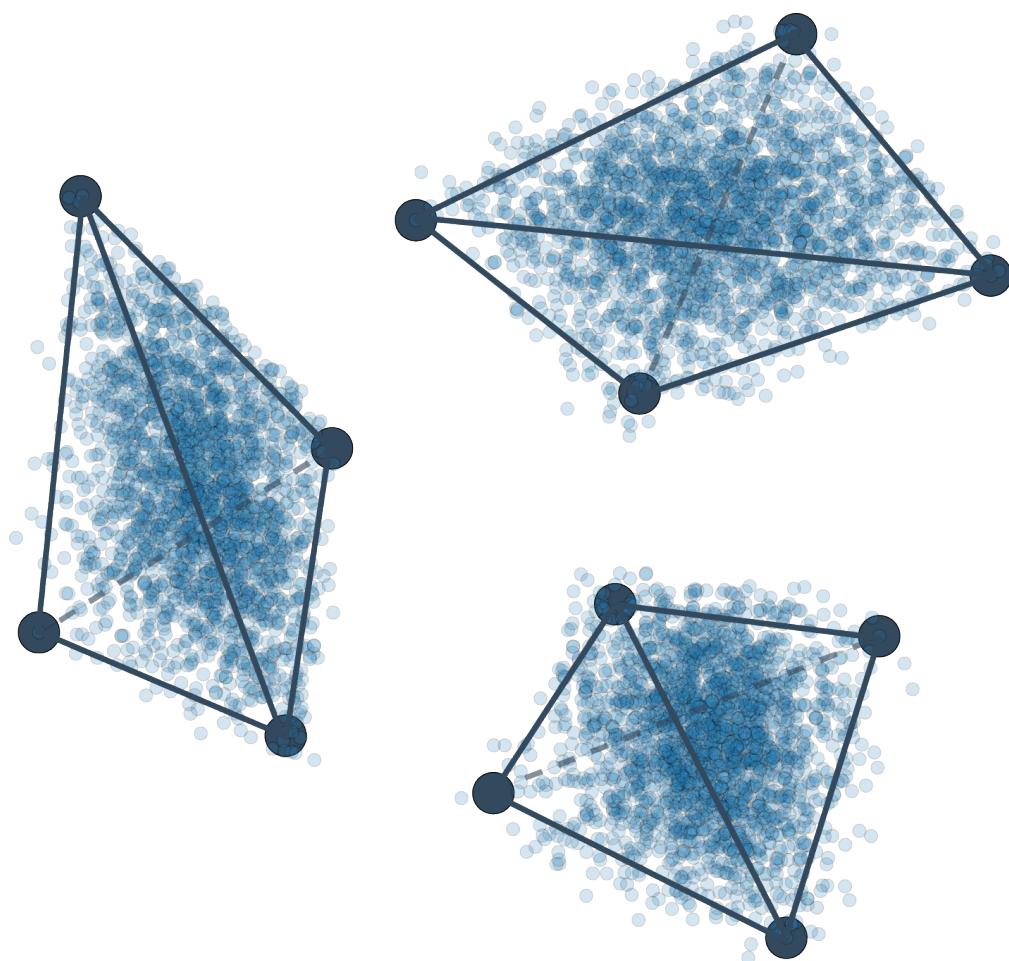


Personality Archetypes Reveal Fundamental Evolutionary Strategies for Social Behavior

Master Thesis

Ulf Aslak



Preface

This thesis was submitted as a final requirement for obtaining the degree of Master of Digital Media Engineering (cand. polyt.) at the Technical University of Denmark (DTU). It was carried out from january to july of 2016 as a collaboration between the Uri Alon Lab, at the Weizmann Institute of Science (WIS) in Israel, and the Social Fabric project at University of Copenhagen and DTU Compute. Throughout the project I lived in Israel and appeared in the lab where I, aside from conducting my own research, participated in group meetings, project sparing, and various other research activities. The project was mainly supervised by Uri Alon (Professor, WIS, Dept. of Molecular Cell Biology) and Sune Lehmann (Associate Professor, DTU, DTU Compute). Hila Sheftel and Pablo Szekely (PhD students, WIS, Dept. of Molecular Cell Biology) also took part in supervising and mentoring.

I want to thank Uri for actively engaging in my work with childlike curiosity and genuine interest, and for never failing to identify the paths to the most interesting research objectives. I thank Sune for always believing in me and my motivations, for facilitating the whole arrangement and for his infinite resource of spirit. I thank Hila and Pablo for invaluable scientific guidance and mentoring, Avi Mayo for technical discussions of various sorts, all the members of the Theatre Lab for inspiration and enlightenment, and everyone in the Uri Alon Lab for accepting me as a member of their nurturing research environment. I also thank my dear friend Daniel Svendsen for endless videocall discussions about the hardship of research and the meaning of life.

For economical support I want to thank Oticon Fonden, Reinholdt W. Jorck og Hustrus Fond, Ingeniørforeningen i Danmarks og Berg-Nielsens Studie- og Støttefond, KnuHøjgaards Fond, Dansk Tennisfond, Otto Mønsteds Fond and Dansk-Israelsk Studiefond.

Finally, I thank Vivi Lai for her tremendous courage and faith, and for letting me travel so far away for so long and still be right there, in the hammock by the window, upon my return.



Ulf Aslak
July 4th, 2016

Abstract

Evolutionary psychology suggests that different strategies for performing evolutionarily important tasks shaped our personality differences. But what these strategies are remain guesswork. Here, we apply Pareto Task Inference to infer the strategies and their behavioral correlates. We find the same six personality archetypes to consistently emerge from archetypal analysis applied to a diverse selection of datasets describing personality of millions in broad western culture. Enrichment of demographic attributes, personality facets, values and behavior near the archetypes reveal behavioral strategies, some of which are mentioned in the literature. We test how well distance from personality archetypes can be predicted for individuals using behavioral data, and compare the resulting accuracy to that of prediction models for Big Five traits and Big Five questionnaire principal components. We find that distance from personality archetypes can be predicted more accurately than the other personality targets. Based on the findings and supported by literature in evolutionary psychology we argue that personality archetypes can be viewed as fundamental behavioral strategies.

Resume

Evolutionær psykologi foreslår at forskellige strategier for at udføre evolutionært vigtige opgaver i livet har formet vores individuelle forskelle som mennesker. Men hvad disse strategier er har indtil videre været gætværk. Her anvender vi Pareto Task Inference til at udlede strategierne og deres adfærdsmæssige korrelater. Vi finder, at de samme seks personlighedsarketyper opstår, påkonsistent vis, på tværs af et diverst ensemble af datasæt, som til sammen beskriver personligheder af millioner af mennesker i den vestlige verden. Berigelse nær arketyper af variable som demografiske attributter, personlighedsfacetter, menneskelige værdier og adfærd målt gennem smartphones, giver evidens for forskellige personlighedsstrategier hvoraf nogle ræsonnerer med kendte strategier fra litteraturen. Vi sammenligner hvor godt individuel afstand fra personlighedsarketyper, Big Five samt principiel-komponenter af varians i spørgeskema-svar kan modelleres ud fra adfærdsdata. Simuleringen viser at afstand fra personlighedsarketyper kan modelleres bedst. På baggrund af disse resultater, og understøttet af litteratur i evolutionær psykologi, foreslår vi at de fundne personlighedsarketyper kan betragtes som fundamentale adfærdsstrategier.

Table of Contents

Preface	i
Abstract	ii
Resume	iii
1 Introduction	1
2 Background Theory	4
2.1 Evolutionary Psychology	4
2.1.1 What is personality and how does it relate to behavior?	5
2.1.2 How is personality measured?	6
2.1.3 How does evolutionary psychology treat personality?	9
2.2 Pareto Task Inference	13
2.2.1 Finding traits that are optimized for objectives	13
2.2.2 Inferring objectives from enrichment near archetypes	16
2.2.3 Inference of archetypes	17
3 Datasets	20
3.1 Sensible DTU	20
3.2 myPersonality	22
3.3 Midlife Development in the U.S	22
3.4 SAPA-Project	23
4 Analysis and Results	24
4.1 Six consensus archetypes emerge across seven datasets	24
4.2 Enrichment and correlation reveal distinct strategies	27
4.2.1 Analysis and data	28
4.2.2 Summary and strategy suggestions	31
4.3 Archetypes separate data better than traits	35
5 Discussion	37
6 Conclusions	39
6.1 Future Research	40
7 Methods	42
7.1 SensibleDTU data pre-processing	42
7.1.1 Feature Engineering	43
7.1.2 Variable Transformation	44
7.1.3 Outlier Detection	45
7.1.4 Personality Data Extraction	45

7.2	Analysis	45
7.2.1	Computing consensus archetypes	46
7.2.2	Computing enrichments	47
7.2.3	Computing correlations	48
7.2.4	Classification	48
	References	50
	Appendices	
	A Full List of Extracted Indicators	56
	B A Note on Feature Engineering	59
	C Questionnaire Enrichment Table	61
	D Basic Human Values Enrichment Table	63
	E Behavior Correlation Table	64
	F Theory of Basic Human Values	65
	G Distributions of Big Five datasets	66
	H Code for enrichment of feautres near archetypes	67
	I Online	69

1 Introduction

From one perspective humans are all the same. We think with our brains and act with our bodies, feel sadness and share laughter. Yet we are also different from each other. No two people are exactly the same, and even for identical twins the firstborn will always be the firstborn and will never know what it's like to be the secondborn. Somewhere between our common humanity and individual differences lies *personality*. We are all variations of the same basic template, and personality captures those differences: how our shared human nature can manifest itself in different behaviors, feelings and ways of thinking.

Personality can be measured. Throughout the 20th century, many have concerned themselves with understanding and quantifying personality. Rooted in the assumption that personality attributes that are truly important should manifest themselves in language^[1], various systems for measuring personality using questionnaire methods have been developed. Models like the *Big Five*, the *16PF* and the *Eysenck* model score personality along dimensions such as *extraversion* and *neuroticism* and are well established and widely applied. The Big Five model especially provides an effective and intuitive framework and is generally recognized as the gold standard instrument of measurement^[2]. It measures personality along five nearly orthogonal dimensions called *traits* which are *openness to experience*, *conscientiousness*, *extraversion*, *agreeableness* and *neuroticism*. It is argued to represent the most evolutionarily plausible way of representing human personality, due to the property that its traits are nearly independent^[3,4].

Personality can be viewed as behavioral strategy. Past studies in evolutionary biology did not recognize personality as important to an organism's ability to propagate its genes, i.e. its *fitness*, because other factors like physical attributes were considered far more important. Recent findings from twin-studies have, however, shown that personality is 30 – 50% inherited suggesting that there are fitness benefits to different genetically inheritable aspects of personality^[5,6]. This finding, joined with the observation that personality differs between people, has roused significant interest in the emerging field of evolutionary psychology (EP). EP established the following notion: if a human has to perform a certain set of tasks to propagate its genes - such as self-protection, status affirmation and mate acquisition - and these tasks can be performed, from an evolutionary perspective, equally well in a multitude of ways, the personality held by an individual is its adopted strategy for performing these tasks^[2]. Hereinafter 'personality' and 'strategy' will be used interchangeably.

Frequency dependence causes individual personality differences. Evolutionary psychologists argue that variation in personality between individuals arise because there is no single optimal strategy for performing life's many tasks^[2]. In fact there are many, which causes people's personalities to vary greatly. This argument

is best presented using an example. A fundamental strategy to cope with life's tasks could be to rely on trust and benevolence to establish a network to rely on, conform with society's rules and opt for stability over risk. Another fundamental strategy might then be to take advantage of people's weaknesses, lie to get ahead and trust no one. If everyone adopted the first strategy it would be too easy to succeed using the second, and if everyone adopted the second there would be no one for the second to take advantage of. The fitness associated with each strategy is therefore *frequency dependent*, which drives constant fluctuations in fitness as the frequency of each changes. Furthermore, because these fluctuations occur asynchronously across local environments, variation will be perpetuated globally.

Individual personality is a trade-off between fundamental strategies. To fully adopt one fundamental strategy is rarely a robust solution due to the fitness fluctuations that result from frequency dependence. An individual will therefore tend to adopt a strategy that is a mix of many fundamental strategies. This can be thought of as a multi-objective optimization problem: individuals must ensure maximal fitness but personal strategies must obey inherent trade-offs between performance of fundamental strategies. It confines solutions to a surface in performance space called the *Pareto front*. The Pareto front defines the boundary where increasing performance in one strategy comes at the expense of performance in others. Evolution tends to not allow individuals to stay off the Pareto front because their fitness advantage would be inferior to those on it, and as such personalities of people should be confined to the Pareto front where they remain *Pareto optimal* trade-offs between fundamental strategies.

Fundamental personality strategies can be inferred from archetypes. Recent theoretical advances due to Uri Alon et. al. enable the inference of evolutionary biological objectives subject to performance trade-offs^[7,8]. The Pareto Task Inference (ParTI) principle developed by Alon et. al. establishes that for a representatively large population of individuals, the distribution of points in trait-space will fall on a simplex (a polytope of $d_{dimensions} + 1$ vertices that evaluates to a triangle in 2D, a tetrahedron in 3D, a 4-simplex in 4D, etc.) either in the original trait-space or in a lower-dimensional subspace spanned by the first principle components. The vertices of this simplex are *archetypes* and each represents a combination of traits that are suited for a specific objective. Depending on the system that is analyzed the objectives that are inferred can be single tasks such as nut cracking or branch grasping in the case of Darwinian finches, or something more complex like personality strategies which are optimized to solve multiple tasks in different ways in the case of humans. The ParTI principle has proven successful in explaining functions of cancer cells^[8], morphology of ammonite shells^[9], mass and longevity design of mammals^[10], and more^[7,11].

This study sets out to infer fundamental personality strategies. In the established context there is a gap in our knowledge about fundamental personality strategies. Given the ready availability of instruments to measure personality, a whole field of theory which treats personality as behavioral strategy, and a proven theoretical principle for inferring evolutionary objectives from traits there is strong motivation to pursue an understanding of fundamental personality strategies.

The research project documented here is novel and ongoing, and as such this thesis presents the trajectory of first exploration with strong emphasis on outlook. The research idea is due to Hila Sheftel, Uri Alon and Ulf Aslak. In brief, the reader will be presented with the following narrative: Six personality archetypes are found across a multitude of independent Big Five datasets using the ParTI principle. By studying how different personal attributes and questionnaire items have higher/lower means or are in over/under-representation for individuals near each archetype (i.e. are *enriched/depleted*) and how certain types of behavior correlate with distance from archetypes, an understanding of possible strategies for each archetype emerges. For this purpose the book *Evolutionary Psychology* (2014) by Workman et. al.^[2] is used as a literary reference framework for establishing sound arguments for which scientifically investigated evolutionary strategies that the archetypes may correspond to. To infer the viability of the archetypes a supervised machine learning approach is taken to infer the predictive power between behavior and distance from the archetypes, as compared to Big Five traits as well as principle components of questionnaire items. The result of this analysis gives a preliminary picture of what the landscape of fundamental personality strategies looks like. Since the discoveries are of sensitive nature, due discussion is made to clarify doubts about their implications. Furthermore, results relating to race and gender are omitted for moral reasons.

The relevant theory is presented in Section 2 which primarily treats personality psychology, EP and ParTI. Section 3 describes the datasets used for analysis. Section 4, which is the main section, presents the conducted research. It reads like an investigative research story separated into shorter subsections that each has a single premise. Section 5 discusses the results, their uncertainties and limitations, implications and impact. Section 6 sums up the outcomes of the study in a list of conclusions, and gives an outlook on future directions for the project. The methods are presented last in Section 7 as supporting information and serves as reference to consult for clarification on aspects of the analysis, and provides necessary information to allow for future reproducibility.

2 Background Theory

This chapter introduces the most relevant theory used in the current study. It first examines the concept of *personality* in evolutionary psychology, giving an account of the related field personality psychology from which it uses many ideas. It then introduces the reader to the Pareto Task Inference principle which is the fundamental research paradigm on which this study is based. Although a large number of computational tools from machine learning, and concepts from statistics are used in the analysis, this chapter does not treat them because, although important to understand for the purpose of usage, they are not a central part of the research problem. For reference on topics in machine learning used in this thesis the reader is referred to [12].

2.1 Evolutionary Psychology

This section is as much about personality psychology (PP) as it is about evolutionary psychology (EP), but that is only because EP is as much about any subfield of psychology as any subfield is about itself. EP boldly sets out to treat all parts of psychology from the fundamental assumption that all psychological processes result from Darwinian *natural selection* - a mechanism which over time pressures "unfit" traits out of existence. It views the brain as an evolved organ designed by natural selection to guide the individual in making decisions that increase the chance of survival and reproduction. EP was conceived in 1992 following the publication of *The Adapted Mind* by Leda Cosmides et. al., and is still considered an emerging field. Cosmides et. al. defines the discipline as "*simply psychology that is informed by the additional knowledge that evolutionary biology has to offer...*"^[13]. Despite having earned a substantial amount of controversy and criticism (which even has a dedicated Wikipedia article to it^[14]), EP has proven useful in explaining many concepts in e.g. language, cognition, intelligence, learning, adaptation, perception and personality^[2,15].

Since the author of this thesis has not been given the appropriate schooling, in the field of psychology, to lecture any reader beyond the mere introductory aspects of EP, each part of this section is presented as a query rather than a block of knowledge. The first sections treat topics in PP mainly dealing with personality and behavior, how personality can be measured and what inherent limitations remain in the existing methods. Following this, personality is presented as treated by EP and the view of personality as a behavioral strategy is given due motivation. The material relating to PP relies mostly on Icek Ajzen's textbook *Attitudes, Personality and Behavior* (2005)^[16], while the material on EP uses Lance Workman and Will Reader's textbook *Evolutionary Psychology* (2013)^[2].

Nature of response	Source of information about responses		
	Observation	Person	Acquaintances
Overt	Motor acts, nonverbal cues, verbal response	Self-reports of motor acts, nonverbal cues	Peer-reports of motor acts, nonverbal cues
Covert	Physiological responses, personal device monitoring	Self-reports of thoughts, feelings, needs, desires	Peer-reports of thoughts, feelings, needs, desires

Table 2.1: Responses used to infer personality. As the columns indicate, behavioral responses may be recorded by observing the individual from the outside, by personal questioning or by interviewing acquaintances of the individual. For each of these approaches, researcher can either aim at measuring overt or covert behavioral responses. This table is an excerpt from [Ajzen 2005]. The bolded text indicate information which has been added by this author.

2.1.1 What is personality and how does it relate to behavior?

In both psychology and everyday reasoning, human behavior is often explained by referencing underlying personality traits. When people cancel appointments they are called unreliable, when they make others laugh they are said to be charming, and when they are caught lying they are considered dishonest. As such, personality should be understood as that which you *are* while behavior is that which you *do*. More elaborate definitions characterize personality as “the set of emotional qualities, ways of behaving, etc., that makes a person different from other people”, and behavior as “the way a person or animal acts or behaves” or “the manner of conducting oneself”¹. Given these definitions it stands to reason that personality and behavior should be highly coupled constructs. Undoubtedly, the personality of an individual will affect its behavior, and in turn, its behavior will reveal information about its personality. This particular notion is of central concern to Icek Ajzen’s *Attitude, Personality and Behavior* (2005, 2nd Ed.)^[16]. The remainder of this discussion follows that discourse.

Ajzen argues that personality consists of latent, hypothetical characteristics which are not directly accessible, and can only be inferred from records of observable responses. Table 2.1, which is an excerpt from Ajzen’s book, establishes that response measurements can come from three sources: an observer, the subject itself, or someone who is acquainted with the subject. Responses can either be *overt* - meaning directly observable acts/cues/expressions - or *covert* - meaning not directly observed but inferred from interaction with the subject or from electronic measurements. It is important to note that responses need not be behavioral, but can also be answers

¹“personality”, “behavior”. Merriam Webster.com. Merriam-Webster, 2011. Sat. 7 May 2016.

to direct or indirect questions about the subject's personality, which is, for example, the case when inferring personality from questionnaires. In this study, measurements of behavior (for the sake of measuring behavior) are obtained through personal device monitoring (*observation/covert*) while measurements of personality are obtained through a questionnaire inventory (*person/covert*).

Trait theory is a field of inquiry that seeks to model personality as a set of traits (or characteristics) that each has influence on a range of behaviors. This, along with the earlier stated notion that personality and behavior are coupled, implies a degree of correlation between traits and behavioral manifestations, and so it is obvious to seek to construct predictive models between personality and behavior. In fact, a large parts of trait theory and an extensive amount of research, both historically and contemporary, addresses the modeling problem^[17–21]. Common for all, is that in no case has it been possible to produce statistically significant correlations between behavioral indicators and personality traits that exceeded 0.3 in absolute correlation. This limit has been observed so frequently that it has come to be known as the "personality coefficient" (coined by Walter Mischel (1969)^[22]). Regardless of its broad acceptance, however, it is important to note that the personality coefficient is an empirical value.

