74(6), 371-378. <https://doi.org/10.1159/000087785>
- Pickering, A. D., Corr, P. J., & Gray, J. A. (1999). Interactions and reinforcement sensitivity theory : A theoretical analysis of Rusting and Larsen (1997). Authors' reply. *Personality and individual differences*, 26(2), 357-372.
- Psederska, E., Yankov, G. P., Bozgunov, K., Popov, V., Vasilev, G., & Vassileva, J. (2020). Validation of the Levenson Self-Report Psychopathy Scale in Bulgarian Substance-Dependent Individuals. *Frontiers in psychology*, 11, 1110.
- Pullmann, H., & Allik, J. (1999). The Rosenberg Self-Esteem Scale: its dimensionality, stability and personality correlates in Estonian. *Personality and individual differences*, 28(4), 701-715. [https://doi.org/10.1016/S0191-8869\(99\)00132-4](https://doi.org/10.1016/S0191-8869(99)00132-4)
- Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of research in personality*, 44(1), 53-61. <https://doi.org/10.1016/j.jrp.2009.10.005>
- Raskin, R., & Terry, H. (1988). A principal-components analysis of the Narcissistic Personality Inventory and further evidence of its construct validity. *Journal of personality and social psychology*, 54(5), 890.
- Reid-Seiser, H. L., & Fritzsche, B. A. (2001). The usefulness of the NEO PI-R positive presentation management scale for detecting response distortion in employment

- contexts. *Personality and individual differences*, 31(4), 639-650.  
[https://doi.org/10.1016/S0191-8869\(00\)00168-9](https://doi.org/10.1016/S0191-8869(00)00168-9)
- Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The Structure of Conscientiousness: an Empirical Investigation Based on Seven Major Personality Questionnaires. *Personnel psychology*, 58(1), 103-139. <https://doi.org/10.1111/j.1744-6570.2005.00301.x>
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale. *Personality & social psychology bulletin*, 27(2), 151-161.  
<https://doi.org/10.1177/0146167201272002>
- Robins, R. W., Tracy, J. L., Trzesniewski, K., Potter, J., & Gosling, S. D. (2001). Personality Correlates of Self-Esteem. *Journal of research in personality*, 35(4), 463-482.  
<https://doi.org/10.1006/jrpe.2001.2324>
- Rosenberg, M. (1979). *Conceiving the self*. Basic Books.
- Rosenthal, S. A., Matthew Montoya, R., Ridings, L. E., Rieck, S. M., & Hooley, J. M. (2011). Further evidence of the Narcissistic Personality Inventory's validity problems: A meta-analytic investigation—Response to Miller, Maples, and Campbell (this issue). *Journal of research in personality*, 45(5), 408-416.  
<https://doi.org/10.1016/j.jrp.2011.06.004>
- Salgado, J. F. (2017). Bandwidth-fidelity dilemma. *Encyclopedia of Personality and Individual Differences; Zeigler-Hill, V., Shackelford, TK, Eds*, 1-4.
- Saucier, G., & Goldberg, L. R. (1996). Evidence for the Big Five in analyses of familiar English personality adjectives. *European Journal of Personality*, 10(1), 61-77.  
[https://doi.org/10.1002/\(SICI\)1099-0984\(199603\)10:1<61::AID-PER246>3.0.CO;2-D](https://doi.org/10.1002/(SICI)1099-0984(199603)10:1<61::AID-PER246>3.0.CO;2-D)
- Saucier, G., & Goldberg, L. R. (1998). What is beyond the Big Five? *Journal of personality*, 66, 495-524.
- Saucier, G., & Goldberg, L. R. (2001). Lexical Studies of Indigenous Personality Factors: Premises, Products, and Prospects. *Journal of personality*, 69(6), 847-879.  
<https://doi.org/10.1111/1467-6494.696167>
- Saucier, G., Hampson, S. E., & Goldberg, L. R. (2000). Cross-language studies of lexical personality factors. In S. E. Hampson (Ed.), *Advances in personality psychology* (Vol. 1, pp. 1-36).
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing Optimism From Neuroticism (and Trait Anxiety, Self-Mastery, and Self-Esteem): A Reevaluation of the Life Orientation Test. *Journal of personality and social psychology*, 67(6), 1063-1078. <https://doi.org/10.1037/0022-3514.67.6.1063>
- Schroeders, U., Kubera, F. R., & Gnambs, T. (2021). The Structure of the Toronto Alexithymia Scale (TAS-20): A Meta-Analytic Confirmatory Factor Analysis.