Ajzen attributes the low correlation to effects governed by behavioral inconsistency, that is, even though someone might consider themselves highly altruistic, they may not always act accordingly due to natural fluctuations in mood, circumstance, attitude, etc. Furthermore the notion of behavioral multi-determinism - that traits express themselves through multiple behaviors, and reversely that behavior results from multiple trait expressions - is argued to be a contributing factor to low predictive validity between personality and behavior^[16].

It is important to note that two variables with low correlation can still be statistically significant, and that often times this is the only thing which matters. In the case of this study, the purpose of correlating personality with behavior is not to build predictive models, but rather to study whether there is a significant relationship at all.

2.1.2 How is personality measured?

Human efforts to derive a taxonomy for character traits predate our calendar. Early efforts include the four *Cardinal Virtues* - prudence, temperance, courage and justice - due to Plato and Aristotle^[23], as well as the four humors - blood, yellow bile, black bile and phlegm - said to have originated in Ancient Egyptian medicine but formalized by Hippocrates^[24]. The foundation of the field now known as *trait theory* is, however, not considered rooted in this way thinking. Rather it started in 1884 when Sir Francis Galton established the *lexical hypothesis* which states that any trait that is truly important to humans will become encoded into their language^[1]. Galton reported finding

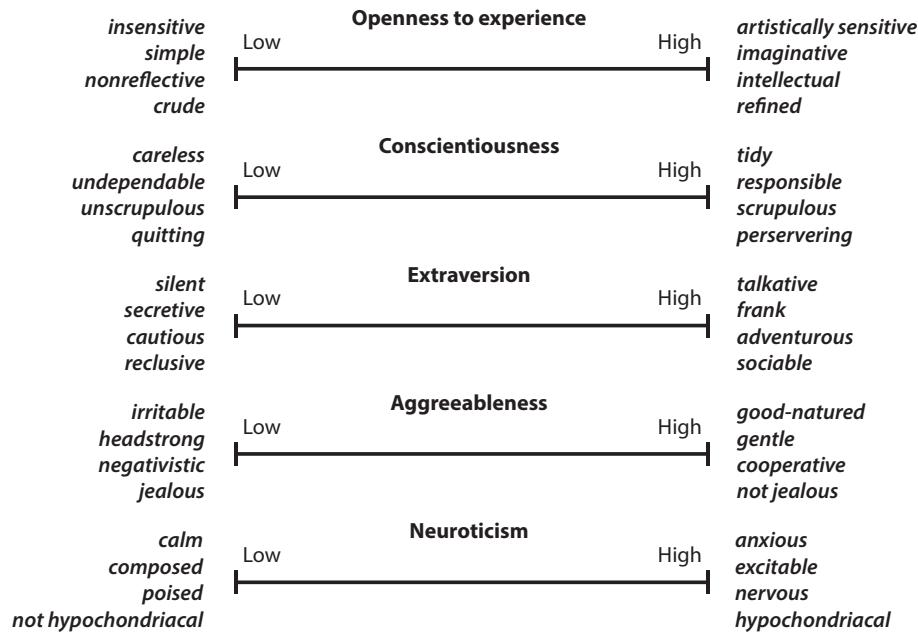


Figure 2.1: Visualization of the BFM. Traits are shown in the order 'OCEAN', each constituting a continuous domain between low trait value, associated with adjectives that oppose the name of the trait, and high trait value, associated with adjectives that support the trait name. The model explains an individual's personality as a vector with one score per trait. Scored are typically obtained through a questionnaire inventory.

exactly 1000 personality descriptive adjectives in the English language. A similar effort was later made by Gordon Allport and Henry Odbert (1936) using the *Webster's Unabridged Dictionary* (2nd Ed.), which resulted in a 17,953 long list of unique words to describe personality and behavior^[25]. The first effort to understand the underlying structure of this vocabulary was taken by Raymond B. Cattell (1943) who applied factor analysis to the covariance structure of words when used to describe individuals in large test groups, and found 16 primary and eight secondary traits^[26]. This result was refined by Donald. M. Fiske (1949) who, in an effort to replicate Cattell's result, arrived at a five factor solution^[27]. Supported by later research, this is still the prevailing result. It was coined the **Big Five model**² (BFM) by Lewis Goldberg (1981)^[28], but credit for its invention is typically given to Fiske^[29]. The model is shown in Figure 2.1. It offers five high-level personality traits: *openness to experience* (O), *conscientiousness* (C), *extraversion* (E), *agreeableness* (A) and *neuroticism* (N)^[30], which are typically recalled using acronyms OCEAN or CANOE. Traits in the BFM is typically measured using a questionnaire inventory such as the Big Five Inventory or the IPIP100 Inventory which relies on honest self-report by the subject^[31]. Rather than reducing the rich tapestry of the human psychology to mere five traits, the model seeks to provide

²Also commonly referred to as the Five Factor model (FFM).

a compelling framework for organizing the enormous variety of psychological differences characterizing humans. It assumes a hierarchical view of psychological traits, where the Big Five traits (BFT) are independent roots that each branch into many facets and subfacets.

Criticism The BFM is widely accepted among a substantial number of researchers, by some even considered a universally valid structure that transcends culture and language [32,33]. Unsurprisingly this particular viewpoint is not unchallenged. Recent findings have shown the BFM to only explain variance in a subset of traits (C, E, N) for indigenous societies [34], which resonates with the finding that *humility* emerges as a trait when studying populations in Asia (as captured by the HEXACO model, see below). Furthermore the lexical hypothesis has been argued to have the two following problems: (1) since language itself is not developed by experts, the inherent ambiguity of words causes any model based on language sampling to, at best, reflect only a lay perceptions of the traits [35,36], and (2) words can't accurately capture the spectrum of personality because some very important traits are tacit [37].

Limitations It has been shown that the BFM only captures up to 56% of a persons personality spectrum [38]. This is very important knowledge to have when using the BFM in any study that attempts to make general statements about personality. In the current undertaking, it too must be respected that certain fine-grained aspect of personality won't be captured and therefore shouldn't be speculated about. Unlike the Eysenck model (presented below) the BFM doesn't for example capture personality variance due to *psychotism*. It thus stands to reason that any strategies inferred from a dataset of BFTs cannot possibly be purely psychopathy, although it is in fact a proven evolutionarily stable strategy (ESS, see Section 2.1.3) [2].

Alternative models While the BFM is the most commonly used system for measuring personality traits, it is not the only one. The HEXACO model of personality structure is another widely accepted model which has deeper roots in European and Asian languages. Using factor analysis it finds six traits, where five are similar to those in the BFM, and one is related to honesty and humility [39]. Another important model is Cattell's 16PF questionnaire which measures personality on 16 out of his original 24 personality dimensions [40]. Finally, the Eysenck personality questionnaire reduces personality into mere three traits: *psychoticism*, *extraversion* and *neuroticism*. In this model psychoticism captures the aggressive and more extreme nuances of neuroticism. Rather than capturing a wide spectrum of personality the Eysenck model aims only at capturing personality variance associated with temperament.

2.1.3 How does evolutionary psychology treat personality?

Evolutionary psychologists are interested in personality because there are clues that suggest it might play an important role in natural selection. The first such clue is that much research shows personality to exhibit stability over time and contexts^[4, 41, 42]. The second is that studies have shown individual personality traits to be 30 – 50% inherited^[5, 6]. This strongly suggests that there are fitness benefits to different genetically inheritable aspects of personality. EP suggests that due the heredity of personality it may be seen, from an evolutionary point of view, as a **type of behavioral strategy**, i.e. a motivational system which predisposes people to seek out particular situations and respond in particular ways^[2]. This argumentation, however, gives rises to a paradox: in trying to explain the heredity of personality traits by assuming that the traits increase fitness, the question of why personality is *only* 30 – 50% genetic must be addressed. EP has several theories to explain this, of which the two central ones are: (1) the brain develops as it does due to *chance* because the genome only contains ~20,300 genes of information which cannot possibly code exactly for a brain that has ~100 billion complexly interacting neurons and (2) the individual is born with an adapted capacity for environmental calibration because inherent uncertainty of the world renders a fixed strategy on birth suboptimal^[2].

EP's primary instrument for measuring personality is the BFM because, as argued by Buss, it represents the most evolutionarily plausible way of carving up human personality^[3, 4]. The main reason for this is because the traits are orthogonal and they capture a wide array of individual differences. To compare, in one end of the scale Cattell's 16PF model elaborately cover many aspects of personality at the cost of using a very non-orthogonal system of traits, whereas in reverse the Eysenck model is too narrow and disregards evolutionarily important traits like *agreeableness*, *conscientiousness* and *openness to experience*.

Trade-offs cause trait variations Trade-offs between physical and psychological traits are inherent to almost all biological systems. It can be understood intuitively as the problem of having to choose between things that are good for something but can't be had simultaneously. Like being big or small: each has advantages, but obviously an organism must somehow choose to be either like one or a *trade-off*. An investigated example from Darwinian evolutionary theory explains the practical value of this problem. Ground finches on the Galapagos are typically sectioned into three types: (1) long beak/medium body, (2) large thick beak/large body and (3) small thick beak/small body. Shoval et. al. shows that each of these trait configurations are in fact archetypes of traits exhibited by all ground finches, and that each archetype is optimized to perform a specific task (see Figure 2.2.b)^[7]. Archetype-(1) is optimized for picking insects and extracting nectar, archetype-(2) is optimized for cracking large

hard seeds and archetype-(3) is optimized for crushing small soft seeds. Shoval et. al. observes that most ground finches are, however, not exactly like any archetype, rather each seems to *trade off* archetypal traits such as to be able to perform all tasks to some ability. This is thought to be a strategy for coping with changing ecological pressures, such that when a niche becomes unavailable (e.g. small soft seeds disappear) the organism has other ways of leveraging its fitness.

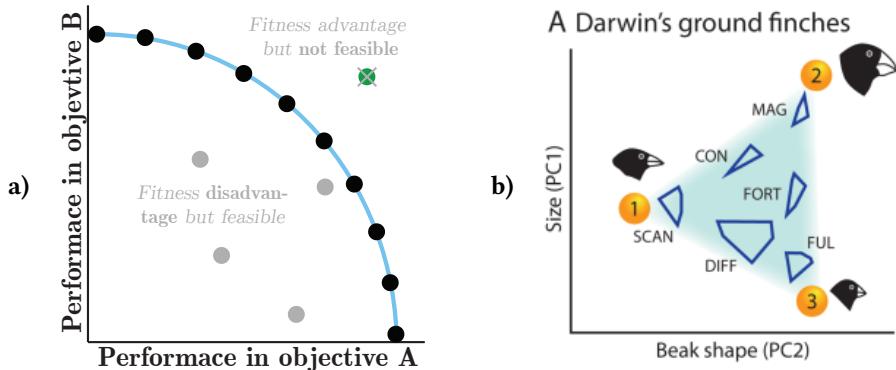


Figure 2.2: a) Illustration of distribution of points in a two-objective performance space where the black points are Pareto optimal and the blue curve signify the Pareto front where no objective can be optimized further without decreasing performance in the other. b) Illustration of how members of a species is distributed in trait-space to respect trade-offs between archetypal traits (source: Shoval et. al. [7]).

It is reasonable to ask what the best trade-off configuration of traits is, but the immediate answer is that there are infinite. To explain what this means, the concept of *Pareto optimality* must be introduced. Pareto optimality is a concept that stems from Economics and characterizes the set of solutions to a multi-objective optimization problem where no objective can be further optimized without having to compromise performance of another. For two objectives this means that solutions that are Pareto optimal are distributed along a curve in performance space, as illustrated in Figure 2.2.a. For an individual organism, traits will be adopted in such a way the the organism remain Pareto optimal, i.e. good enough at all evolutionarily important tasks to remain fit. Those that are not on the Pareto front will have a fitness disadvantage and be selected against in the long run, thus not passing on the genes that caused the disadvantage. The propensity to seek niches of least competition, i.e. *Niche filling*, drives populations to distribute traits somewhat evenly across the Pareto front.

The configuration of traits that a ground finch inherits from its parents can, in an evolutionary view, be considered a *strategy* for coping with the environment. This view is conceptually similar to the way evolutionary psychologists treat personality. For this reason, an individuals configuration of personality traits is also considered to be under influence of trade-offs. For example a person cannot (disregarding personality disorders) be both extroverted and introverted at the same time, and if life in

		Opponent	
		Hawk	Dove
You	Hawk	$(V - C)/2$	V
	Dove	0	$V/2$

Table 2.2: Reward table for the hawk-dove game. It models an exchange between 'you' and an opponent where each chooses e.g. to be like a dove (cooperate) or hawk (cheat). Each combination of moves has a different reward for each player.

general can be viewed as a complex collection of tasks then a persons level of intro-/extraversion poses an evolutionary trade-off problem just as much as beak shape does to seed-cracking.

Frequency dependence drives emergence of multiple strategies EP treats the predominance of different personalities/behavioral strategies using the concept of *frequency dependence* from evolutionary theory. It argues that the reason why all humans have not simply adapted the *single best strategy* is that, while at times there may be one strategy which is the best, this quickly changes when too many individuals adopt it. This can be compared to how driving to work early is a good strategy for avoiding traffic as long as not too many people do it. Another example is that of *psychopathy*. Workman et. al. argues that psychopathy is in fact (purely evolutionarily speaking) a very effective behavioral strategy to maximize personal fitness³, provided there are not too many psychopaths around^[2]. In most populations only 1-3% of individuals will adopt this strategy to a sub-clinical level whereas clinical psychopathy is as low as 0.2%^[43]. The psychopath strategy revolves around gaining peoples trust and then using them for own goals, and imaginably this strategy wouldn't work well if there were so many psychopaths that otherwise trusting people were used to getting cheated by psychopaths and therefore in general learned only to trust kin or long-term friends.

Game theory can explain some behavioral strategies The suggestion that personality is a behavioral strategy must be understood in terms of what evolutionary psychologists already identify with the term 'strategy'. It has its roots in the field of game theory, which examines how people behave depending on the strategies of others. Here the term is simply synonymous with *algorithm*. A central goal of applying game theory is to find a strategy for an individual which, given what everybody else is doing, cannot make the individual better off. Such a strategy is said to represent a state of *Nash Equilibrium*. A commonly treated paradigm is the so-called *hawk-dove game* which is illustrated in Table 2.2. It models an interaction between two players, *you* and an *opponent*, who have to share a resource V . Each player can choose

³Also: *exclusive fitness* - only increasing fitness of self as opposed to *inclusive fitness* which measures fitness of all kin.

to cooperate or defect, i.e. act as a *dove* or as a *hawk*. If both cooperates V is shared equally, but if one chooses to cooperate while the other doesn't the gullible player gets nothing while the defector gets all. In case both players choose the hawk-strategy the resources are shared equally but at the cost of battle, C . This game has been studied extensively, and for repeated games - between rational players - the winning strategy is a so-called *tit-for-tat* strategy^[2], where one simply copies the opponents' last move. This strategy constitutes a Nash equilibrium because, provided the opponent doesn't change strategy, there is nothing to gain from changing strategy^[44]. Other strategies such as a *generous-tit-for-tat* which chooses cooperation with some probability even if the opponent has defected, or *win-stay lose-shift* which (even simpler) just repeats past moves if they were successful and tries something new if not, have been shown to work slightly better when the opponent acts irrationally^[45]. Other less sophisticated strategies which are not in equilibrium are *uncompromising cooperation* as well as *constant defecting*.

A generalization of the Nash equilibrium to describe evolutionary stability which is used in biology, behavioral ecology and EP is the concept of the evolutionarily stable strategy (ESS). An ESS is a behavioral strategy that, if it exists abundantly enough in a population, it cannot be replaced by an invading strategy^[46]. It links to the concept of frequency dependence examined above, when multiple ESSs exist alongside, because they depend on each other in complex ways (recall the example with psychopathy), and can therefore be said to be *inter-frequency* dependent (this terminology is not used in literature but resonates with notions raised by Workman et. al.^[2] and Buss^[42]).

EP considers a number of behavioral strategies to be ESS. Firstly the psychopathy strategy is thought to be ESS. Another is *reciprocal altruism* which has been extensively studies in tribe cultures^[2], where individuals tend to give presents to non-kin with the expectation that presents will be returned to them at another time of greater need. Reciprocal altruism is formally treated as either: (1) *direct* where reciprocity is expected from the same person/group and of equal value as was donated (compared to the tit-for-tat strategy) or (2) *indirect* which relies on good reputation, established through gossip of altruistic acts, to eventually provide the giver with resources from like-minded (cooperation strategy)^[45]. Another important strategy which is thought to be ESS is the *freeloader*, which utilizes the cooperators trusting nature to cheat as frequently as possible within the limit of remaining an ESS. Evolutionary psychologists speculate the *freeloader* to have significant importance to community formation through exemplifying morally wrong behavior^[2].

2.2 Pareto Task Inference

Pareto Task Inference (ParTI) is a research paradigm developed at the Uri Alon lab to solve a number of problems related to evolution of morphology in various species. ParTI has been applied to understand systems in biology involving cancer cells^[8], morphology of ammonite shells^[9], mass and longevity design of mammals^[10], and more^[7,11]. It is founded in principles relating to evolutionary trade-offs and Pareto optimality as explained in Section 2.1.1, and combines these to provide a novel approach for inference of biological functions.

The goal of ParTI is to infer the *evolutionary objectives* of an organism. Objectives can be anything from *tasks* like cracking nuts and picking insects which a finch must perform well to remain evolutionarily fit, or even behavioral *strategies* that a human can employ to perform all of life's evolutionarily important tasks in a way that maximizes fitness in a given niche. **The value proposition of ParTI is that it provides evidence of what these objectives are only by considering the traits of the organism in question.** ParTI works in two parts:

- 2.2.1** Find archetypes, i.e combinations of traits that are optimized for specific objectives (*Pareto*).
- 2.2.2** Infer objectives by studying the combinations of traits for each archetype (*Task Inference*).

The central paradigm of ParTI exists in the first part, while the second part employs statistical methods to produce scientific results. In the following, each of these are examined.

2.2.1 Finding traits that are optimized for objectives

This part of ParTI is based on two core abstractions:

- 1. Inherent trade-offs between traits that are evolutionarily important to increase performance in different objectives, drive variation in a population to distribute along a Pareto front.
- 2. The Pareto front in trait-space is the hypervolume of a simplex. Each vertex of this simplex represents combinations of traits that are optimal for performing its respective objective.

From (2) it is noted that vertices of a simplex may also be treated as *archetypes*. Abstraction-(1) is explained in Section 2.1.3. Abstraction-(2) requires further explanation. When organisms have a single performance objective natural selection leads to traits that maximize performance of that objective. Any deviation from the optimal trait configuration causes performance, and thus fitness, to drop and in turn individuals who fall too far from the optimum will be selected against. This is illustrated in

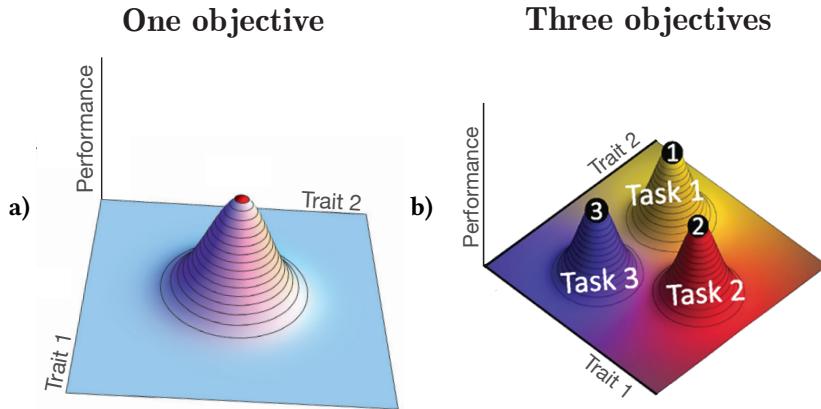


Figure 2.3: *a)* Performance function of a single objective/task in trait-space (source: [47]). *b)* Performance function of three objectives/tasks in trait-space. Archetypes are the local maxima (source: [10]).

Figure 2.3.a. When multiple objectives exist, no set of traits can optimize performance in all tasks and as such organisms must trade off traits relating to optimal performance of each objective. In Figure 2.3.b a performance function of three objectives is illustrated. It shows that in order to maximize performance, leaving the triangle spanned by the three peaks is disadvantageous. In fact, it corresponds to moving away from a position of Pareto optimality since performance in each objective is decreased without any gain. Therefore, the triangle shaped by the peaks, or *archetypes*, constitutes the region where traits are Pareto optimal, i.e. the Pareto front. Figure 2.4 shows how the number of objectives shapes the Pareto front. If there are two tasks, organisms must have traits that are distributed somewhat along a line between the archetypes in order to remain Pareto optimal, and if there are four tasks the Pareto front will be the volume spanned by the corresponding four archetypes shaping a tetrahedron. This is important because it shows if objectives are *evolutionary tasks* or *behavioral strategies*, these can be inferred just by considering the vertices of the triangle, i.e. the archetypes that emerges from the distribution of points in trait-space.

The ParTI principle generalizes to any number of dimensions Since a triangle is just a special case of a d -simplex composed of $d + 1$ vertices, the ParTI principle can be used to infer objectives in any number of dimensions. However, ParTI provides no "idiot-proof" solution to any dataset. For very high-dimensional data like gene expressions which may span a thousand dimensions, one shouldn't expect to find a thousand and one objectives. There is a purely mathematical reason for this. The convex set for any distribution of points grows exponentially with the number of dimensions as $\mathcal{O}(\log(n)^{M-1})$ [48]. Points on the convex hull are more likely to be

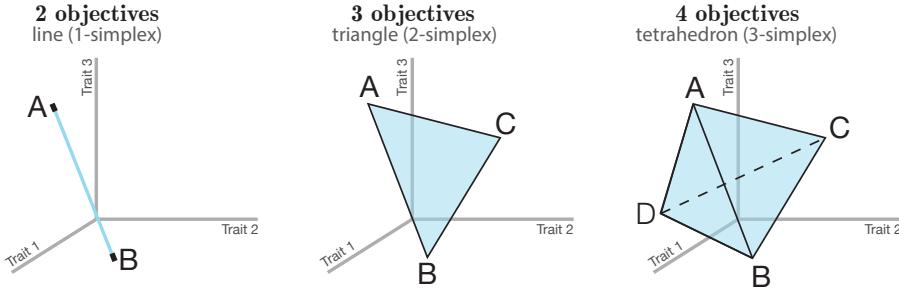


Figure 2.4: Pareto front regions (blue) for up to three objectives in trait-space. In the **2 objectives** case the Pareto front spans the edge connecting objectives A and B. For **3 objectives** and **4 objectives** Pareto fronts are the area and volume, respectively, of the simplex spanned by the set of objectives.

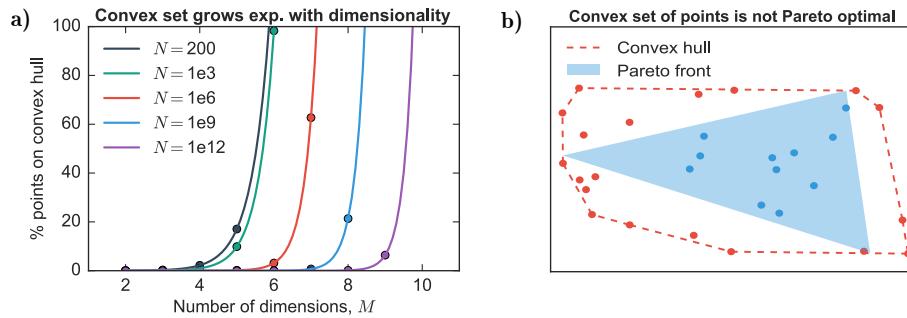


Figure 2.5: **a)** The size of the convex hull for N points scale exponentially with the number of dimensions M as $\mathcal{O}(\log n)^{M-1}$. Above five dimensions, many points are required in order to obtain a low ratio between the size of the convex set and the number of points. Note that the M -axis is discrete and that curves are plotted on a continuous scale for the sake of illustration. **b)** The Pareto front is a simplex anchored in $d+1$ points on the convex hull, and as such points on the hull that are not archetypes are not Pareto optimal.

outside of the Pareto front, and as such the possibility of finding valid simplex geometries (where *valid* means that most points are on the Pareto front) in high-dimensional spaces is severely diminished. Figure 2.5 illustrates that for over five dimensions, a very large number of datapoints are required to maintain a low ratio between the number of points in the convex set to the number of points in total.

Is this problematic? Somewhat yes, but recall from Section 2.1.3 the concept of ESS. In most biological systems multiple ESSs can only exist alongside if they are frequency dependent and *inter*-frequency dependent, and arguably there must be an upper limit to the number of complexly interacting strategies which can be ESS and exist alongside. That said, the above analysis cannot inform us about where this limit lies as biology has little regard for computational shortcomings.

Projections enable inference in high-dimensional data There is a solution to the problem pointed out above and illustrated in Figure 2.5. In high-dimensional

systems such as cancer cells which are typically measured in terms of their gene expressions, objectives can be found using dimensionality reduction methods. In ParTI literature the prevailing such method is principle component analysis (PCA)^[8]. PCA allows the analyst to project points in a high-dimensional space onto a subspace spanned an orthogonal basis of vectors which explain the most variance in the data. These vectors are called principle components (PC) and are computed as the eigenvectors of the covariance matrix. Informally, PCA can be thought of as casting a shadow of an object in such a way that the object can best be recognized. If the object was a sword, the first PC would follow the length of the sword and the second (being orthogonal to the first) would follow the guard and as such the best way to show that it was a sword just from looking at the shadow would be to project it onto these two components.

For objects that exist in high-dimensional spaces such as a point-cloud of gene expressions this method of "casting the shadow of highest explained variance" works methodically in exactly the same way as the sword example. Here, however, the analyst has more freedom to choose the dimensionality of the space projected to (the shadow). When ParTI treats high-dimensional data it looks for shadows that are shaped like simplexes (see Figure 2.4). For a low-dimensional projection where the points appear to fall inside a simplex, the vertices can be taken as the combinations of traits that allow for optimal performance of objectives, and each can in turn be identified by further analysis into how organisms with that combination of traits behave.

2.2.2 Inferring objectives from enrichment near archetypes

In the current context the word *inference* simply means to uncover evidence of certain features or behaviors of an archetype to the point where it is valid to make an interpretation of what the main objective of the archetype is. As such, ParTI is really a method for *providing evidence* rather than giving an ultimate answer.

Since archetypes in trait-space are optimal configurations of traits for performing specific objectives that are evolutionarily important, the problem of inferring what those objectives are can be solved in many ways. For very simple systems one could unobtrusively observe how organisms employ their traits and take special notice of behavioral differences between those that resemble archetypes^[7]. But for systems that can't be directly observed, other methods must be used. A relevant example is personalities of people and which behavioral strategies for performing evolutionarily important tasks they lead to. In this example traits are personality factors such as the Big Five (see Section 2.1.2) and the objective is the behavioral strategy that the personality is optimized to perform. Given the archetypes it would be to time-consuming and obtrusive to monitor people, who were similar to each archetype, in their natural habitat. Using qualitative interviewing might work better but would also be time

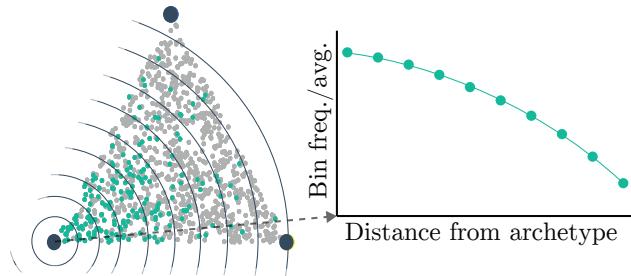


Figure 2.6: Enrichment of variables near archetypes are estimated across a number bins characterized by distance to the archetype. If the studied feature is continuous the

consuming and possibly introduce qualitative research biases.

A better approach is to study the statistical properties of individuals who are similar to each archetype. Using *extra data* such as demographic variables, response items from surveys and behavioral data, layers of information can be added such as to reveal what people similar to each archetype typically do for a living, whether they tend to be married or what age range they occupy. This approach is general. If the system were cancer cells and traits were gene expressions, extra data such as tissue type, cell cycle period and protein production level would provide evidence to aid inference of objectives for different cells [8].

There are different methods for doing this. The simplest method is to correlate the feature value with distance from the archetype and then use the slope and correlation coefficient as measures of relatedness. Another approach which provides a better visual image is to analyze the *enrichment* of features near the archetypes. As illustrated in Figure 2.6 this method works by binning points according to their distance from the archetype and then taking the average of the feature in each bin to create a curve which shows whether the feature tends to be enriched or *depleted* near the archetype. If the feature is categorical the method estimates the frequency of each category in every bin, such as to create separate histograms for the feature-categories. This method is applied widely in ParTI literature [7–11].

2.2.3 Inference of archetypes

So far *archetype* has been used synonymously with *simplex vertex*, and surely simplex vertices can be considered archetypes - but archetypes are not necessarily simplex vertices. Figure 2.7 gives two examples of archetypes in clusters of simulated points, and shows that archetypes are in fact just the intuitive "corner-points" of any point cloud. The method for inferring these corner-points is called archetypal analysis (AA). It was introduced by Adele Cutler and Leo Breiman in 1994 and developed as an unsupervised learning method for clustering data that is not well-separated [49]. The central idea it to derive a set of archetypes from which all points can be approximated

as linear combinations. Below is given a brief introduction to the method.

The AA factorization problem AA solves the following matrix factorization problem. Consider a data matrix $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \in \mathbb{R}^{m \times n}$. The task of AA is to estimate the two factor matrices $\mathbf{S} \in \mathbb{R}^{k \times n}$ and $\mathbf{C} \in \mathbb{R}^{n \times k}$ such as to satisfy the following equation with minimal approximation:

$$\mathbf{X} \approx \mathbf{AS} = \mathbf{XCS} \quad (2.1)$$

The number of archetypes is controlled by k and can be chosen by the analyst, while the archetypes themselves are column vectors in $\mathbf{A} \in \mathbb{R}^{m \times k}$. \mathbf{C} and \mathbf{S} are column stochastic matrices subject to $|\mathbf{c}_j|_1 = |\mathbf{s}_i|_1 = 1$. Columns in \mathbf{C} give the coefficients of each archetype as a linear combination of the data points, and reversely \mathbf{S} give the coefficients of each data point as an approximate linear combination of the archetypes. This leads to the property that archetypes may be derived from data points as $\mathbf{A} = \mathbf{XC}$ and data points approximated from archetypes as $\mathbf{X} \approx \mathbf{AS}$. This introduces symmetry into the matrix factorization problem consistent with the following statement: *archetypes are convex combinations of the data points and data points are approximated by convex combinations of archetypes*. \mathbf{C} and \mathbf{S} are derived by minimizing the residual sum of squares (RSS):

$$RSS = \|\mathbf{X} - \mathbf{XCS}\|_F^2 \quad (2.2)$$

subject to $|\mathbf{c}_j|_1 = 1 \forall j \in \{0, \dots, k\}$ and $|\mathbf{s}_i|_1 = 1 \forall i \in \{0, \dots, n\}$, where $\|\cdot\|_F$ is the Frobenius norm. Furthermore, AA imposes the constraint that all coefficients of \mathbf{C} and \mathbf{S} are greater or equal to zero. Note that RSS goes to zero in the limit $\mathbf{AS} \rightarrow \mathbf{X}$ where the data points can be reconstructed exactly as convex combinations of the archetypes. It can be shown that Equation (2.2) is non-convex, and that the optimization problem has no closed form solution [49], yielding numerical methods necessary for solving the problem.

Numerical methods for estimation of archetypes Cutler et. al. solved the AA problem using an alternating optimization approach. Starting from a random initialization of \mathbf{S} the method alternates between finding the best \mathbf{C} for an \mathbf{S} and finding the best \mathbf{S} for a \mathbf{C} , for as necessary until the reduction in RSS is sufficiently small. Each iteration requires solving several convex least squares problems and convergence is slow for many points in high dimensions because the number of points on the convex hull increases exponentially (see Figure 2.5.a). Recent work due to Mørup and Hansen introduces an effective initialization procedure and a projected gradient approach to solving the AA matrix factorization problem, which however reduces the computation time by limiting the number of points on the convex hull that can con-



Figure 2.7: Two generic examples of archetypes in multidimensional data. Hollow circles represent the archetypes which are typically located in corner-like regions of the data.

stitute solutions to the optimization problem^[48]. It furthermore allows the analyst to compute archetypes with slack such that they may fall outside of the convex hull, diminishing the problem illustrated in Figure 2.5.b.

Available implementations A number of implementations of AA are available across different software languages. Morup et. al. gives an implementation of their proposed projected gradient version of AA^[48] which has furthermore been translated into a Python package, by this author, available on the Python Package Index (PyPI). The Python package PyMF also gives an implementation of the, however, slightly slower alternating optimization version of AA^[50]. Finally, there is an implementation in R that also uses the traditional approach^[51].

Alternative methods As far as the PartI principle is concerned the only requirements of archetypes is that they constitute the geometrical corners of the dataset. As such other methods for computing archetypes are just as valid as AA. Four alternative methods estimate archetypes using minimal volume simplex and spectral unmixing methods using different approaches^[52,52–54]. It is, however, not within the scope of this thesis to further discuss any of these methods.

3 Datasets

Four different datasources are used in this study. Each one contains multiple datasets and are used for slightly different purposes. Table 3.1 gives an overview of which datatypes each source provides. They all have in common that they provide Big Five data for a large number of survey responders which is useful for the analysis presented in Section 4.1. Appendix G gives a comparative overview of all Big Five datasets.

	Big Five	Facets	Values	Attributes	Behavior	N (appr.)
SensibleDTU	✓	-	-	-	✓	1E3
myPers.-project	✓	-	✓	-		2.5E6
SAPA-project	✓✓✓	✓		✓		3 – 78E3
MIDUS	✓✓✓	-	-	-		3 – 7E3

Table 3.1: Overview of the types of data contained in each used dataset. '✓' symbolizes that the source contains a given type, and if there are multiple it means that the source contains multiple sets of the type. '-' symbolizes that the type is contained in the datasource but is not used in the analysis.

3.1 Sensible DTU

Sensible DTU was an experiment conducted as part of the ongoing interdisciplinary research project *Social Fabric* at Copenhagen University (UCPH) and DTU Compute of The Technical University of Denmark (DTU), which ran from September 2013 to January 2015 (2.5 years). In the experiment, 863 undergraduate students at DTU gave informed consent to participate. Participants were predominantly male (78.8%), of average age 21.7 (SD 2.8) years, distributed across all study lines, and were primarily 1st year students (59.2%) and 2nd year students (30.2%). Remaining students were either 3rd year students or students for which a year was not reported. Participants were provided a NEXUS smartphone for personal use, which was reprogrammed to continuously collect social behavioral data through multiple channels^[55,56]. Tab. 3.2 gives an overview of which social behavioral data were collected.

The experiment saw its largest volume of activity in the spring of 2014, where 526 students were actively using the provided smartphone as their primary device (active on more than 75% of days through the period, see Fig. 3.1).

As described in Sec. 2.1.1, behavioral inconsistency sets a limit for the level of mutual information, or correlation, there may be found between a personality trait obtained through a questionnaire, and a specific type of behavior. To mitigate such influence, data originating only from school weeks in the spring semester of 2014 is selected for analysis, based on the assumption that this provides the highest degree of self-similarity in the data. Similar subsets, e.g. from exam periods or holiday weeks, could also have been chosen, but would contain less data since these are narrower spans of time.

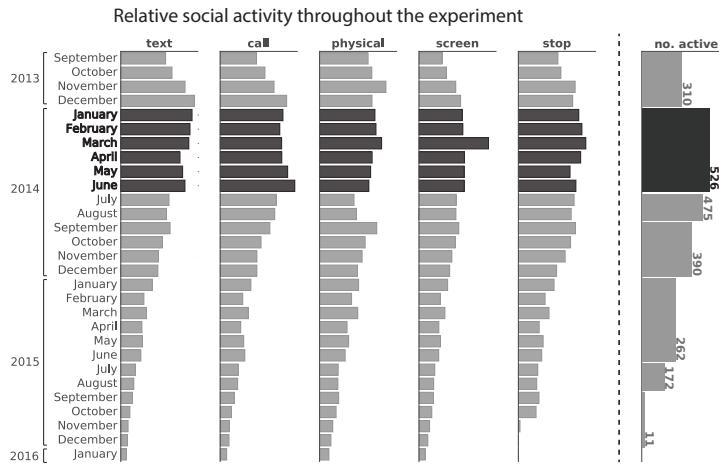


Figure 3.1: An overview of the amount of activity tracked in the respective months of the experiment. The months marked in bold represent those with which the data used in this study is associated. The far right bar plot show the number of students in each semester and summer holiday, which are active on more than 75% of days through the period.

Before receiving the phone, participants were required to fill out a baseline questionnaire. It included the Big Five (BF) questionnaire which was used in this study to obtain BF profiles for each subject (see Sec. 2.1.2). The BF questionnaire was filled out for each participant no later than December 20th of 2013 [55]. The notion raised in Section 2.1.1 that behavior and personality are not constants provides a compelling argument that they should be measured as close in time as possible, which gives another reason for choosing to use only behavioral data from the spring semester of 2014.

Researchers in the Social Fabric project also has access to administrative data such as birth year, courses, drop-out and most notably grades, none of which, however, were

Datatype	Sample period	Used in this study
GPS-location	15 minutes	Yes
WiFi-scans	10 minutes	No
Bluetooth-scans	5 minutes	Yes
Phone calls	-	Yes
Text messages	-	Yes
Facebook data	-	No
Screen on/off events	-	Yes

Table 3.2: Overview of the datatypes which were collected through the administered smartphones, and which are used in this study. Neither channel provides written or spoken content. The Facebook channel includes friends, groups, interests, likes, political orientation, religion and more.

used in this study.

Privacy and data access Given the sensitive information contained, or easily extracted, from this data despite all reference to personal identity having been removed, strict access rules to the data are imposed. Researchers are obliged to sign a non-disclosure agreement (NDA) and raw data is not to be contained on private or public machines outside of the EU. Data is accessed in JSON format through the Data Viewer API at `sensible.dtu.dk/apps/data_viewer` with a valid research token, or through a virtual IPython Notebook environment hosted at `raw.sensible.dtu.dk`, where data is served as Numpy MaskedArray type for efficiency which easily converts to Pandas DataFrame type, Numpy 2D array or similar matrix-like types (see Sec. 7.1 for technical reference). Researchers working outside of the EU are restricted to work with the data remotely through the virtual IPython Notebook environment, which is, conveniently, also the faster alternative.

3.2 myPersonality

The myPersonality research project maintained a Facebook application through which users could answer various personality questionnaires and donate personal data from their profiles^[57]. It reached great popularity because it allowed users to compare results with friends and even fill out personality questionnaires for others. From 2007 to 2012 the research project managed to collect questionnaire responses from nearly 7.5 million Facebook users. 900.000 users retook the questionnaire providing longitudinal data, 300.000 provided friend ratings enabling self-report accuracy assessment and 40% gave full access to private profile information.

In the current study, access to approximately 2.5 million BF profiles and 9000 'Basic Human Values' (BHV) questionnaire responses were provided. BHVs stem from the Theory of Basic Human Values due to Shalom Schwartz^[58,59], which juxtapose the 10 fundamental values: *benevolence, tradition, conformity, security, power, achievement, hedonism, stimulation, self-direction, universalism*. An overview of value/facet-relationships of their position relative to each other is given in Appendix F.

3.3 Midlife Development in the U.S

Midlife Development in the U.S (MIDUS) was a longitudinal telephone/email survey study conducted in three waves in 1995, 2004/06 and 2013^[60]. Its goal was to create a dataset which could be analyzed to understand how various life variables affect each other. It successfully collected repeated questionnaire responses for 3294 randomly selected Americans, which, among many other metrics, included personality profiles measured with the BFM. For the BF items in the survey participants rated how well different words described them on a four tick scale. In this study, BF datasets from

each of the waves are used. MIDUS 1 has 6261 valid datapoints, MIDUS 2 has 3971 and MIDUS 3 has 2715. Importantly the MIDUS 2 and MIDUS 3 datapoints are measurements of participants who are also measured on in the previous years. For the intended purpose this is not thought to have any damaging influence on the results, but is still important to keep in mind. Considering Appendix G is can furthermore be seen that the C, E, and A traits do not distribute normally, but seem highly shifted towards higher values. This may be due to bias in the collection method which relied on an interviewer to collect the responses over telephone, or possible an artifact of the inventory used.

3.4 SAPA-Project

SAPA stands for 'Synthetic Aperture Personality Assessment' which is a survey collection method that feeds the respondent with semi-random questions from a large pool of questionnaire items from a variety of inventories, until the respondent stops answering. This creates a sparse dataset of response values, but for a large number of participants, covariance matrices between questionnaire items across people with different attributes (e.g. country/gender) can be computed to infer differences between groups. The SAPA-Project is an effort by computational psychologist David Condon, which works as a web application allowing Internet users to take a personality test and get insights about themselves^[61]. In turn it is also a tools for researchers to gather personality data. Participants disclose various personal demographic information and respond to questions from three BF inventories: IPIP100, IPIP-NEO and SPI. The first two are internationally recognized pools of questionnaire items^[62], while the last one is a product if Condon's thesis^[61]. Each one derives the BFTs using slightly different approached but largely produce the same factors. Using only responders who answered at least four items for each trait, the IPIP100 inventory has 77,685 valid datapoints, the SPI inventory has 2928 valid datapoints and the IPIP-NEO has 37,554 valid datapoints. 100% of valid responders to the SPI inventory are also valid responders to the IPIP-NEO inventory and 47% of valid responders to the IPIP-NEO inventory are also valid responders to the IPIP100 inventory.

An excerpt of the SAPA data is freely available through the Harvard Dataverse at <https://goo.gl/vrX83G>.