- Segarra, P., Poy, R., López, R., & Moltó, J. (2014). Characterizing Carver and White's BIS/BAS subscales using the Five Factor Model of personality. *Personality and individual differences, 61-62*, 18-23. <https://doi.org/10.1016/j.paid.2013.12.027>
- Sharp, E. S., Suthers, K. M., Crimmins, E., & Gatz, M. (2009). Does "no" mean "sometimes"? How older adults respond to the same depression symptoms with different response formats. *Clinical gerontologist, 32(4)*, 371-378.
- Sharpe, J. P., Martin, N. R., & Roth, K. A. (2011). Optimism and the Big Five factors of personality: Beyond Neuroticism and Extraversion. *Personality and individual differences, 51(8)*, 946-951. <https://doi.org/10.1016/j.paid.2011.07.033>
- Smits, D. J. M., & Boeck, P. D. (2006). From BIS/BAS to the big five. *European Journal of Personality, 20(4)*, 255-270. <https://doi.org/10.1002/per.583>
- Strus, W., Ciecuch, J., & Rowiński, T. (2014). The Circumplex of Personality Metatraits: A Synthesizing Model of Personality Based on the Big Five. *Review of general psychology, 18(4)*, 273-286. <https://doi.org/10.1037/gpr0000017>
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International journal of medical education, 2*, 53-55. <https://doi.org/10.5116/ijme.4dfb.8dfd>
- Taylor, G. J., Ryan, D., & Bagby, R. M. (1985). Toward the Development of a New Self-Report Alexithymia Scale. *Psychotherapy and psychosomatics, 44(4)*, 191-199. <https://doi.org/10.1159/000287912>
- Torrubia, R., Ávila, C., Moltó, J., & Caseras, X. (2001). The Sensitivity to Punishment and Sensitivity to Reward Questionnaire (SPSRQ) as a measure of Gray's anxiety and impulsivity dimensions. *Personality and individual differences, 31(6)*, 837-862. [https://doi.org/10.1016/S0191-8869\(00\)00183-5](https://doi.org/10.1016/S0191-8869(00)00183-5)
- Trull, T. J., Widiger, T. A., Lynam, D. R., & Costa, P. T. (2003). Borderline Personality Disorder From the Perspective of General Personality Functioning. *Journal of abnormal psychology (1965), 112(2)*, 193-202. <https://doi.org/10.1037/0021-843X.112.2.193>
- Tupes, E. C., & Christal, R. E. (1961). Recurrent Personality Factors Based on Trait Ratings. *Technical Report, U.S. Air Force, Lackland Air Force Base, TX, 60*. <https://doi.org/10.1111/j.1467-6494.1992.tb00973.x>
- U, S., Census Bureau. (2019). *American Community Survey 5-year estimates*. Retrieved from Census Reporter Profile page for Eugene-Springfield, OR Metro Area. Retrieved 30.05.2021 from <http://censusreporter.org/profiles/31000US21660-eugene-springfield-or-metro-area/>
- Ueno, M., Maeda, M., & Komaki, G. (2014). Different subgroups of high-scorers on the TAS-20 based on the big five personality traits. *Personality and individual differences, 68*, 71-76.



- Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The general factor of personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of research in personality*, 44(3), 315-327.
- Visser, B. A. (2018). Narcissism and the Big Five/HEXACO Models of Personality. In (pp. 205-212). Springer International Publishing. [https://doi.org/10.1007/978-3-319-92171-6\\_22](https://doi.org/10.1007/978-3-319-92171-6_22)
- Vittersø, J. (2001). Personality traits and subjective well-being: Emotional stability, not extraversion, is probably the important predictor. *Personality and individual differences*, 31(6), 903-914.
- Wallace, J. C. (2004). Confirmatory factor analysis of the cognitive failures questionnaire: evidence for dimensionality and construct validity. *Personality and individual differences*, 37(2), 307-324.
- Wallace, J. C., Kass, S. J., & Stanny, C. J. (2002). The cognitive failures questionnaire revisited: dimensions and correlates. *The Journal of general psychology*, 129(3), 238-256.
- Watters, C. A. (2018). *Investigations into the Structural Validity and Divergent Domain Scoring Methods of the Personality Inventory for DSM-5 (PID-5)*
- Watters, C. A., & Bagby, R. M. (2018). A meta-analysis of the five-factor internal structure of the Personality Inventory for DSM-5. *Psychological Assessment*, 30(9), 1255.
- Watters, C. A., Sellbom, M., Uliaszek, A. A., & Bagby, R. M. (2019). Clarifying the Interstitial Nature of Facets From the Personality Inventory for DSM-5 Using the Five Factor Model of Personality. *Personality disorders*, 10(4), 330-339. <https://doi.org/10.1037/per0000327>
- Wilhelm, O., Witthöft, M., & Schipolowski, S. (2010). Self-reported cognitive failures. *Journal of Individual Differences*.
- Williams, K. M., Nathanson, C., & Paulhus, D. L. (2003). Structure and validity of the self-report psychopathy scale-III in normal populations. 111th annual convention of the American Psychological Association,
- Williams, K. M., & Paulhus, D. L. (2004). Factor structure of the Self-Report Psychopathy scale (SRP-II) in non-forensic samples. *Personality and individual differences*, 37(4), 765-778. <https://doi.org/10.1016/j.paid.2003.11.004>
- Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the Four-Factor Structure of Psychopathy in College Students Via Self-Report. *Journal of personality assessment*, 88(2), 205-219. <https://doi.org/10.1080/00223890701268074>
- Wilmot, M. P., DeYoung, C. G., Stillwell, D., & Kosinski, M. (2016). Self-Monitoring and the Metatraits. *Journal of personality*, 84(3), 335-347. <https://doi.org/10.1111/jopy.12162>

- Zuckerman, M., Joireman, J., Kraft, M., & Kuhlman, D. M. (1999). Where do motivational and emotional traits fit within three factor models of personality? *Personality and individual differences*, 26(3), 487-504. [https://doi.org/10.1016/S0191-8869\(98\)00161-5](https://doi.org/10.1016/S0191-8869(98)00161-5)
- Zuckerman, M., Kuhlman, D. M., Joireman, J., Teta, P., & Kraft, M. (1993). A comparison of three structural models for personality: the big three, the big five, and the alternative five. *Journal of personality and social psychology*, 65(4), 757.