4 Analysis and Results

The research objective is to find fundamental personality archetypes and understand which behavioral strategies they correspond to in terms of their personality facets, demographic attributes, values and correlated behaviors. As stated in the introduction this thesis documents the first efforts in an ongoing research project, and as such the above goal is bold and will not be exhaustively satisfied. The results do however offer clues towards *which* evolutionary behavioral strategies can be inferred from data.

The analysis takes the following trajectory: Six archetypes are found to emerge consistently across multiple datasets. For each, values, personality facets, personal attributes and behavior is inferred. The features of each archetype differ in fundamental ways and provide clues towards strategies for performing life's many tasks. These are speculated to be: (1) *self-concern* (power/achievement driven, emotionally stable, unsympathetic), (2) *social achievement* (achievement driven, highly social, emotionally unstable), (3) *reciprocal altruism* (conformist, benevolent, social, emotionally stable), (4) *xenophobia-loyalism* (narrow-minded, loyal, orderly), (5) *freeriding* (unconformed, undutiful, unorganized, unworried, emotionally stable) and (6) *withdrawal* (antisocial, depressed, worried, emotionally unstable). Note that for the sake of clarity these strategy labels are used throughout the analysis before they are given fair justification.

The analysis is separated into parts that each has a single premise. Each motivates the next and together they discover the strategies. The first part presents the analysis leading to the discovery of *consensus archetypes*, a relaxed interpretation of an archetype. The second part investigates the types of personal attributes that are enriched near each consensus archetype, and which stable behavioral indicators correlate with distance from the archetypes. The third part investigates the predictive power of the consensus archetypes compared to raw BFTs and PCs of BF questionnaire items (QI). Technical details about the analysis are explained in Section 7.2.

4.1 Six consensus archetypes emerge across seven datasets

The search for six strategies has two motivations. The first is that, as was shown in Section 2.2.1 Figure 2.5, for more than five dimensions most points are in the convex set of points and therefore not inside of the Pareto front, unless the number of points becomes very large. As such, this justifies six as an upper bound. The lower bound is slightly harder to justify, but an argument can be found in the fact that there are five personality factors. In Section 2.1.2 the Big Five model was presented and it was explained that the covariance of personality descriptive adjectives applied to subjects within a large enough population across a multitude of studies leads to five significant factors of personality. The ParTI principle states that objectives (here: strategies) are the $d + 1$ vertices of the simplex on which points fall in the d -dimensional subspace of greatest variance, and since personality varies strongly along five factors, there

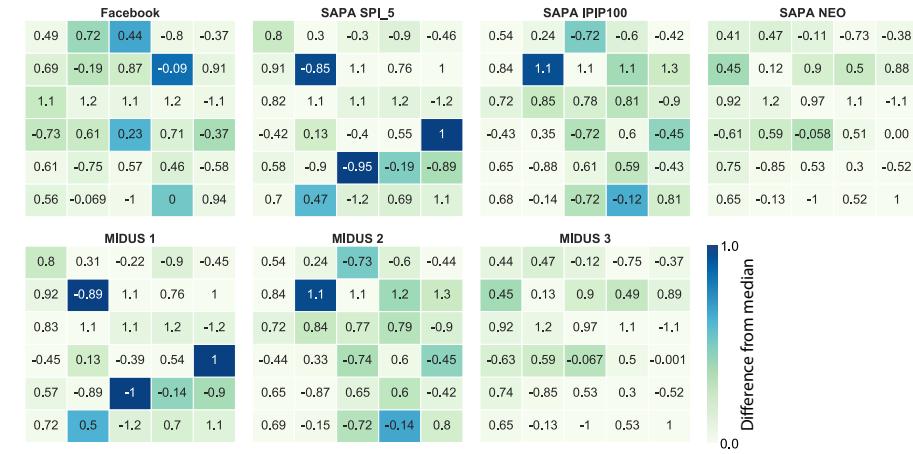


Figure 4.1: Archetypes emerging from each dataset. Color map scales with archetype trait distances from archetype trait median across all datasets. Annotated values are rows of archetypes and BFTs as columns ordered by O, C, E, A and N.

should not be less than six fundamental personality strategies.

Archetypes are computed for multiple datasets to ensure generality. Under the assumption that no questionnaire inventory and sample population can provide enough generality to find the fundamental personality archetypes, the ParTI principle is applied to seven different datasets (see Section 3). The Sensible DTU BF dataset is left out because the population is deemed to be specified to produce general enough archetypes. For each dataset archetypes are computed as median values over 5000 bootstrapping iterations with sample size 200. The emerging sets of archetypes (median values) are presented in Figure 4.1. Section 7.2.1 explains in detail how the archetypes are computed.

Archetype deviations across datasets reveal consensus archetypes. Figure 4.1 shows that roughly the same archetypes emerge from the different datasets. However, certain archetype traits are far from the median. Since these are large datasets from mixed populations it is not believed that any one produces archetypes that are severely different due to noise. Rather the differences are believed more likely to arise due to characteristics relating to the population and the questionnaire inventory. To investigate the severity of the differences a 6×5 array of SDs across the datasets (stacked in depth and calculated along 2nd axis, see Section 7.2.1) is computed, and shown in Figure 4.2.a. Due to the observation that some archetype traits appear to have much higher SD than others a *consensus threshold* of maximum SD, λ , is set for a significance level $\alpha = 0.05$. To find the threshold, an SD value array is computed multiple times for shuffled archetype traits (columns shuffled independently). All resulting SDs are then arranged in a list and sorted to ascending order, and the largest value of the lower 5th percentile is chosen as the threshold. This method results in

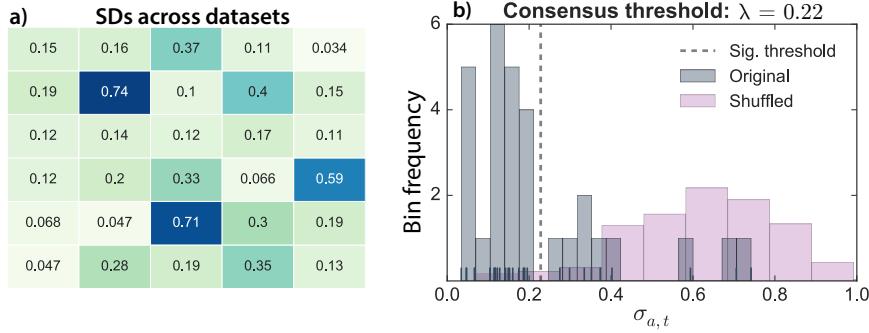


Figure 4.2: **a)** SD values for archetype traits across datasets. It is observed that some vary greatly while others are very stable. **b)** A consensus threshold λ emerges at 0.22, which gives rise to p-values less than 0.05 for all archetype traits below the consensus threshold. The purple histogram shows the normalized distribution of SDs for shuffled archetype traits. The blue histogram shows the normalized distribution of SDs of the actual archetypes, where the rug shows distribution of individual SDs.

$\lambda = 0.22$. The distribution of SDs and shuffled archetype SDs are visualized in Figure 4.2.b. Clearly there is a group of archetype traits with low SD and one with high SD. λ gives rise to a mask that decides which archetype traits are in consensus and which are not. Applying the mask to the median archetypes across all datasets reveals the *consensus archetypes* (CA) presented in Figure 4.3.a. For consistency the CAs are here named according to their later speculated strategies.

The strategies reveal themselves somewhat from the CA traits. The *self-concern* CA is particularly unagreeable, which typically correlates with distrust for others and self-centered motives. The *withdrawal* CA is highly introverted and neurotic which could indicate a *constantly-aware-of-danger* strategy. In fact, it seems, each archetype is quite distinct from the others. It is definitely worth taking a short break from reading to get acquainted with the CAs and their individual differences and similarities (see Figure 4.3.a).

A remark on the mathematical implications of CAs should be made. While there are no strict mathematical definitions of what an archetype is, the discourse so far has implied that it is a *point*. However, once a consensus mask is applied, certain of the trait values for some of the archetypes are said to be "not in consensus". What this really means is that when e.g. computing the distance to a CA the weighted euclidean distance must be used, adding zero weight to non-consensus traits while all other weights sum to 1. As such, if, for an archetype, a single trait is in non-consensus the archetype is effectively a line. If two traits are in non-consensus it is a plane, if three a volume, if four a hyperplane in five dimensions, etc. In the following analysis the fact that they are CAs and not just normal archetypes only has importance in terms of computing the distance to them, which is done using the weighted euclidean ap-

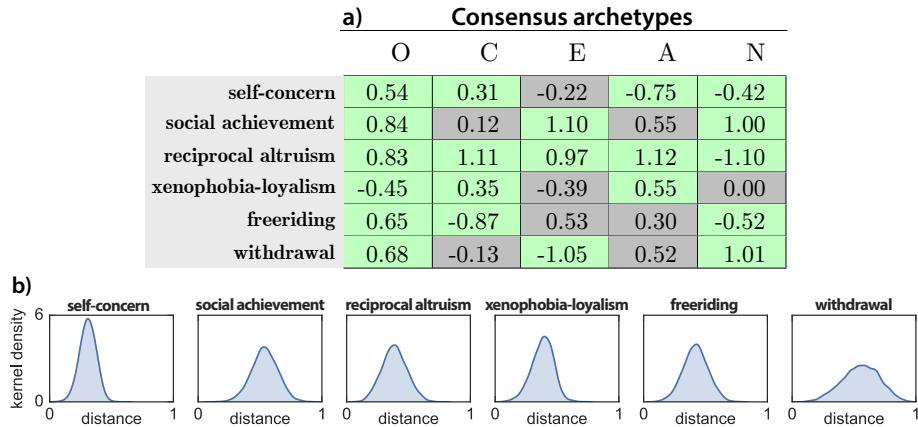


Figure 4.3: a) **Main result of Section 4.1.** Consensus archetypes emerge as median archetypes where trait values are colored with respect to the consensus threshold: green signifies consensus and gray signifies above threshold non-consensus, i.e. undefined dimensions. b) Distributions of weighted euclidean distances from points in the Facebook dataset to consensus archetypes. Axes are scaled equally.

proach as described. Therefore CAs will be referred to as *archetypes* unless discussion requires them to be explicitly stated as archetypes in consensus between datasets.

4.2 Enrichment and correlation reveal distinct strategies

From Figure 4.3.a it is observed that some archetypes share very similar trait values. Comparing e.g. the *social achievement* and *withdrawal* archetypes they only differ on the E trait. Does this mean that the two may correspond to the same behavioral strategy? No. In the following the differences between archetypes are investigated by studying how facets, values, demographic attributes and behaviors are enriched/depleted⁴ for- or correlate with distance from the archetypes. Enrichment analysis is used because it provides an intuitive way of interpreting whether a variable is influenced by distance from an archetype, and correlation analysis is used for behavior where there are not enough datapoints to provide statistically significant results using enrichment analysis (see Section 2.2.2). It is important to stress, however, that either of these approaches inform only about statistical *trends* and not absolute differences. For example, as was stated in Section 2.1.1 there is a limit to how well personality and behavior can correlate, empirically fixed around 0.3. And while the same rule may not necessarily apply for demographic attributes, values and particularly not facets (for reasons explained later), it still raises the important notion that no relationship between *anything* and personality is absolute, because after all (as was raised as criticism in Section 2.1.1) the best available method can only explain 56% of personality variance. The results presented in the following are therefore only *indicative* and their

⁴Herinafter just *enriched* to simplify the language.

implications are strictly useful for supporting claims that strategies exist and suggesting what those might then be. Since the analysis yields many results, this section presents them in summary form. The full collection of results is presented in Appendices C-E.

4.2.1 Analysis and data

Enrichment methods for continuous and categorical variables, as described in Section 2.2.2, and Pearson product-moment correlation are applied. Table 4.1 gives a qualitative summary of all enrichments as they are presented in Appendices C-E. Figure 4.4 visualizes the strongest enrichments of QIs and BHVs, as well as enrichment of professional disciplines for each archetype, to give the reader an intuition of the method. Because both enrichment and correlation methods uses *distance* from archetype as the independent variable, negative slopes signify enrichment and positive slope signifies depletion. Error bars have been omitted because they vanish due to the large sample sized, however SDs are typical in the order of a quarter of the second axis range which is not insignificant to the later interpretation of the enrichments. All presented enrichments and correlations in Table 4.1 and Figure 4.4 are significant after Bonferroni and Benjamin-Hochberg corrections, respectively, for a significance level of 0.05.

Two datasets provide facets, values and attributes and one provides behavior. The SAPA-project dataset is analyzed for facets and demographic attributes. It comprises personality test responses from a large number of Internet users, as well as demographic information such as age, gender, country, education and marital status. This provides a unique opportunity to study how personality - or in this paradigm, distance to an archetype - links to demographic attributes. Furthermore the property that the SAPA collection method asks respondents questions from multiple inventories is taken advantage of. When analyzing enrichment of QIs care has been taken to avoid "circular enrichment" since distance is measured in units of BFTs and BFTs are composed of QIs. By analyzing enrichment of QIs from the SPI inventory for individuals close to each archetype, where distance is measured using BFT values computed from the IPIP100 inventory, circularity is mitigated.

The *myPersonality* Facebook dataset is analyzed for values. The SensibleDTU dataset is analyzed for behaviors. Raw data is recombined into 38 behavioral indicators for 412 study participants, using an - for the purpose of this study - extended version of the behavioral feature extraction Python software *bandicoot*. The indicators can be viewed in Appendix A, and the implementation pipeline is presented in Section 7.1.1.

Facets	Demographic attributes						Mobile behavior		
	Values	Highlights	Discipline	Rel.ship	Job status	Exerc.	Smok.	Age	Highlights
self-concern	Emotionally stable, unsympathetic, self-concerned, exploitative, arrogant, hard to offend	Power, not benevolence, not conformity, not tradition, achievement, stimulation, hedonism	Computer sci., language and lit., arts, natural sci. eng. and communication, edu., med. and health, social services	No	Unemployed	Yes	-	Younger (0.5 years/bin), not retired	-
social achievement	Talkative, social, lively, moody, worrying, emotional, insecure/need reassurance, easily hurt	Stimulation, hedonism, power, achievement, universalism, security, not conformity	Insurance sales, operations manager/specialist, CEO - not food preparation worker, cashier, computer programmer	Yes	Employed	Yes	Yes	Not retired	Freq. calls/texts, brief stops, goes new places, long calls, sends more texts than receives, response-rate depends on receiver, socializes at night
reciprocal altruism	Happy, hard-working, dutiful, well-tempered, trusting, talkative, social, lively, unworried	Conformity, benevolence, stimulation, achievement, tradition, security, universalism	Business, education, community and social services - not arts, lang. and lit.	Yes	Employed, not student	Yes	No	Older (-1.01 years/bin)	Socializes off-campus, freq. calls/texts, goes out at night, unperiodic social pattern, brief stops, gets final word
xenophobia-loyalism	Uncreative, simple-minded, ignorant, unintellectual, slow	Not self-direction, conformity, not stimulation, tradition, not universalism, security, not hedonism	Healthcare, office support, military - not computer sci. and arts	Yes	Student, unemployed	-	No	Younger (0.3 years/bin)	Periodic social pattern, returns calls/texts, immediate text responses, has long ongoing chats, shows up for class
freeriding	Undutiful, not hard-working, disorderly, rule-breaker, don't plan ahead, calm, emotionally stable	Not security, not achievement, not tradition, not power, not benevolence, stimulation, hedonism	Arts, cultural and regional studies, lang. and lit. studies, communication, social sci. - not business and med./health	No	Student, unemployed	No	Yes	Younger (0.6 years/bin), not retired	Short calls, unperiodic social pattern, very nocturnal, interacts with strangers, none/slow text/call responses, doesn't conclude chats
withdrawal	Not talkative, shy, antisocial, reclusive, depressed, worried, moody, irritated, panicking, emotional, lost	Not stimulation, not achievement, not conformity, not power, not security, not benevolence, not hedonism, not tradition	Arts, lang. and lit., computer sci., natural sci. - not business, health-care and communication	No	Unemployed	No	No	Retired	Doesn't meet people at campus, not nocturnal, periodic social pattern, rarely calls/texts/looks at phone, few contacts, stays home

Table 4.1: Summary table of the strongest significant enrichments for facets, Basic Human Values, demographic attributes and mobile behaviors for each archetype. Considering the rows one by one there are clear differences between the archetypes. Facets, values and behaviors are sorted by descending enrichment/correlation strength. The following words are shortened: *sci[ence], techn[ology], engineering, literature, lang[uage], education]*, *med[icine]*. '-' symbolizes no significant enrichments. Facets are single-word descriptions of longer questionnaire items (see Appendix C). Demographic attributes strongly influenced by age, such as education level and marital status, are omitted due to the strong age differences between archetypes.

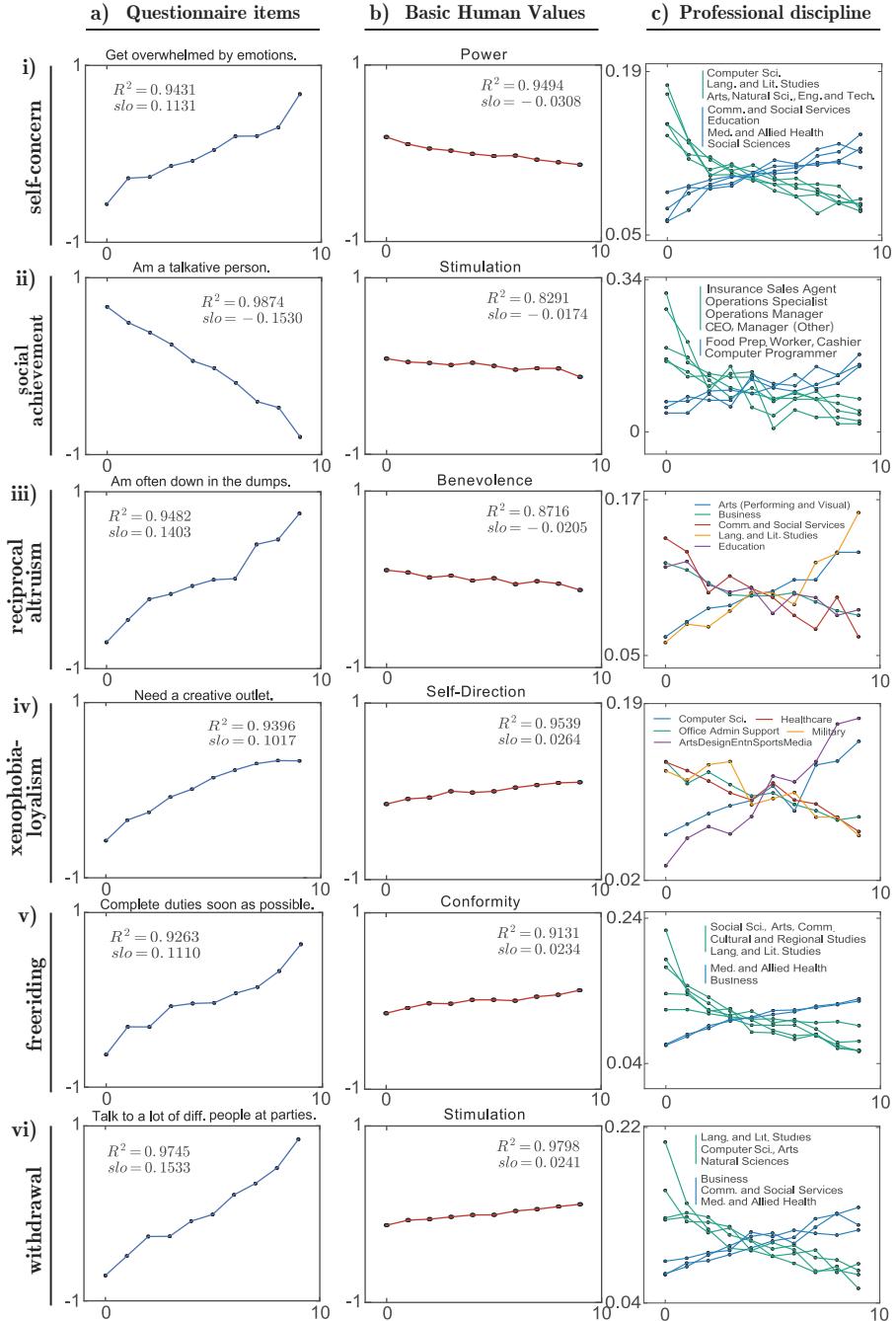


Figure 4.4: Enrichments for each archetype. **a-b)** are single strongest enrichments of questionnaire items (blue) and Basic Human Values (red). **c)** are enrichments of professional disciplines. First axes are bin-distances from archetype and second axes for a-b) are average item response values rescaled from [1 ; 5] to [-1 ; 1] and normalized frequency for c). 'slo' is the slope of a first order approximation of the curve, so **negative slope is enrichment and positive slope is depletion**.

4.2.2 Summary and strategy suggestions

The analysis yields a large number of enrichments and correlations for each archetype, of which an excerpt is visualized in Figure 4.4, a shortened overview is shown in Table 4.1, and elaborate presentation is given in Appendices C-E. Here, each archetype is presented in prose highlighting their individual characteristics. Each is examined and proposed a behavioral strategy with reference to existing literature when possible (primarily *Evolutionary Psychology* (2014) by Workman et. al. [2]).

Self-concern The *self-concern* archetype is unemotional and mostly interested in power and achievement. Occupied in fields which commonly don't involve working directly with people, whereas the depleted professional disciplines are the reverse. Lacks sympathy, tends to exploit others for own ends and looks down on others. Has a high degree of control through emotional stability and is hard to offend. Highest valued BHVs other than power and achievement are stimulation and hedonism and least valued BHVs are benevolence, conformity and tradition. Tends to be slightly younger (~5 years younger than furthest bin), particularly not retired individuals, not committed in relationships and has a tendency to exercise often. Comprises many unemployed individuals. There were found no significant correlations between distance from this archetype and behaviors measured through smartphones.

Discussion These are traits that are often prevalent in individuals with antisocial personality disorders (APD) and are typically lumped together using the term *psychopathy* by evolutionary psychologists [2]. People suffering from psychopathy, i.e. *psychopaths*, can use their manipulative tendencies to exploit others, gaining resources including sex, which is supported by studies that show ADP traits to correlate positively with an accelerated mating strategy [63,64]. Psychopathic traits are furthermore shown to lead to success in high-level business positions [65], which the current results, however, don't support. Enriched disciplines are rather those of solitary nature, and as such it is just as reasonable to speculate this archetype to have a form of passive aggressive withdrawal strategy, that don't necessarily attract only psychopaths and individuals with other personality disorders. Furthermore, considering Figure 4.3.b distances to this archetype are typically small, which in other words mean that many people somewhat resemble this archetype, which is inconsistent with the converging agreement between psychologists that only up to 3% of any larger population should suffer from sub-clinical psychopathy [43]. As such it might be an artifact of a possible tendency in many humans to be competitive and, depending on context, feel emotions of contempt and superiority towards others, all of which may only surface when filling out a survey. The safest label for this strategy is therefore *self-concern*.

Social-achievement The *social achievement* archetype is highly social, talkative and lively, yet suffers from mood swings, insecurity and gets hurt easily. Values are mostly stimulation, hedonism, power and achievement. Has high level of employment and works high-pace jobs that directly involve people, such as sales, management and CEO positions. Does not hold unskilled positions (food prep. worker, cashier) which resonates with the valuing of achievement and power. Curiously both smokes and exercises, which is somewhat contradictory but can be explained by the high held value of stimulation and hedonism. It may also indicate an ability to hold contradictory beliefs simultaneously, which is undoubtedly a useful feature for individuals whose behavioral strategy is to use social engagement to advance in hierarchies. Maintains social relations through mobile phone by calling and texting a lot, socializes and moves around frequently both day and night and explores new places (visits places only once). Typically makes long calls, sends more texts than receives and response-rate to texts or missed calls tends to depend on who the receiver is.

Discussion The BFTs and enrichments for this archetype resonate with known literature. Recall that this archetype is characterized by high levels of O, E and N. N is found to correlate positively with competitiveness and academic success for those who are resilient enough to cope with it^[66,67], which agrees with the finding that this archetype works in fast-pace/high-status jobs such as CEO and operations manager. Provided that the goals of this archetype are mainly achievement and power (curiously similar to the *self-concern* archetype), it seems clear that this strategy focuses on using social interaction as a tool for acquiring resources and power. To the knowledge of this author there is no strategy treated in EP literature which strongly resembles this one. It can be argued that this strategy may attract individuals with multiple complex profiles, since it is a fairly universal drive to be highly social - and if one is also sensitive interacting with people may be a matter of earning popularity and using vulnerability to get ahead. This is discussed by Nettle^[15]. On a highly speculative note, given the occupation profiles, it might even attract psychopaths. However, for now it is simply named after its main facet *social* and value *achievement*.

Reciprocal altruism Happy, lively and social, well-tempered, not worried, trusting, hard-working and dutiful describe the most diverse set of enriched facets for the *reciprocal altruism* archetype. It highly values conformity, benevolence, stimulation and achievement, and is strongly enriched for relationship commitment and employment. This archetype is more mature in age (closest bin is ~10 years older than furthest bin), and tends to work in disciplines that have direct interaction with people. Smoking is depleted and exercise is enriched, indicating self-discipline. Has many friends outside of the university, frequently uses phone for communication, goes out at night, doesn't meet people in predictable weekly pattern, goes to many different places and

tends to send the last text in chat conversations.

Discussion EP build on the notion reciprocal altruism first introduced by Trivers^[68]. The idea, which is also mentioned briefly in Section 2.1.3, is that altruism may have developed as a mechanism for optimizing group fitness, or simply as a personal strategy for storing value by giving it away and expecting a return of similar value at a later time of greater need. Much about this archetype resemble this, and as such it is deemed to represent *reciprocal altruism*.

Xenophobia-loyalism The *xenophobia-loyalism* archetype is uncreative, simple-minded and slow, values conformity, tradition and security and neither self-direction, stimulation nor hedonism. Works in health-care, office administration support and military, and particularly not computer science and arts. Slight tendency to be younger (~3 years younger than furthest bin), non-smoker, committed in relationship and also often students. Meets people in predictable weekly pattern, immediately responds to missed calls/texts, has long ongoing chats and shows up for class.

Discussion Workman makes the remark that "Social psychologists have long known that the roots of xenophobia, or hatred of strangers, may be traced back to identifying strongly with one's own group and negatively stereotyping those of other groups.". Going further back Darwin once noted: "The tribes inhabiting adjacent districts are almost always at war with each other [and yet] a savage will risk his own life to save that of a member of the same community"^[69]. Several enriched features of this archetypes resonate with characteristics commonly associated with xenophobia and loyalism. So the question of why this seemingly negative-sounding characterization should be an ESS for performing life's many tasks in such a way to increase fitness is fairly difficult to answer. Possibly this is remanence of a time where we lived in small tribes and had to react strongly to outsiders to survive, as noted by Darwin. However, others suggest that the ability to perceive sharp distinctions between "us and them" group members is crucial to the formation of lasting coalitions and is therefore an adaptive strategy^[70].

Freeriding The *freeriding* archetype is undutiful, disorderly, not hard-working, breaks rules, unworried, emotionally stable and does not plan ahead. Only values stimulation and hedonism, and particularly doesn't value security, achievement, tradition, power nor benevolence. Professional disciplines that are enriched near the archetype include performance and visual arts, cultural studies, language and literature studies as well as communications and social science, while there is depletion for business and medicine/health related disciplines. This archetype is younger (~6 years younger than furthest bin and not retired), not committed in relationship, tends

to smoke and not exercise, and is enriched for student and unemployed. Makes short calls, doesn't meet people in predictable weekly pattern, is very nocturnal, interacts with strangers and bad at responding to missed calls/texts.

Discussion Workman remarks: "*The evolution of cooperation among non-kin is greatly compromised by the existence of freeriders. Freeriders reap the benefits of cooperation without paying the costs and thus place themselves at a competitive advantage by exploiting cooperators. Many (e.g. Tooby et al., 2006)^[71] have argued that, given the inherent disadvantages of cooperating with freeriders, cooperation could not have evolved unless mechanisms for detecting and dealing with freeriders also evolved.*" While the behavior of the freerider - other than its immediate strategy - are not examined in literature, this archetype could be a candidate, to help probe further into the understanding of this strategy which relies on others to do the heavy lifting.

Withdrawal The *withdrawal* archetype is antisocial and reclusive, worried, depressed, suffers from mood swings and gets very emotional. Values none of the BHVs, and particularly not stimulation, achievement, conformity and power. Disciplines include arts, language and literature, computer science and natural sciences, and particularly not human interaction dependent disciplines in business, health-care and communication. No significant enrichment of age, but enriched for unemployment and retirement which indicates that this strategy might attract older people. Neither smokes nor exercises, which resonates with the low valuing of stimulation and hedonism, and is not committed in relationship. Doesn't meet people at campus, not nocturnal, meets people in predictable weekly pattern, rarely calls/texts/looks at phone, has few contacts, stays at home.

Discussion This archetype may also simply be called *depression*. Models in literature view depression as adapted response to adverse conditions, which may help the individual accept defeat^[72]. Others attribute it more complexity, arguing that it is an involuntary strategy to signal yielding towards the surroundings such as to terminate unnecessary conflict (which would otherwise likely have been lost by the depressed)^[73]. If the latter is true, the label *withdrawal* is appropriate.

4.3 Archetypes separate data better than traits

To investigate how well the archetypes, which perform the suggested strategies, separate data, the predictive power of archetypes as a target variable in a supervised learning paradigm is investigated. The prediction accuracy is compared to that of plain BFTs as well as PCs of QIs, of which there are also six. Data from the SensibleDTU experiment is used. As is shown in Section 4.2.1, different strategies tended to be adopted by people in different age groups, and since the population in the SensibleDTU dataset is strictly composed of students, conclusions must be drawn with a degree of respect for this bias. Details of the analysis are discussed in Section 7.2.4.

Optimized prediction models are trained on 38 indicators to predict low, medium or high distance to each of the archetypes. Figure 4.6 visualizes the accuracies, and explains high-level aspect of the applied procedure. It is compared to prediction accuracies for Big Five traits (Figure 4.5) and PCs of Big Five Inventory QIs (Figure 4.8) using the same approach. There are discovered to be six PCs in QI-space as shown in Figure 4.7. The remainder of this section presents the results.

Distances to archetypes are predicted with highest accuracies. Average prediction accuracies above baseline (AB) reported in Figure 4.6, 4.5 and 4.8 are 5.8%, 2.4 and 4.1%, respectively. As such PCs can be predicted better than BFTs and distance from archetypes can be predicted better than PCs. Prediction accuracy compares to, but is not better than what has already been reported for prediction of BFTs in previous studies [^{20,21}]. This may be upsetting in terms of the predictive promise of the behavioral indicators computed using the extended version of the bandicoot software (see Section 7.1.1), but on the other hand it also shows that archetypes do split the data better.

Archetypes are predicted better than their parts. In the current approach only C and E can be predicted well. Considering the *self-concern* archetype, prediction accuracy is 6.8% AB, yet out of these two BFTs this archetype only varies on C which

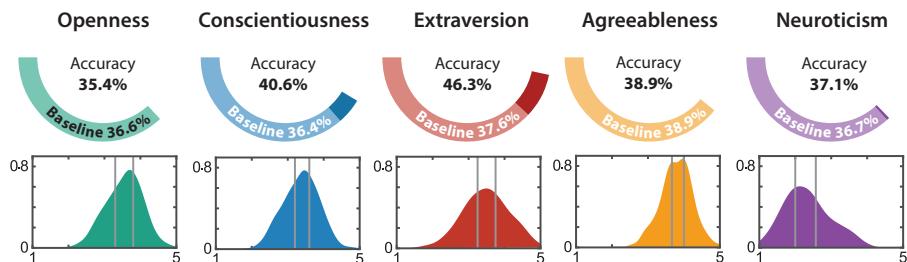


Figure 4.5: Three-class classification accuracy of BFTs predicted from behavioral indicators. Only C and E can be predicted. Same procedure as described in Figure 4.6.

has prediction accuracy 4.2% AB. It is unclear what accounts for this difference, and further analysis is required to understand it.

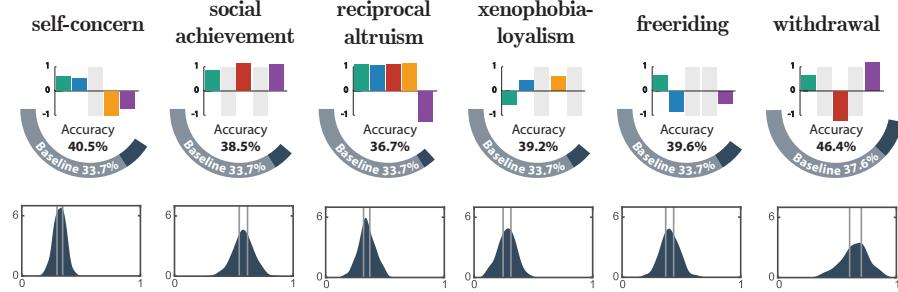


Figure 4.6: Three-class classification accuracy of archetypes predicted from behavioral indicators. Samples are split into three semi-balanced classes based on their target values (low, medium, high). Sample distributions are shown as KDE plots below each score diagram, and vertical lines show decision boundaries. Reported accuracies are averages over 100 shuffled stratified cross validation folds of parameter optimized random forest models (random search and grid search), with standard errors in the range of .2 to .5 percent. Baselines are calculated as the size of the largest class divided by number of samples. The bar diagrams show how each component/archetype is interpreted in terms of BFTs, where height is the normalized archetype position in BF space and non-consensus dimensions are shown in grey.

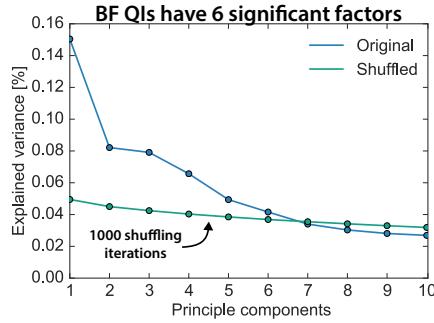


Figure 4.7: Percentage variance explained of PCs for original and domain shuffled Big Five questionnaire items. The green curve results from 1000 shuffling iterations which are plotted with high transparency. It shows that the six first principle components in Big Five questionnaire item space are significant, because none of the shuffling iteration can produce components of higher explained variance than the first six.

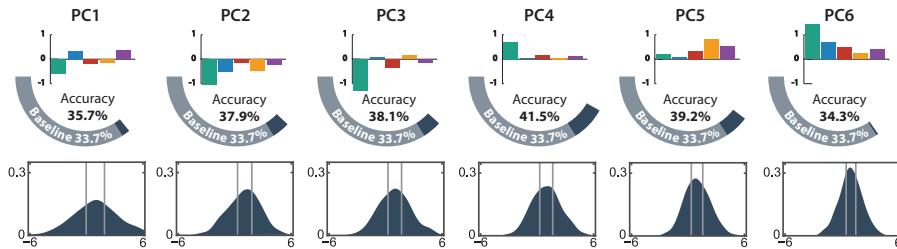


Figure 4.8: Three-class classification accuracy of PC components of Big Five Inventory QIs [74]. Same procedure as described in Figure 4.5. The bar diagrams show how each component/archetype is interpreted in terms of BFTs. Bar length correspond to inner products between PC and BF trait.

5 Discussion

Out of the six archetypes which have been suggested based on the conducted analysis, all except the *social achievement* archetype appears to fit a strategy known from the literature. There are, however, a number of aspects which are important to the validity of the results which needs to be discussed. Some are scientific in nature and some regard the ethical implications.

The Big Five model is "too crude" to capture fine-grained strategies. The Big Five factors result from factor analysis based in as many as 17.000 personality descriptive adjectives (English language), each of which was encoded into language arguably to fill some hole in comprehension. The model then lumps together all these words, into five factors that admittedly comprise a whole range of facets, but at the same time disregards much of the variance inherent to the language. For this reason, very fine-grained personality differences cannot be explained. For example, there is no trait in the Big Five model that successfully captures a persons psychopathic tendencies to form close relationships with people for the sake of personal gain. Trying to describe this behavior using the Big Five model is like trying to grab a coin with a boxing glove. As such, the strategies which are suggested in this work, are never less crude than the model they were inferred from.

Archetypes are not real people. There is an important notion to be made about the nature of these archetypal personalities. When presenting this work to people who are not familiar with it, and in particular when showing people the archetypes they have a strong tendency to quickly identify with one. Why people do this remains an open question but a speculation is that people may find it easier to understand things such as this by projecting their own self-image and feelings onto it. However, while people tend to do this, no one can really be described by just one archetype. In fact, by a geometrical argument (and as can be seen from Figure 4.3.b), no one is very close to an archetype - we are all described by *all* archetypes at the same time. That is to say, since the archetypes were derived from the point distribution of personalities in Big Five space using the PCHA algorithm run with some allowed slack, points are generally closer to the middle of the simplex than they are to the archetypes.

People are not necessarily stuck with a strategy. This study makes no claims about human nature beyond the strategies which can be inferred from the data. That is to say, whether a person can change strategy over a short period of time, or instantly depending on context, this study does not know. However, it does say that no person can fully assume more than one strategy *at a time*. Picturing the Pareto front in 3D trait-space, this makes sense - a point can only be in one location at a time. Yet, the result that the *reciprocal altruism* archetype tends to be far older than e.g. the *freeriding* suggests that people do change strategies over a lifetime.

The type of people who take surveys on the Internet may not represent a

general population. This very artifact may be what causes the distribution of traits to be so different between datasets that were collected online and per phone, as shown in Appendix G. However, it is unknown exactly how severe this biases the data, or if it biases it at all.

Archetypes separate data better than traits do, but is this really so surprising? This short discussion would have been replaced by a simulation if time had permitted it. However, since that is not the case a few remarks on the result that archetypes are better prediction targets than BFTs and QI-PCs are due. First: Maybe it's not so strange? Archetypes are geometrical extremes and the data is not well-separated so possible the best way to distinguish points is my learning which their nearest archetype is.

Confirmation bias may have influenced the analysis. We are all humans. And as humans analyzing humans we are constantly at risk of making analysis choices that are under bias of this fact. So while care has been taken to withdraw excitement over preliminary results and "read too much into things", it cannot be rejected that some scientifically unhealthy humanity has seeped into the interpretations of the results.

6 Conclusions

Rooted in the notion from evolutionary psychology that personality is an adapted behavioral strategy, the goal of thesis was to discover these strategies using the Pareto Task Inference paradigm. In the process of conducting this project a number of conclusions were drawn. Each is presented below.

- Simulation examples revealed the search for more than six archetypes computationally unfeasible unless a very large number of datapoints is available. This was further exemplified by illustrating that points on the convex hull of a data cluster are not contained within the Pareto front. While this constitutes an upper limit to the number of archetypes, it was argued that for Big Five personality data in five dimensions, the lower bound on number of archetypes should also be six.
- Using the same approach for computing archetypes in seven different datasets, six very similar personality archetypes emerged. The most deviating traits were taken to be "not in consensus" meaning they were left variable such that computing distance to the archetypes would disregard non-consensus trait.
- Archetypes were found to influence different variables to become enriched or depleted near them. Similarly many behavioral indicators measured through personal smartphones and computed using an extended version of the *bandicoot* software were found to correlate both positively and negatively with distance from archetypes.
- Six evolutionary behavioral strategies were suggested for the archetypes based on enrichments correlations. By reviewing the literature, four were to correspond to speculated strategies known to evolutionary psychologists already. This implies that: (1) the Big Five model is in fact an evolutionarily plausible way of carving up personality as argued by Buss^[3,4], and (2) two strategies *self-concern* and *social achievement* which are not accounted for in the literature may be a discovery that can lead to a better understanding of personality as behavioral strategy.
- There are significant age differences between behavioral strategies, indicating that humans adopt different strategies over the course of a lifetime.
- An individuals' distance to the inferred archetypes can be predicted with greater accuracy than its BF trait values and questionnaire item principal component values. Whether this points to something "special" is discussed, yet remained unanswered and requires more analysis.

- Principle components of personality questionnaire items provide a better prediction target indicating that strategies may belong in a PC-basis space rather than a BF-basis space.

Furthermore, the project yielded the following general value propositions.

- An extended version of the *bandicoot* behavioral indicator software for Python which can readily be applied by other researchers with access to the SensibleDTU dataset, to compute out-of-the-box measures of behavior.
- A fast Python implementation of the AA algorithm developed by Mørup et. al. [48], which is made available on PyPi^[75] and can be installed in any Python environment using the *pip* package manager with the following command: `pip install py_pcha`.

6.1 Future Research

Since this is an ongoing research project, there is a long list of problems to address which were not investigated or reported in this thesis. The list below presents the most important of those.

- Conduct a further investigation into the nature of the two strategies *self-concern* and *social achievement*, since the existing literature could not provide them with convincing strategies known to evolutionary psychology.
- Quantify how extreme individuals in each dataset are. What is the distribution of how many archetypes each individual is *close* to, where *close* is defined by some distance, such as a quarter of the average distance between archetypes. The analysis could be expanded to quantify the percentage of individuals with e.g. a certain job-type or education background are close to a particular archetype.
- In Section 7.2.4 it was shown there are six significant principal components in questionnaire inventory space and that these are better prediction targets than Big Five traits, in terms of increasing prediction accuracy. Personality archetypes are points in BF-space, but in light of this result could it be that they might rather belong in PC-space? This could be analyzed by collapsing multiple datasets and using PCA to find the common vectors of most explained variance, and then computing consensus archetypes similar to how it is explained in Section 7.2.1 but in 5D PC-space (not 6 - recall Figure 2.5).
- Investigate whether the discovered archetypes are predicted with higher accuracy than equally extreme well-separated random points on the convex hull.

- Employ a supervised basis-rotation optimization scheme that finds the projection of Big Five inventory items which provide the highest accuracy. Using this is an upper limit for how well the behavioral data can predict personality is found, and can be compared to the currently obtained accuracy for the archetypes.
- While the shape of performance functions of objectives has not been discussed in this thesis it has been of sporadic interest to this project to treat them as such. Sheftel explores this idea^[47], but considering Figure 2.3 one can imagine that the performance could have elliptic contours rather than circular. Using e.g. the standard deviations for each archetype trait as axes for the contours of the performance functions, shaped performance functions for each objective could be estimated and a more complex geometry of the emerging Pareto front could be computed.
- Based in the SensibleDTU dataset, an investigation into how archetype resemblance influences personal relationships is due. A hypothesis could be e.g. that the *freeriding* archetype should be observed in many work relationships with the *reciprocal altruism*.

7 Methods

For most data-driven projects, the typical pipeline includes extraction of data from some source and cleaning and reformatting of that data; there is typically feature engineering involved, which is the process of designing higher-level features based on a raw or semi-raw extract of data. After that comes outlier detection and typically also some variable transformation in order to balance the scales and distributions of signals in the data. All of these steps are called preprocessing, but will often require more work than any other task in the pipeline, simply because it constitutes a bottleneck to the quality of the scientific analysis, where errors made in the early, result in poor or, even worse, misleading results. Because of this, preprocessing must be carried out with the desired analysis in mind, and analysis must be conducted with awareness of the problems that might result from errors in preprocessing.

While this project mainly relies on personality data, for which the extraction pipeline is simple, a significant amount effort has been invested in development of a reusable pipeline to extract *behavioral indicators* from the SensibleDTU experiment dataset (see. Section 3.1). This pipeline constitutes an alternative value proposition of this research project because it can readily be applied by other researchers within Sensible DTU to extract high-level indicators for various purposes.

In the following, it is explained how the pipeline is implemented and behavior- and personality data from the SensibleDTU experiment is extracted. Then follows a brief section that explains the preprocessing of personality data from the remaining three datasources: *myPersonality*, *SAPA* and *MIDUS*. Finally technical details about the analysis are presented.

7.1 SensibleDTU data pre-processing

This pipeline transforms raw data from the Sensible DTU experiment into ordered, clean, datamatrices of behavioral indicators, personality traits and personality archetype distances for each participant, labeled \mathbf{X} , \mathbf{Y} and \mathbf{D} respectively. It is constructed using principles from software engineering such as reusability and version control. The pipeline is implemented as a set of smaller modules that each execute a task. All modules rely on the virtual IPython Notebook environment made available to researchers associated with the Social Fabric project (see Sec. 3.1). The pipeline is designed to allow the researcher to extract data for specific date ranges and with hour-of-day- and day-of-week- restrictions. It relies on a modified verision of the Python module *bandicoot* for computing behavioral indicators.

Consider Fig. 7.1 as a reference for the following introduction to the pipeline. The researcher first enters the `create_user_records` IPython notebook to create data records for each user in a desired time range using the `load_sensible_data` module. When data is loaded for a new time range, `load_sensible_data` caches

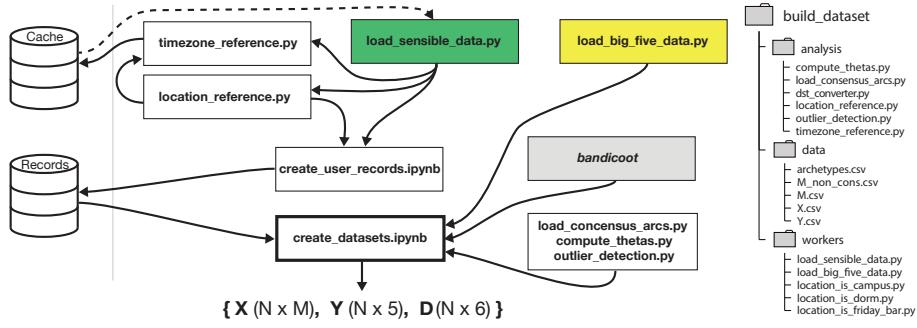


Figure 7.1: Preprocessing pipeline implemented on the virtual IPython notebook environment (see Sec. 3.1). The researcher first creates records for all users in a format that is accepted by the modified version of bandicoot, then uses bandicoot and pipeline submodules to compute three datamatrices: behavioral indicators \mathbf{X} , Big Five personality traits \mathbf{Y} and weighted euclidian distances from precomputed consensus archetypes \mathbf{D} .

the raw data to save time and API load in case of future uses, then looks for timezone and location references for each user which, if not found, are automatically created. The module then returns semi-raw datasets for each datatype (text, call, F2F, stop, screen) where timestamps are corrected according to timezone and DST, and where a location label ('home', 'campus', 'dorm', 'friday_bar' or 'other') is added to stop-records using a *point-inside-polygon* algorithm^[76]. Each semi-raw dataset is split into separate datasets for each user, and stored. Using the `create_datasets` IPython Notebook which relies on the `bandicoot` library to compute behavioral indicators for each user constructing a datamatrix \mathbf{X} which has a row for each user and a column for each indicator. For each user the corresponding Big Five personality traits are computed using the `load_big_five_data` module. Furthermore distances from precomputed archetypes are computed using the `compute_thetas` module.

The analysis pipeline is made available on Github at:

<https://github.com/ulfaslak/MScProject-build-dataset>

7.1.1 Feature Engineering

Feature engineering is carried out with focus on producing statistically independent behaviroal indicators, such as to capture as much of the behavioral spectrum with as little redundancy as possible. Views on importance of this is discussed in Appendix B. For extracting behavioral indicators (herinafter just *indicators*) the Python package `bandicoot` is used, and extended to handle multiple data-types. `bandicoot` is a call-detail-record (CDR) analysis package that offers phone companies and researchers to extract a comprehensive array of measurements of behavior based on text and call records^[20]. It works by loading records for individual users as comma separated

values (CSV) which it expects to contain certain attributes such as `timestamp`, `interaction`, `correspondent-id`, into `Record` objects and creating `User` objects which contain their `Records` as parameters as well as other parameters inferred from the loaded records such as `home` if the data contains location labels. The analyst then specifies whether to compute indicators for each day, week or month or the entire period comprising the data, whether to treat day and night or weekdays and weekends separately and finally which indicators to compute. Some of these rely on records being grouped into *conversations* which are defined as dyad exchanges that are not delayed by more than some time constant (typically 30 minutes or 1 hour). Two extensions to bandicoot were made:

1. Generalization of the `Record` object to be any kind of data which only requires to have `timestamp` as an attribute.
2. Modifying existing indicators to accept new data-types and adding new indicators that combine data from multiple channels.

An example of (2) is that using multiple data channels allows for creation of indicators that show how much use their phone when they are social (channels: F2F and screen), how much they tend to go in groups rather than be in dyads (time overlap between F2F *conversations*) and whether they tend to socialize with friends from the university *outside of the university* (F2F and stop-locations). A full list of the indicators which are used is shown in Appendix A, where it is marked which are added for this study and which already exist in the standard bandicoot version. The project code for the standard and extended versions of bandicoot is available, respectively, at:

```
github.com/yvesalexandre/bandicoot
github.com/ulfaslak/bandicoot
```

In this study, features were computed across the entire period since it was expected that the filtering criteria explained in Figure 3.1 would provide sufficient behavioral consistency in the data.

7.1.2 Variable Transformation

Two steps of variable transformation is employed:

1. Domain scaling
2. Form transformation

For (1): The indicators that the extended version of bandicoot returns are scaled in different domains. Most of them either return ratios or count variables, which can be problematic to analysis methods that are not scaling invariant. To correct for this,

the data are structured as a matrix with rows for users and columns for indicators, and the columns are subtracted by their means and divided by their standard deviations such as to enforce zero mean and unit standard deviation for all columns. The same approach is taken for the BF data (but not the distances from consensus archetypes, as this would produce negative distance which confuses the analysis).

For (2): Some indicators are not normally distributed which can cause problems in analysis because a density shift to one end of the domain can make trends in the data arise although there is none. In most cases, generally, and for all cases in this study, variables that are not normally distributed can be transformed using a scaling function like the logarithmic, square root, square, inverse, etc.. Out of the 38 behavioral indicators used in this study 12 were log-transformed.

7.1.3 Outlier Detection

As Figure 3.1 shows, the activity of participants in the SensibleDTU study wasn't constant over the entire period, and although care was taken only to use data from the period of most activity (containing 526 participants), all participants cannot be expected to provide useful data that informs the analysis. Therefore two levels of filtering was set up. The first level of filtering removes outliers using indicator activity thresholds based in careful assumptions about how people use smartphones. The aim is to filter out participants who were not actually using the phone. The assumptions are: (1) not everyone likes calling and texting, but everyone carries their phone with them on most days, (2) average over time people are within 1.5 meters distance of at least three people every day and (3) changes location at least twice daily. These conditions remove 84 participants from the analysis. The second level of filtering uses kernel density estimation (KDE) for removing outliers [77]. It employs a very simple scheme that computes kernel density (KD) (using the Python module `sklearn`) for each datapoint and removes those that fall below an empirically chosen threshold. This removes 30 participants from the analysis. After both levels of filtering **412 valid datapoints remain**.

7.1.4 Personality Data Extraction

The BF personality questionnaire used in the SensibleDTU study is a version of the BF Inventory [78], translated from English into Danish. It contains 44 individual items and each BFT is computed as the average of 7-10 items. To extract the traits, reference [74] was used.

7.2 Analysis

In the following, details about how analysis is conducted, which cannot be understood from Section 4, is explained. It gives an account of the details which enables others to

reproduce, however for very exact details and code (of which this section consciously includes very little), the reader is referred to the Github repository at [79].

7.2.1 Computing consensus archetypes

Section 4.1 explains that six consensus archetypes are found to emerge across seven datasets. In the following the methods used to arrive at this result is explained.

Employing archetypal analysis on a for a large number of points prove problematic when the domain is constrained (in this case from [-1 ; 1]), because noise in the data causes all subdomains of the space to fill up. For BF datapoints which exist in five dimensions, the full dataset therefore constitute a 5D cube (a hypercube), which leads the archetypes to snap the the vertices of this box. Therefore a different approach is taken. Considering Figure 2.5 the convex hull in five dimensions of just 200 points only contains about 20% of the datapoints in a point cloud. Furthermore, 200 points are found to constitute a small enough subset that points rarely go to the geometrical extremes of the domain. However, it is observed that using such a small number of points often leads to archetypes that are very different between iterations, and as such the following approach for finding the archetypes is taken (all references to row sorting use the algorithm shown in Python code in Listing 1):

1. Compute archetype 5000 times for subsets of 200 points for each of the seven datasets, using Python implementation of AA [75]. This yields a tensor \mathbf{X} which has dimensions $(N = 6) \times (M = 5) \times (K = 5000)$, where N are archetypes, M are BFTs and K are iteration slices, indexed i, j, k , respectively (see Figure 7.2).
2. For each dataset:
 - (a) For each pair of iteration slices (of which there are $(K - 1)^2 / 2$) sort rows to minimize average distance between points, and record the distance between the pairs. Remove the 50% of all iteration slices starting with the ones of highest average distance to other slices.
 - (b) In two-iteration steps, sort rows of \mathbf{X}_{k+1} with respect to rows of \mathbf{X}_k such as to minimize average distance between archetypes. Replace pair with average along 2nd axis. Repeat until there is only one slice left. This constitutes the archetypes of the dataset.
3. Sort resulting rows for each dataset using one as a reference. This yields the archetypes shown in Figure 4.1. Stack them in depth to form a matrix of shape $6 \times 5 \times 7$ and compute the SD and median matrices along the 2nd axis, Σ and \mathbf{A} respectively.

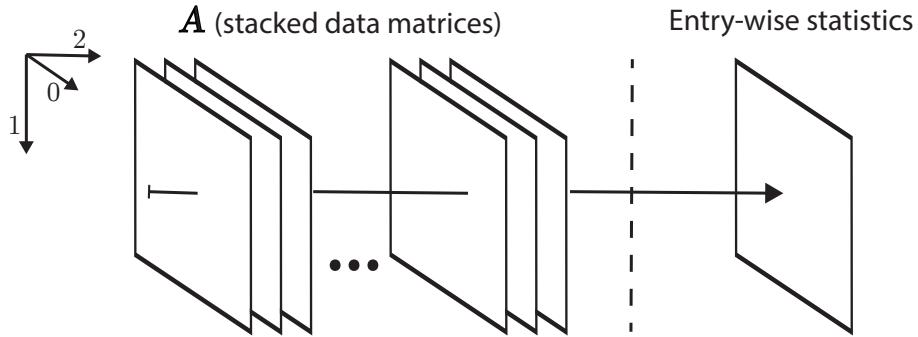


Figure 7.2: Example of how 2D matrices are stacked to shape a 3D tensor. The coordinate system explains the use of axis labels in the text, and the 'Entry-wise statistics' matrix illustrates how e.g. SD, mean or median values result from computing such along the 2nd axis.

4. Repeat step (3) 1000 times where columns in each slice are shuffled, and record the distribution of SDs along the 2nd axis. Take the maximum value of the 5th percentile as a threshold, λ , for maximally allowed SD between archetype traits (consensus threshold, see Figure 4.2).
5. Enforce λ on Σ and apply the resulting mask to A to reveal the consensus archetypes.

```

1 import numpy as np
2 def reorder_to(A, B):
3     """Return order of rows in A the best match rows in B."""
4     distance_matrix = np.ones((6, 6))*np.inf
5     for i, a in enumerate(A):
6         for ii, b in enumerate(B):
7             ba = (b-a)
8             distance_matrix[i, ii] = np.sqrt(np.dot(ba, ba))
9     reorder = [[] for _ in range(6)]
10    for _ in range(6):
11        ind = np.argmin(distance_matrix)
12        i, ii = ind//6, ind%6
13        reorder[ii] = i
14        distance_matrix[i, :] = np.inf
15        distance_matrix[:, ii] = np.inf
16    return reorder

```

Listing 7.1: "Python code for sorting rows of matrix A to match those of B "

7.2.2 Computing enrichments

For the early research conducted in this project, enrichment was analyzed using the ParTI software package developed by the Uri Alon group^[8]. However as Figure 2.6 illustrates the concept of enrichment is simple, and as such an implementation was

made in Python which is shown in Appendix H. It takes a matrix of precomputed distances between samples and archetypes D and a matrix of feature-values corresponding to the same samples, and returns a JSON object that contains enrichments for each feature near each archetype (example given in Appendix H). Standard errors are computed for each associated bin as SD/\sqrt{N} where N is the number of samples. Since the number of samples is large the standard error goes to 0 and is therefore not shown on the enrichment plots in Figure 4.4.

Statistical significance of enrichments is computed for the null hypothesis that: *any random permutation of bins will result in greater R^2 and slope values for a first order polynomial fitted to the bin values*. The significance level is set to $\alpha = 0.05$ and a Bonferroni multiple hypothesis correction is used^[80]. For enrichment of attributes and facets, of which there are collectively 780 which all stem from the SAPA dataset, enrichments are therefore only accepted if they have a p -value smaller than $\alpha/780$, since there are 780 hypotheses being tested. For BHVs of which there are 10 the threshold is $\alpha/10$.

7.2.3 Computing correlations

For computing linear relationships between behavioral indicator values and distance from archetypes, the Pearson product-moment correlation method is used^[81]. It estimates the ratio between the covariance of two variables normalized by the product of their individual variances, and as such does not provide an estimate for the slope of the covariance, only its strength and sign. This is, however, considered valid because the analysis goal is only to find evidence of the existence of relationships and not to model them.

Statistical significance is tested using a permutation test for the null hypothesis that: *randomly redefined coordinate pairs from the original coordinates yields greater correlation coefficients than the original data*. A significance level, $\alpha = 0.05$, is used and a Benjamin-Hochberg (BH) multiple comparisons correction is used. The BH correction is less strict than the Bonferroni correction mentioned in Section 7.2.2 but is widely accepted, especially by those that consider the Bonferroni correction too harsh. It sorts the p -values of N tests such that each is denoted p_i and states that in order for a test i to be significant it must obey:

$$p_i = \frac{\alpha}{N} i \quad (7.1)$$

Any test i which is significant, renders all preceding tests significant also.

7.2.4 Classification

It is outside of the scope of this thesis to give an exhaustive account of classification in machine learning and as such no dedicated section to explain this is included in

the theory chapter, but for good measure a brief introduction to the subject is given below. Supervised learning constitutes a very rigid analytical paradigm: given a set of *observations* $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]^\top \in \mathbb{R}^{n \times m}$ which has a corresponding set of *labels* $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n]^\top \in \mathbb{R}^{n \times 1}$ supervised learning aims to learn the labels that typically result from different types of observations. If the labels for example were "cat" or "dog", and features were "paw size", "body weight", "has whiskers", "barks", a supervised machine learning algorithm would be trained to learn the *decision boundary* that separates the classes and be able to predict whether new unseen data corresponded to a cat or a dog. Modeling is, however, not always the primary concern when applying supervised machine learning. As is the case with the current analysis, it is often desired to get an estimate of the *accuracy* with which a target variable can be predicted from a given type of data. To do this, the original dataset \mathbf{X} is split into a *training fold* $\mathbf{X}_{\text{train}}$ and a *test fold* \mathbf{X}_{test} . The model is then trained on $\mathbf{X}_{\text{train}}$ and the reported accuracy is the percentage of labels it guess correctly when it is tested on \mathbf{X}_{test} . This sectioning of training and test data can be done in a variety of ways to get the most information out of the available data. In the current analysis a *cross validation* approach is used for doing this. There exist many different models for classification one of the simplest being *decision trees* that learns to separate data by sequential decision rules which best separates it into different classes. An extension of this model is the *random forest* which trains many trees that in turn "vote" for which class a given sample should belongs to.

Section 7.2.4 investigates whether consensus archetypes provide better prediction targets than Big Five traits and QI PCs using a supervised machine learning approach. For each target accuracy is reported for a three-class optimized random forest classifier which is validated using a random sampling cross-validation approach (explained below). The optimization scheme optimizes the parameters `n_estimators` and `max_features` which are generally considered the most important parameters to tune for random forest models^[82]. `n_estimators` decides the number of trees which the forest model trains and uses for voting, and `max_features` is the maximal number of randomly selected features which are considered at each split. The scheme works in two steps: (1) estimate validation scores 100 times for random parameter settings and record the parameter combination $(p_1^{\text{opt}}, p_2^{\text{opt}})$ which lead to the highest score, then (2) do a 10×10 grid-search in the quadratic range of $[p_1^{\text{opt}} - \frac{p_1^{\text{opt}}}{2}; p_1^{\text{opt}} + \frac{p_1^{\text{opt}}}{2}]$ and $[p_2^{\text{opt}} - \frac{p_2^{\text{opt}}}{2}; p_2^{\text{opt}} + \frac{p_2^{\text{opt}}}{2}]$ to find the optimal parameters.

Cross-validation is performed over 100 folds. In each fold the model is trained on a subset of 80% randomly selected samples from the dataset and tested on the remaining 20%. The reported accuracy is the average of this. SD of accuracies across folds are around 2% so for 100 fold the standard error evaluates to around 0.2%.

References

- [1] Francis Galton. The measurement of character. 1949.
- [2] Lance Workman and Will Reader. *Evolutionary psychology*. Cambridge University Press, 2014.
- [3] David M Buss. Evolutionary personality psychology. *Annual review of psychology*, 42(1):459–491, 1991.
- [4] Randy J Larsen and David M Buss. Personality psychology. *Jastrebarsko: Naklada Slap*, 2008.
- [5] Denise L Newman, Auke Tellegen, and Thomas J Bouchard Jr. Individual differences in adult ego development: sources of influence in twins reared apart. *Journal of personality and social psychology*, 74(4):985, 1998.
- [6] Timothy C. Bates. Behavioral genetics (5th edition) robert plomin, john c. defries, gerald e. mcclearn, peter mcguffin. (2008). worth publishers, new york, 505 pp., us\$115.95, isbn 10 1 4292 0577 6. *Twin Research and Human Genetics*, 11:245–246, 4 2008.
- [7] Oren Shoval, Hila Sheftel, Guy Shinar, Yuval Hart, Omer Ramote, Avi Mayo, Erez Dekel, Kathryn Kavanagh, and Uri Alon. Evolutionary trade-offs, pareto optimality, and the geometry of phenotype space. *Science*, 336(6085):1157–1160, 2012.
- [8] Yuval Hart, Hila Sheftel, Jean Hausser, Pablo Szekely, Noa Bossel Ben-Moshe, Yael Korem, Avichai Tendler, Avraham E Mayo, and Uri Alon. Inferring biological tasks using pareto analysis of high-dimensional data. *nature methods*, 12(3):233–235, 2015.
- [9] Avichai Tendler, Avraham Mayo, and Uri Alon. Evolutionary tradeoffs, pareto optimality and the morphology of ammonite shells. *BMC systems biology*, 9(1):1, 2015.
- [10] Pablo Szekely, Yael Korem, Uri Moran, Avi Mayo, and Uri Alon. The mass-longevity triangle: Pareto optimality and the geometry of life-history trait space. *PLoS Comput Biol*, 11(10):e1004524, 2015.
- [11] Yael Korem, Pablo Szekely, Yuval Hart, Hila Sheftel, Jean Hausser, Avi Mayo, Michael E Rothenberg, Tomer Kalisky, and Uri Alon. Geometry of the gene expression space of individual cells. *PLoS Comput Biol*, 11(7):e1004224, 2015.
- [12] Christopher M Bishop. Pattern recognition. *Machine Learning*, 128, 2006.

- [13] Jerome H Barkow, Leda Cosmides, and John Tooby. *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford University Press, 1995.
- [14] Criticism of evolutionary psychology. <https://goo.gl/hWu6I5>. Accessed: 2016-01-07.
- [15] Daniel Nettle. The evolution of personality variation in humans and other animals. *American Psychologist*, 61(6):622, 2006.
- [16] Icek Ajzen. *Attitudes, personality, and behavior*. McGraw-Hill Education (UK), 2005.
- [17] Richard D Mann. A review of the relationships between personality and performance in small groups. *Psychological Bulletin*, 56(4):241, 1959.
- [18] Albert Richard Behnke and Jack H Wilmore. *Evaluation and regulation of body build and composition*. Prentice Hall, 1974.
- [19] Sampo V Paunonen. Big five factors of personality and replicated predictions of behavior. *Journal of personality and social psychology*, 84(2):411, 2003.
- [20] Yves-Alexandre de Montjoye, Jordi Quoidbach, Florent Robic, and Alex Sandy Pentland. Predicting personality using novel mobile phone-based metrics. In *Social computing, behavioral-cultural modeling and prediction*, pages 48–55. Springer, 2013.
- [21] Bjarke Mønsted, Anders Møllgaard, and Joachim Mathiesen. Phone-based metric as a predictor for basic personality traits. *arXiv preprint arXiv:1604.04696*, 2016.
- [22] Walter Mischel. Continuity and change in personality. *American Psychologist*, 24(11):1012, 1969.
- [23] Sarah Broadie. *Ethics with Aristotle*. Oxford University Press, 1991.
- [24] Geoffrey Ernest Richard Lloyd, John Chadwick, and William Neville Mann. *Hippocratic writings*. Penguin (Non-Classics), 1978.
- [25] Gordon W Allport and Henry S Odber. Trait-names: A psycho-lexical study. *Psychological monographs*, 47(1):i, 1936.
- [26] Raymond B Cattell. The description of personality: Principles and findings in a factor analysis. *The American Journal of Psychology*, 58(1):69–90, 1945.
- [27] John M Digman. Personality structure: Emergence of the five-factor model. *Annual review of psychology*, 41(1):417–440, 1990.

- [28] Lewis R Goldberg. Language and individual differences: The search for universals in personality lexicons. *Review of personality and social psychology*, 2(1):141–165, 1981.
- [29] A Festschrift Honoring Donald W Fiske. Personality research, methods, and theory.
- [30] Ernest C. Tupes and Raymond E. Christal. Recurrent personality factors based on trait ratings. *Journal of Personality*, 60(2):225 – 251, 1992.
- [31] International personality item pool. <http://ipip.org/>. Accessed: 2016-01-07.
- [32] Thomas J Bouchard Jr and John C Loehlin. Genes, evolution, and personality. *Behavior genetics*, 31(3):243–273, 2001.
- [33] Robert R McCrae and Paul T Costa Jr. Personality trait structure as a human universal. *American psychologist*, 52(5):509, 1997.
- [34] Michael Gurven, Christopher von Rueden, Maxim Massenkoff, Hillard Kaplan, and Marino Lero Vie. 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, 2013.
- [35] Daniel Cervone and Dylan T. Lott. Language and the languages of personality. *European Review, Eur. Rev, Eur Rev*, 15(4):419–437, 2007.
- [36] Drew Westen. A model and a method for uncovering the nomothetic from the idiographic: An alternative to the five-factor model? *Journal of Research in Personality*, 30(3):400–413, 1996.
- [37] M ARGYLE. Silent messages - mehrabian,a. *Contemporary Psychology, Contemp. Psychol, Cont Psycho, Contemp Psychol, Contemporary Psychology Apa Review of Books*, 18(5):239–240, 1973.
- [38] Gregory J Boyle, Lazar Stankov, and Raymond B Cattell. Measurement and statistical models in the study of personality and intelligence. In *International handbook of personality and intelligence*, pages 417–446. Springer, 1995.
- [39] Michael C Ashton, Kiboom Lee, Marco Perugini, Piotr Szarota, Reinout E De Vries, Lisa Di Blas, Kathleen Boies, and Boele De Raad. A six-factor structure of personality-descriptive adjectives: solutions from psycholexical studies in seven languages. *Journal of personality and social psychology*, 86(2):356, 2004.
- [40] Heather EP Cattell and Alan D Mead. The sixteen personality factor questionnaire (16pf). *The SAGE handbook of personality theory and assessment*, 2:135–178, 2008.

- [41] Colin Cooper. *Individual differences and personality*. Routledge, 2015.
- [42] David M Buss. How can evolutionary psychology successfully explain personality and individual differences? *Perspectives on Psychological Science*, 4(4):359–366, 2009.
- [43] Ann M Kring, Gerald C Davison, John M Neale, and Sheri Johnson. *Abnormal psychology*. Wiley, 2005.
- [44] John F Nash et al. Equilibrium points in n-person games. *Proc. Nat. Acad. Sci. USA*, 36(1):48–49, 1950.
- [45] Martin A Nowak. Five rules for the evolution of cooperation. *science*, 314(5805):1560–1563, 2006.
- [46] J Maynard Smith. The theory of games and the evolution of animal conflicts. *Journal of theoretical biology*, 47(1):209–221, 1974.
- [47] Hila Sheftel, Oren Shoval, Avi Mayo, and Uri Alon. The geometry of the pareto front in biological phenotype space. *Ecology and evolution*, 3(6):1471–1483, 2013.
- [48] Morten Mørup and Lars Kai Hansen. Archetypal analysis for machine learning and data mining. *Neurocomputing*, 80:54–63, 2012.
- [49] Adele Cutler and Leo Breiman. Archetypal analysis. *Technometrics, Technometrics, Technomet*, 36(4):338–347, 1994.
- [50] Python implementation of the aa algorithm. <https://pypi.python.org/pypi/PyMF>. Accessed: 2016-02-07.
- [51] Manuel Eugster and Friedrich Leisch. From spider-man to hero-archetypal analysis in r. 2009.
- [52] José M Bioucas-Dias. A variable splitting augmented lagrangian approach to linear spectral unmixing. In *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, pages 1–4. IEEE, 2009.
- [53] Jun Li and José M Bioucas-Dias. Minimum volume simplex analysis: A fast algorithm to unmix hyperspectral data. In *IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages III–250. IEEE, 2008.
- [54] Tsung-Han Chan, Chong-Yung Chi, Yu-Min Huang, and Wing-Kin Ma. A convex analysis-based minimum-volume enclosing simplex algorithm for hyperspectral unmixing. *IEEE Transactions on Signal Processing*, 57(11):4418–4432, 2009.
- [55] Social Fabric. Data documentation of the social fabric survey.

- [56] Arkadiusz Stopczynski, Vedran Sekara, Piotr Sapiezynski, Andrea Cuttome, Mette My Madsen, Jakob Eg Larsen, and Sune Lehmann. Measuring large-scale social networks with high resolution. *Plos One, Plos One*, 9(4):e95978, 2014.
- [57] Michal Kosinski, Sandra C Matz, Samuel D Gosling, Vesselin Popov, and David Stillwell. Facebook as a research tool for the social sciences: Opportunities, challenges, ethical considerations, and practical guidelines. *American Psychologist*, 70(6):543, 2015.
- [58] Shalom H Schwartz. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in experimental social psychology*, 25:1–65, 1992.
- [59] Shalom H Schwartz. An overview of the schwartz theory of basic values. *Online readings in Psychology and Culture*, 2(1):11, 2012.
- [60] Orville Gilbert Brim, Carol D Ryff, and Ronald C Kessler. *The MIDUS National Survey: An Overview*. University of Chicago Press, 2004.
- [61] David Condon. *An Organizational Framework for the Psychological Individual Differences: Integrating the Affective, Cognitive, and Conative Domains*. PhD thesis, NORTHWESTERN UNIVERSITY, 2014.
- [62] M Brent Donnellan, Frederick L Oswald, Brendan M Baird, and Richard E Lucas. The mini-ipip scales: tiny-yet-effective measures of the big five factors of personality. *Psychological assessment*, 18(2):192, 2006.
- [63] Peter K Jonason, Norman P Li, Gregory D Webster, and David P Schmitt. The dark triad: Facilitating a short-term mating strategy in men. *European Journal of Personality*, 23(1):5–18, 2009.
- [64] Peter K Jonason, Katherine A Valentine, Norman P Li, and Carmelita L Harbeson. Mate-selection and the dark triad: Facilitating a short-term mating strategy and creating a volatile environment. *Personality and Individual Differences*, 51(6):759–763, 2011.
- [65] Robert D Hare. Psychopathy: A clinical and forensic overview. *Psychiatric Clinics of North America*, 29(3):709–724, 2006.
- [66] Scott R Ross, Jane Stewart, Molly Mugge, and Brandy Fultz. The imposter phenomenon, achievement dispositions, and the five factor model. *Personality and Individual Differences*, 31(8):1347–1355, 2001.
- [67] James McKenzie, Mahdad Taghavi-Khonsary, and Gary Tindell. Neuroticism and academic achievement: the furneaux factor as a measure of academic rigour. *Personality and Individual Differences*, 29(1):3–11, 2000.

- [68] Robert L Trivers. The evolution of reciprocal altruism. *Quarterly review of biology*, pages 35–57, 1971.
- [69] Charles Darwin. *The descent of man, and selection in relation to sex*, volume 1. Murray, 1888.
- [70] Dennis L Krebs and Kathy Denton. Social illusions and self-deception: The evolution of biases in person perception. *Evolutionary social psychology*, pages 21–48, 1997.
- [71] John Tooby, Leda Cosmides, and Michael E Price. Cognitive adaptations for n-person exchange: the evolutionary roots of organizational behavior. *Managerial and Decision Economics*, 27(2-3):103–129, 2006.
- [72] Lewis Wolpert. Depression in an evolutionary context. *Philosophy, Ethics, and Humanities in Medicine*, 3(1):1, 2008.
- [73] John Price, Leon Sloman, Russell Gardner, Paul Gilbert, and Peter Rohde. The social competition hypothesis of depression. *The British Journal of Psychiatry*, 164(3):309–315, 1994.
- [74] John Srivastava. Big five inventory (bfi). <http://goo.gl/ayINfy>.
- [75] Python implementation of the fast aa algorithm. https://pypi.python.org/pypi/py_pcha. Accessed: 2016-02-07.
- [76] Patrick Jordan. Python implementation point-inside-polygon algorithm. <http://goo.gl/yjirNa>. Accessed: 2016-03-07.
- [77] Longin Jan Latecki, Aleksandar Lazarevic, and Dragoljub Pokrajac. Outlier detection with kernel density functions. In *International Workshop on Machine Learning and Data Mining in Pattern Recognition*, pages 61–75. Springer, 2007.
- [78] Oliver P John and Sanjay Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. *Handbook of personality: Theory and research*, 2(1999):102–138, 1999.
- [79] Ulf Aslak. Project code repository (github). <https:// goo.gl/h6oD9O>.
- [80] Olive Jean Dunn. Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293):52–64, 1961.
- [81] Karl Pearson. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58:240–242, 1895.
- [82] Sci-kit learn ensamble methods. <http://scikit-learn.org/stable/modules/ensemble.html>.

Appendices

A Full List of Extracted Indicators

Complete list of features extracted from the data and used in the analysis, including explanations. Features may be extracted from the sources `call`, `text`, `F2F`, `screen`, `stop`, either solely, or as combinations utilizing the interaction between measurements. Below, feature names are highlighted in bold, and the types used for each feature is listed in brackets: `[]`. A bracket in a bracket signifies that the feature is computed from the combination of the datatypes listed in the bracket. A short exemplary explanation is given for the first feature. Furthermore features are either computed as *summary* type which yields mean and standard deviation of a distribution, or as *scalar* type yielding a scalar. **Id** and **Data** are consistent with function and datatype names in project code. Asterisk '*' marks that the indicator is only a part of the extended bandicoot version, however note that most of the standard indicators have also been extended to accept new data-types.

Number of contacts Number of individuals interacted with. Requires more than one meeting for an individual to count. Is computed for `text` and `call` (together) and `stop`, resulting in two measures of *number of contacts*. For `stop` it simply measures the number of different places an individual goes. This feature captures the size of an individuals social circle. `F2F` is not included because correspondent Ids are inconsistent for non-study-participant correspondents.

Id: number_of_contacts Data: [[text, call], stop] Type: scalar.

Number of interactions Number of interaction made.

Id: number_of_interactions Data: [[text, call], F2F, stop] Type: scalar.

Ratio of call traffic and text traffic* Number of calls made divided by number of texts sent.

Id: ratio_call_text Data: [[text, call]] Type: scalar.

Duration concentration Percent of interactions that account for 80% of interaction duration. Takes basis in the 80/20 "Pareto" principle.

Id: percent_ei_percent_durations Data: [stop] Type: scalar.

Balance of interactions Percent of interactions directed outwards.

Id: balance_of_interactions Data: [text] Type: scalar.

Duration Duration of interaction. If datatype has a `duration`-attribute (`call`, `screen`, `stop`), this is used, otherwise interactions are grouped into conversations, i.e. series of breaks no longer than one hour, and the duration of these are used to compute average duration.

Id: `duration` ***Data:*** `[call, text, F2F, screen, stop]` ***Type:*** `scalar`.

Percent initiated conversations Percent conversations initiated. Conversations are inferred from series of interactions grouped by respondent, as clusters with no more than one hour long breaks. Standard deviation across percent initiated for set of correspondents, yielding standard deviation a measure of "selectiveness" in who the individual chooses to initiate interaction with. For `call`, an outgoing missed `call` counts as an initiated conversation, whereas an incoming missed `call`, even though returned, does not.

Id: `percent_initiated_conversations` ***Data:*** `[text, call]` ***Type:*** `summary`.

Percent concluded conversations* Similar to 'Percent initiated conversations' but for last interaction in conversation, i.e. `conclusion`. Call is not used because it produces no informative feature.

Id: `percent_concluded_conversations` ***Data:*** `[text]` ***Type:*** `summary`.

Response delay Response delay in conversations grouped by correspondents. Maximum delay is one hour.

Id: `response_delay` ***Data:*** `[call, text]` ***Type:*** `summary`.

Response rate Response rate to conversations initiated by correspondents, grouped by respondent. Response counts if returned within the first hour, otherwise it counts as a delay.

Id: `response_rate` ***Data:*** `[[call, text]]` ***Type:*** `summary`.

Overlap of social meetings* Proportional to average number of people together with when social. Computed from overlap of conversations. Captures propensity to go in groups.

Id: `overlap_conversations` ***Data:*** `[F2F]` ***Type:*** `scalar`.

Percent nocturnal Percent of activity that occurs between 22pm and 6am.

Id: `percent_nocturnal` ***Data:*** `[F2F, screen, stop]` ***Type:*** `scalar`.

Interevent time Average time between events. For screen this captures the average time between phone usage sessions.

Id: `ratio` ***Data:*** `[screen]` ***Type:*** `scalar`.

Phone use while social* Amount of phone use while social divided by amount of phone use while alone.

Id: *ratio_social_screen_alone_screen* ***Data:*** $[[screen, F2F]]$ ***Type:*** scalar.

Ratio of interactions at campus vs. elsewhere* No. interactions that occur within campus regions divided by no. interactions that occur outside of campus and inferred home regions.

Id: *ratio_interactions_campus_other* ***Data:*** $[[stop, F2F]]$ ***Type:*** scalar.

Percent interactions elsewhere with people from campus* Percent of all interactions that occur outside campus, dorms and inferred home regions, which are made with individuals that were also interacted with at campus. Captures level of non-curricular social engagement with school mates. Excludes home because a large group of study participants live at campus dorms, and would exhibit far greater values in this feature for reasons not necessarily related to social engagement.

Id: *percent_outside_campus_from_campus* ***Data:*** $[[stop, F2F]]$ ***Type:*** scalar.

Percent daily time at campus* Average percent of all daily time (24 h) spent at campus.

Id: *percent_at_campus* ***Data:*** $[stop]$ ***Type:*** scalar.

Daily average number of stranger interactions* Percentage of F2F conversations, i.e. connections grouped into series of less than 24 hour breaks, that occur only once. Captures an individual's propensity to engage in social activities with people outside of their social circle, such as talking with people in bars, or joining a group of mixed friends for a weekend in a summer house. May include some noise from non-social connections made in transportation and other public spaces.

Id: *number_of_contacts_less* ***Data:*** $[F2F]$ ***Type:*** scalar.

First seen response rate* Percent of texts that are responded to during first possible session, i.e. right away when observed by the individual.

Id: *first_seen_response_rate* ***Data:*** $[[screen, text]]$ ***Type:*** scalar.

Interaction autocorrelation* Average autocorrelation of F2F togetherness across an individual's dyadic relationships. Captures a facet of the kind of relationships an individual maintains, where a high value means that most relationships fit into a schedule and are non-spontaneous and low values mean most meetings are impromptu. Only includes dyads that are observed in conversation more than 5 times.

Id: *interaction_autocorrelation* ***Data:*** $[F2F]$ ***Type:*** scalar.

B A Note on Feature Engineering

An important notion to introduce when processing data and designing features for use in machine learning or signal processing, is that the amount of information, or entropy, contained in a data source depends not only on its size but also on the presence of patterns such as mutual information between latent variables and periodicity in series. In the present *Big Data* context this rarely enters as a problem of information scarcity, but rather as a source of error due to the risk of capturing uninformative patterns in the feature design, which may go unobserved thus leading to false conclusions. To give an intuition of how this can cause problems, consider the following example.

`feature A` measures the fraction of time a user is looking at his or her phone while socially engaged, and `feature B` measures the unique number of friends that the user has through a given period. If one were to plot the distribution of these variables for a set of individuals against each other, it would yield an almost perfect inverse correlation, begging the interpretation that looking at your phone while socially engaged means you have less friends. The problem with this analysis is that `feature A` is normalized by the summed time socially engaged, which unsurprisingly correlates with `feature B`. The consequence is that the latent variable the analyst is trying to get at, namely the propensity to interact with one's phone while social, vanishes in product with a variable that so strongly correlates with the target. At the same time, removing the normalization term creates the inverse problem, because the summed time of phone use while socially engaged necessarily depends on how much a user engages socially and thereby number of unique friends. In the given case, a better design of `feature A` would be to measure the *amount of phone use while socially engaged vs. while alone* in which case it should behave independently from other features.

Uninformative patterns are expected for mathematical, systematic or other boring reasons, and are often undesired since they wash out interesting informative patterns, like in the example given above. Patterns which are considered most informative are those that emerge unexpectedly or in line with a hypothesis and demands explanation only through interpretation or further research. In general, careful consideration must be made to capture mostly patterns that informs the analysis. Features must be designed to cultivate patterns that express properties of the system which are not entirely self-evident. This excludes not only features capturing uninformative patterns, but also a large number of other features we as humans know to be mutually informative or in correlation. Consider how, for a set of individuals, the *number of calls made vs. number of calls received* would, aligned with our expectations, be in strong corre-

lation, but that *number of calls made and received vs. percent of calls going out*, would not. Both feature pairs are derived from the same call-log data, but as illustrated in Fig. B.3, a far more informative pattern is captured with the latter pair.

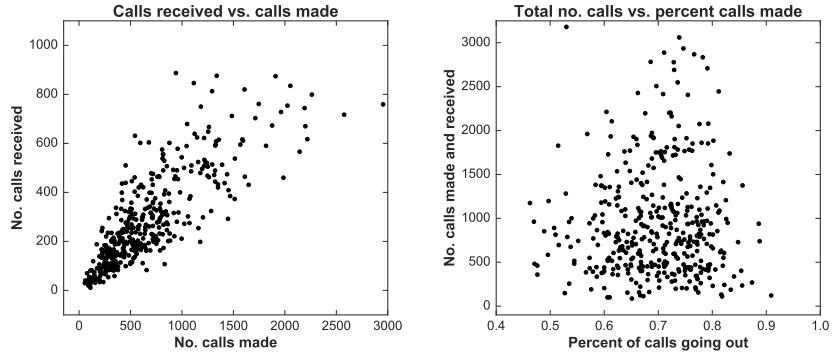


Figure B.3: Casting features differently changes the pattern they capture. Each plot represents a different way of casting the same call-log data, but the right one obviously captures a more informative pattern. It allows one to make genuinely interesting observations, such as how only individuals with an outgoing to incoming call ratio between 0.7 and 0.8 make and receive far above average number of calls.

In practice it is tedious to engineer features to only capture patterns that inform the analysis. In fact, trying to achieve this often creates problems since the task of developing the necessary complex features requires significantly more coding, which increases the risk of generating uncaught bugs. In this work, it is attempted to strike a balance between the informative and the simple. In other words, great care has been taken to understand the data which has led to the realization that it is not feasible to completely avoid capturing uninformative patterns, and as such features are designed in a way that attempts to minimize these, not exclude them.

C Questionnaire Enrichment Table

Slopes of enrichment curves for Big Five questionnaire items near consensus personality archetypes. Negative slope indicate enrichment and positive slope indicate depletion. * denotes statistical significance after Bonferroni correction.

	<i>self-concern</i>	<i>social achievement</i>	<i>reciprocal altruism</i>	<i>xenophobia-loyalism</i>	<i>freeriding</i>	<i>withdrawal</i>
Get overwhelmed by emotions [N]	0.114*	-0.099*	0.091*	-0.0249	0.0762	-0.110*
Indifferent to others' emotions [!A]	-0.112*	0.0374	0.0588	0.0266	-0.0205	-0.0121
Have a soft heart [A]	0.109*	-0.0426	-0.0423	-0.0326	0.0428	-0.0023
Am out for my own personal gain [!A]	-0.107*	0.0068	0.0541	0.0306	-0.0217	0.0028
Believe that I am better than others [!A]	-0.102*	-0.0109	0.0395	0.0518	-0.0178	0.0010
Feel others' emotions [A]	0.101*	-0.0604	-0.0593	-0.0211	0.0316	0.0128
My feelings are easily hurt [N]	0.100*	-0.085*	0.084*	-0.0273	0.0696	-0.097*
Am sensitive to the needs of others [A]	0.100*	-0.0416	-0.0612	-0.0209	0.0303	0.0197
Use others for my own ends [!A]	-0.098*	0.0002	0.0544	0.0407	-0.0310	-0.0015
Take advantage of others [!A]	-0.094*	-0.0090	0.0604	0.0273	-0.0234	-0.0200
Look down on others [!A]	-0.087*	-0.0058	0.0728	0.0297	-0.0097	-0.0353
Inquire about others' well-being [A]	0.084*	-0.0543	-0.0698	-0.0131	0.0201	0.0383
Am a talkative person [E]	0.0287	-0.153*	-0.094*	0.0260	0.0006	0.123*
Can be stirred up easily [N]	0.0587	-0.095*	0.0705	-0.0119	0.0667	-0.0730
Need reassurance [N]	0.0660	-0.082*	0.0707	-0.0227	0.0445	-0.0766
Am often down in the dumps [N]	0.0656	-0.0440	0.142*	0.0069	0.0516	-0.139*
Do too little work [!C]	0.0116	0.0007	0.112*	0.0347	-0.099*	-0.0452
Find it difficult to get down to work [!C]	0.0272	-0.0158	0.107*	0.0294	-0.090*	-0.0537
Waste my time [!C]	0.0167	-0.0077	0.102*	0.0236	-0.0753	-0.0543
Neglect my duties [!C]	0.0131	-0.0105	0.098*	0.0290	-0.089*	-0.0398
Get easily agitated [N]	0.0418	-0.0729	0.098*	-0.0121	0.0591	-0.095*
Make careless mistakes [!C]	0.0193	-0.0305	0.088*	0.0200	-0.0661	-0.0331
Feel that most people can't be trusted [!A]	-0.0599	0.0053	0.078*	0.0155	0.0233	-0.0559
Need a creative outlet [O]	0.0158	-0.0406	-0.0035	0.101*	-0.0245	-0.0123
Am not interested in abstract ideas [!O]	0.0094	0.0062	0.0170	-0.098*	0.0339	0.0063
Know the answers to many questions [O]	-0.0270	-0.0159	-0.0400	0.098*	-0.0207	0.0296
Am full of ideas [O]	-0.0042	-0.0331	-0.0346	0.096*	-0.0156	0.0201
Have a rich vocabulary [O]	-0.0085	-0.0175	-0.0268	0.094*	-0.0226	0.0128
Have a vivid imagination [O]	0.0024	-0.0312	-0.0154	0.087*	-0.0193	-0.0021
Can handle complex problems [O]	-0.0258	-0.0088	-0.0444	0.086*	-0.0144	0.0214
Complete my duties as soon as possible [C]	-0.0036	-0.0034	-0.087*	-0.0349	0.112*	0.0349
Keep things tidy [C]	-0.0224	-0.0002	-0.0732	-0.0417	0.102*	0.0244
Start tasks right away [C]	-0.0149	-0.0109	-0.091*	-0.0154	0.099*	0.0442
Stick to the rules [C]	0.0167	0.0063	-0.0424	-0.0633	0.099*	-0.0104
Do things by the book [C]	-0.0101	0.0236	-0.0314	-0.0577	0.093*	-0.0215
Like order [C]	-0.0049	-0.0028	-0.0369	-0.0352	0.089*	-0.0072
Do things according to a plan [C]	-0.0152	0.0005	-0.0547	-0.0295	0.088*	0.0086
Talk to many diff. people at parties [E]	0.0119	-0.137*	-0.123*	0.0086	-0.0103	0.153*
Enjoy interactions less than others [!E]	-0.0373	0.122*	0.128*	0.0206	-0.0036	-0.148*
Make friends easily [E]	0.0234	-0.117*	-0.124*	-0.0009	0.0009	0.146*
Keep in background on social occasions [!E]	-0.0080	0.121*	0.105*	-0.0023	0.0237	-0.144*
Can easily get some life into a dull party [E]	0.0043	-0.126*	-0.110*	0.0262	-0.0086	0.141*
Prefer reading to meeting people [!E]	-0.0099	0.102*	0.099*	0.0296	0.0027	-0.137*
Am mostly quiet when with other people [!E]	-0.0084	0.126*	0.103*	-0.0052	0.0070	-0.135*
Start conversations [E]	0.0146	-0.116*	-0.119*	0.0125	0.0072	0.135*
Moods go up and down easily [N]	0.0642	-0.088*	0.130*	0.0022	0.0515	-0.132*
Like mixing with people [E]	0.0258	-0.121*	-0.113*	-0.0050	-0.0105	0.132*
Am a worrier [N]	0.0777	-0.0634	0.088*	-0.0337	0.094*	-0.126*
Enjoy meeting new people [E]	0.0260	-0.109*	-0.102*	0.0106	-0.0122	0.120*
Panic easily [N]	0.081*	-0.0641	0.108*	-0.0251	0.080*	-0.117*
Get upset easily [N]	0.0618	-0.083*	0.098*	-0.0234	0.080*	-0.110*
Am rather lively [E]	0.0210	-0.090*	-0.093*	0.0150	-0.0040	0.108*
Get caught up in my problems [N]	0.0582	-0.0622	0.094*	0.0073	0.0450	-0.099*

D Basic Human Values Enrichment Table

Slopes of enrichment curves for Basic Human Values near consensus personality archetypes. Negative slope indicate enrichment and positive slope indicate depletion. * denotes statistical significance after Bonferroni correction.

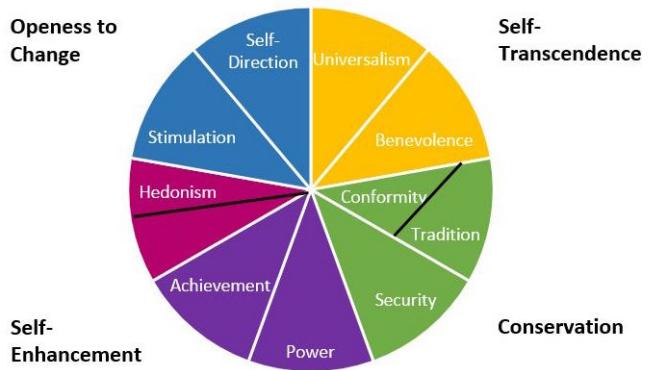
	<i>self-concern</i>	<i>social achievement</i>	<i>reciprocal altruism</i>	<i>xenophobia-loyalism</i>	<i>freeriding</i>	<i>withdrawal</i>
Power	-0.030*	-0.013*	-0.0008	-0.0048	0.013*	0.010*
Benevolence	0.021*	-0.0043	-0.020*	-0.0043	0.009*	0.007*
Hedonism	-0.008*	-0.016*	-0.0004	0.013*	-0.006*	0.006*
Self-Direction	-0.0049	-0.0068	-0.0074	0.026*	-0.0017	0.0025
Conformity	0.014*	0.001*	-0.023*	-0.024*	0.023*	0.011*
Tradition	0.014*	0.0037	-0.014*	-0.021*	0.010*	0.004*
Universalism	0.012*	-0.008*	-0.012*	0.017*	-0.0002	0.0014
Security	0.0003	-0.004*	-0.012*	-0.015*	0.020*	0.008*
Achievement	-0.011*	-0.010*	-0.017*	0.0075	0.018*	0.012*
Stimulation	-0.011*	-0.017*	-0.018*	0.022*	-0.007*	0.024*

E Behavior Correlation Table

Correlation coefficients between distance from CAs and behavioral indicators. * denotes statistical significance after Bonferroni correction.

	<i>self-concern</i>	<i>social achievement</i>	<i>reciprocal altruism</i>	<i>xenophobia-loyalism</i>	<i>freeriding</i>	<i>withdrawal</i>
Response time to missed calls SD	-0.074837	-0.018582	-0.032529	0.055680	-0.045662	0.062105
Frequently calls and texts	0.035789	-0.21798*	-0.13202*	-0.058639	0.069914	0.13381*
Has many phone contacts	0.024015	-0.20536*	-0.091237	0.026240	-0.017387	0.10454*
Visits many places only once	-0.060862	-0.18401*	-0.065878	0.073681	-0.080670	0.062021
Initiates most phone conversations	-0.014207	-0.17718*	0.006292	0.005671	0.088067	-0.061423
Goes places often	-0.010255	-0.14687*	-0.066594	-0.084382	0.080784	0.058721
Visits many different places	0.038093	-0.14040*	-0.099973	0.073829	-0.035228	0.039980
Concludes text conversations	-0.009717	-0.13497*	-0.10267*	-0.080921	0.11701*	0.062394
Initiates most text conversations	-0.017914	-0.12941*	-0.044484	-0.035294	0.086148	0.009535
Getting final word when texting SD	-0.054429	-0.12736*	-0.048257	0.085706	-0.036519	0.071641
Spends short time at locations	-0.013552	-0.12319*	-0.10799*	-0.077596	0.055401	0.076301
Calls some people more than others	-0.061514	0.11050*	0.013780	0.017749	-0.062601	0.030792
Selective text and call response rate	-0.065717	-0.10107*	-0.078203	-0.039442	0.068964	0.040896
Spends long time with people	0.022206	-0.099960	0.044219	0.015243	0.070682	-0.059171
Around people a lot	0.009494	-0.088931	-0.051458	-0.010177	0.053433	0.037501
Mostly goes in groups	0.000515	-0.013708	-0.028966	0.015795	0.023216	-0.021529
Conversation initiation rate SD	-0.039325	-0.063217	0.009150	0.14673*	-0.081758	-0.001845
Returns calls and responds to texts	-0.011807	-0.048248	-0.082415	-0.12769*	0.10628*	0.072745
Has long ongoing text conversations	0.045167	-0.030086	-0.053182	-0.10949*	0.061144	0.022334
Responds quickly to texts	-0.022331	0.092948	0.028020	-0.093492	0.022025	0.031446
Response time to texts SD	-0.027276	-0.025059	-0.007418	0.084515	-0.052311	0.013118
Spends time at campus	0.020200	-0.045188	0.015450	-0.082320	0.053027	-0.019011
Keeps phone sessions short	-0.017042	0.003843	-0.021614	-0.073513	0.036846	0.026214
Makes long calls	0.037902	-0.10913*	-0.009160	-0.018521	0.19279*	-0.085181
Meets with people in weekly patterns	0.019674	0.090653	0.11266*	-0.15568*	0.18400*	-0.14523*
Sends more texts than receives	0.001931	-0.12455*	-0.041272	-0.052693	0.16598*	-0.019525
Interacts with strangers	0.009526	-0.030696	-0.005149	0.048621	-0.15726*	0.051560
Uses phone at night	0.002887	0.049914	-0.025226	0.053391	-0.14350*	0.045076
Responds to texts immediately	-0.024757	0.042422	-0.090559	-0.081023	0.11332*	0.064325
Avoids looking at phone when social	0.018779	-0.006039	-0.057113	-0.007475	0.094036	0.015023
Prefers calling over texting	-0.033625	-0.014632	0.077649	0.075592	-0.084600	-0.052389
Mostly meets people at campus	-0.079546	0.018802	0.14031*	-0.035004	0.044647	0.21230*
Goes many places at night	0.004259	0.008385	-0.11848*	0.076401	-0.14232*	0.18953*
Meets people at night	0.074257	-0.11076*	-0.084527	0.075370	-0.10698*	0.12540*
Doesnt look at phone too often	0.028191	0.10047*	0.083603	0.10003*	-0.046390	0.11075*
Even time-location distribution	0.001487	-0.064028	-0.069595	0.027634	-0.000840	0.10922*
Meets people from uni. outside uni.	0.055556	-0.017454	-0.044660	-0.040767	-0.024956	0.074052
Responds quickly to missed calls	-0.012016	-0.023555	-0.036409	0.041238	0.020437	0.051917

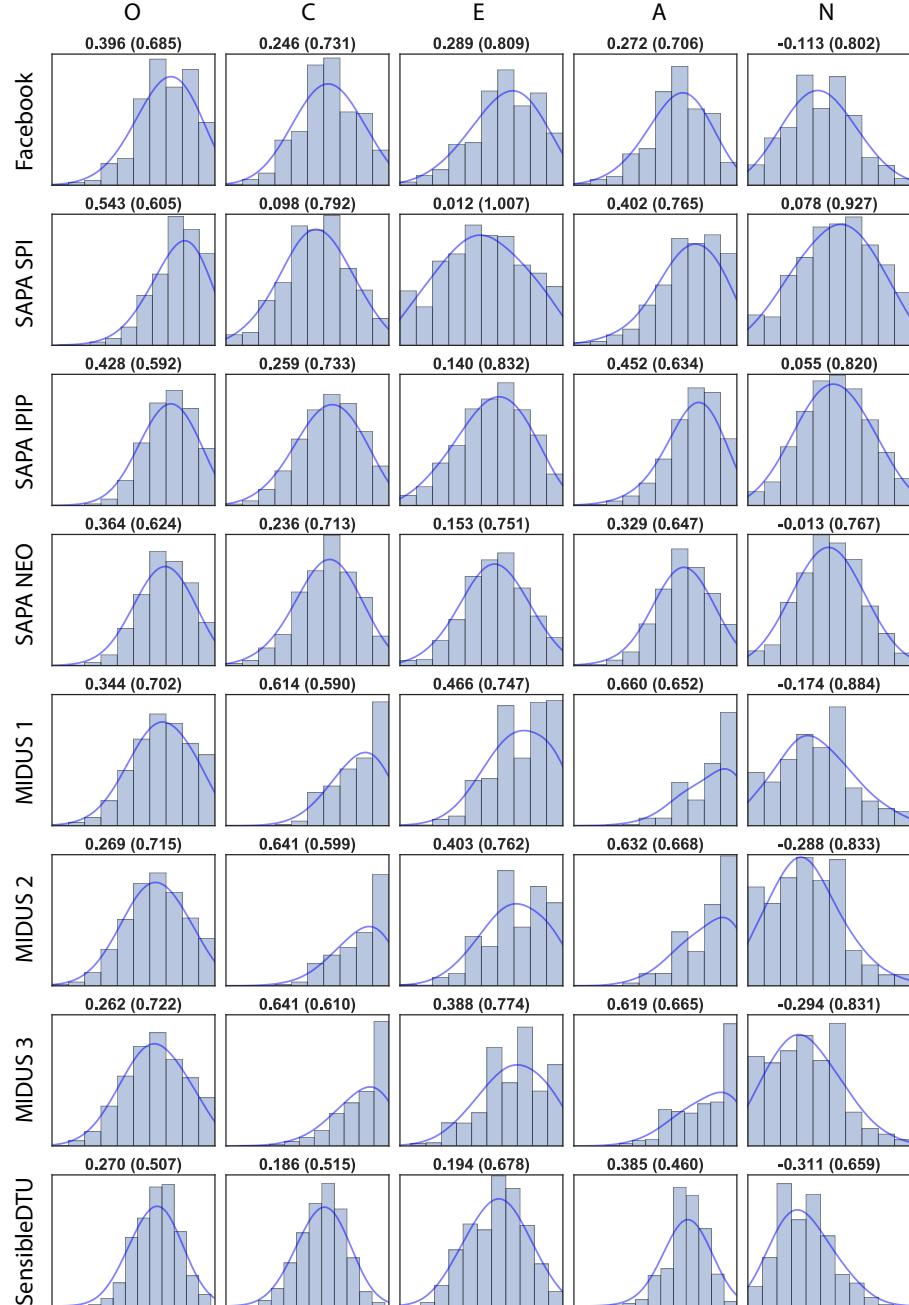
F Theory of Basic Human Values



Self-direction	Power	Spirituality
Freedom	Social power	A spiritual life
Creativity	Wealth	Meaning in life
Independent	Authority	Inner harmony
Choosing own goals	Preserving my public image	Detachment
curious	Social recognition	
Self-respect		
Stimulation	Security	Benevolence
An exciting life	National security	Helpful
A varied life	Reciprocation of favors	Responsible
Daring	Family security	Forgiving
	Sense of belonging	Honest
	Social order	Loyal
Hedonism	Healthy	Mature love
Pleasure	Clean	True friendship
Enjoying life		
Achievement	Conformity	Universalism
Ambitious	Obedient	Equality
Influential	Self-discipline	Unity with nature
Capable	Politeness	Wisdom
Successful	Honoring of parents and elders	A world of beauty
Intelligent		Social justice
Self-respect	Respect for tradition	Broad-minded
Tradition		Protecting the environment
	Devout	A world at peace
	Accepting my portion in life	
	Humble	
	Moderate	

G Distributions of Big Five datasets

Trait distributions of each Big Five dataset. Titles are formatted as mean (SD), histograms are normalized and horizontal axes range from minimum to maximum response value for the respective inventory.



H Code for enrichment of features near archetypes

```
1 import numpy as np
2 from collections import Counter
3
4 def compute_a_c_bins(D, T):
5     """Compute enrichment for features T using precomputed arc-
6     distances, D.
7
8     Parameters
9     -----
10    D : numpy.2darray (N, M)
11        Precomputed distances between N points and M archetypes.
12    T: pandas.DataFrame (N, K)
13        K features for N individuals matching rows in D. Each
14        feature (column) must have an associated column label.
15
16    Output
17    -----
18    a_c_bins : json-dict
19        JSON data structure storing enrichments near each archetype
20        for each
21        feature.
22
23    Example
24    -----
25    {'0': {
26        'age': {
27            # 10 bin averages
28            [34.1, 33.8, 33.2, 33.4, 32.5, 32.0, 32.2, 31.0,
29             30.9, 30.5],
30            32.35, # mean
31            30.5, # minimum
32            34.1, # maximum
33            1.204, # SD
34            0.0231 # standard error
35        },
36        ...
37        'job status': [
38            Counter({'employed': 42, ... , 'student': 12}),
39            ...
40            Counter({'employed': 12, ... , 'student': 42})
41        ], # 10 bin Counter dictionaries
42        set(['employed', ... , 'student']) # set of
43        attributes
44    }
45    }
46
47    """
48    a_c_bins = {}
```

```

45     # Loop over arcs
46     for a in range(D.shape[1]):
47         Da = D[:, a] # distances from arc
48         c_bins = {}
49         # Loop over enrichment features and build bins for each (
50         # b_bins)
51         for c in df.columns:
52             # Sort attributes into bins
53             ca = np.vstack([df[c], Da]).T
54             ca = ca[ca[:, 1].argsort()]
55             ca = ca[np.array(map(lambda v: v==v, list(ca[:, 0]))),
56                     :]
57
58             N = ca.shape[0]
59             bins = []
60
61             bin_intervals = zip(
62                 np.arange(0*N, 1*N, 0.1*N),
63                 np.arange(0.1*N, 1.1*N, 0.1*N))
64
65             # Continuous feature
66             if type(ca[0, 0]) is not str:
67                 bins_std, bins_err = [], []
68                 for p_low, p_upp in bin_intervals:
69                     bins.append(np.mean(ca[p_low:p_upp, 0]))
70                     bins_std.append(np.std(ca[p_low:p_upp, 0]))
71                     bins_err.append(
72                         np.std(ca[p_low:p_upp, 0]) / np.sqrt(len(ca
73 [p_low:p_upp, 0])))
74
75                 c_bins[c] = (
76                     bins, np.mean(ca[:, 0]), np.min(ca[:, 0]),
77                     np.max(ca[:, 0]), bins_std, bins_err
78 )
79
80             # Categorical feature
81             else:
82                 attrs = set()
83                 for p_low, p_upp in bin_intervals:
84                     counter = Counter(ca[p_low:p_upp, 0])
85                     bins.append(counter)
86                     attrs.update(counter.keys())
87                 c_bins[c] = (bins, list(attrs))
88
89             a_c_bins[a] = c_bins
90
91     return a_c_bins

```

Listing H.2: "Python code for computing enrichments near archetypes"

I Online

a) Go to ulfaslak.com/master_thesis/online_appendix1.pdf.