

The Semantic Space of Intellectual Humility

Markus Christen¹, Mark Alfano², and Brian Robinson³

¹ University of Zurich, Zurich, Switzerland
christen@ethik.uzh.ch

² University of Oregon, Eugene OR, USA
mark.alfano@gmail.com

³ Michigan State University, East Lansing MI, USA
prof.brobinson@gmail.com

Abstract. Intellectual humility is an interesting but underexplored disposition. The claim “I am (intellectually) humble” seems paradoxical in that someone who has the disposition in question would not typically volunteer it. There is an explanatory gap between the meaning of the sentence and the meaning the speaker expresses by uttering it. We therefore suggest analyzing intellectual humility semantically, using a psycholexical approach that focuses on both synonyms and antonyms of ‘intellectual humility’. We present a thesaurus-based method to map the semantic space of intellectual humility as a heuristic to support philosophical and psychological analysis of this disposition. We find three semantic clusters that compose intellectual humility: the sensible self, the discreet self, and the inquisitive self; likewise, we find three clusters that compose its contraries: the overrated self, the underrated other, and the underrated self.

Keywords: Intellectual humility, psycholexical analysis, semantics, synonymy, antonymy

1 Introduction

The study of personality and conceptions of personality has been pursued by psychologists and other researchers in various ways, including among others observations in laboratory settings and field experiments, correlational studies of survey responses, and psycholexical analyses. The present research embodies the latter methodology, and is informed by both philosophical theory and mathematical modeling tools developed in physical science.

Psycholexical analysis dates back to Francis Galton’s *Measurement of Character* [1]. The basic idea is that, all else being equal, a natural language is more likely to include a predicate for a property to the extent that the property is important to those who speak the language. English has the word ‘defenestration’ because it’s been important to be able to talk about events in which someone is thrown out a window. It lacks a word for someone being thrown in a window because, over the last several centuries, such events didn’t seem sufficiently worth talking about. This is not to say that *every* phrase or term refers. There are no

unicorns despite the existence of the term ‘unicorn’. Nor is it to say that everything worth talking about is represented by a phrase or singular term. Words are sometimes coined because new phenomena come into existence or become important; words are also sometimes coined because extant phenomena could not otherwise be parsimoniously described and explained. Sometimes a speaker coins words to describe or explain phenomena for which a word already exists, but of which the coiner is ignorant. So words that are synonyms emerge, further emphasizing the importance of the phenomena referred to. Regardless, the rough generalization that there is a strong positive correlation between the importance of phenomena in the lives of the speakers of a language and the probability of the existence of a term in the language that refers to those phenomena is hard to deny. If this is on the right track, studying psychological language is an indirect way of studying the psychological properties people care about.

Psychologists in the psycholexical tradition don’t stop there, though. They also typically argue that the *semantic structure* of a language reflects to some extent the *perceived structure of the phenomena* described by the language. In personality psychology, this insight was used by Allport and Odbert [2] to create a semantic taxonomy of thousands of personality-relevant terms, which they argued represents how people conceive of personality. Of course, the step from language to people’s conception of personality is not identical to the step from their conception of personality to actual personality, but it’s natural to think that there will be at least a positive correlation – if only a weak one – between how we think about personality and how personality actually is. This two-step connection (from language about personality to conceptions of personality, from conceptions of personality to actual personality) has been empirically validated by personality models such as the Big Five [3] and Big Six [4, 5].

The Big Six includes an H factor that represents facets of personality related to honesty and humility. *Intellectual* humility seems to involve a consciousness of the limits of one’s knowledge, including sensitivity to circumstances in which one’s native egocentrism is likely to function self-deceptively [6], though others regard it as more of a “second-order” open-mindedness [7]. In our age of information, intellectual humility has grown all the more relevant. However, little conceptual or empirical work has explored this trait. We think that the psycholexical approach is especially promising in the investigation of intellectual humility because questionnaires are likely to be especially unreliable as measures of this construct. Someone who is genuinely humble is unlikely to report being humble, and someone who reports being humble is unlikely to be humble. Humility – whether intellectual, moral, or otherwise – seems to involve a paradox of self-attribution.

Additionally, our investigation is motivated by Aristotle’s [8] insight, reiterated in contemporary philosophy by Roberts and Wood [6], that a virtue (i.e., a positive value-laden personality disposition or dimension of individual difference) is often best understood in the context of related virtues and the vices they oppose. Put a different way, by contextualizing a term for a virtue in the constellation of its near-synonyms and its near-antonyms, we can create a perspicuous

representation of the meaning of the term. For these reasons, we propose to investigate the trait of intellectual humility psycholexically by comparing ‘intellectual humility’ with both its antonyms and synonyms.

2 Method

Our analysis is based on the assumption that the practice of language is precipitated in dictionaries, lexicons, and other wordbooks. Of particular interest is the thesaurus – a language reference book or database organized to help its users find words related to a concept but having slightly different shades of meaning or connotation. Thesauruses reflect what people in their daily use of language – in particular when writing text – consider semantically similar to a given term. In other words, a thesaurus lists synonyms in a broad sense. Modern thesauruses also list antonyms, which are then again related to a set of their own synonyms.

The present research explores the semantic space of intellectual humility by first identifying the most common synonyms and antonyms of ‘intellectual humility’. Next, by referring to the thesaurus.com database (the largest online thesaurus for American English), we associate each identified term with a *word-bag*, which is the set of synonyms listed for that term. The *semantic constellation* of a term t is thus an ordered pair (t, T) , whose first element is t itself and whose second element is t ’s word-bag $T = \{t, t_{\text{syn}1}, t_{\text{syn}2}, t_{\text{syn}3}, \dots, t_{\text{syn}n}\}$, i.e., the set of synonyms of t (including t itself). By comparing semantic constellations of two terms t^1 and t^2 , we then create a similarity metric S by calculating the relative overlap of each pair of word-bags as follows:

$$S(t^1, t^2) = \frac{|T^1 \cap T^2|}{\min\{|T^1|, |T^2|\}} \quad (1)$$

The similarities calculated in this way are then used in a novel clustering and visualization tool that generates a semantic map of the terms involved. More specifically:

1. We identified potential synonyms and antonyms for ‘intellectual humility’ in three ways:
 - a) We searched philosophy and psychology journals for articles that discuss intellectual humility; we found 24 papers or related texts (such as calls for proposals or abstracts).
 - b) We performed an Internet search for entries on ‘intellectual humility’ and found 20 entries that dealt in a significant way with the concept.
 - c) We identified scales that are used in psychology for constructs that have some similarity to intellectual humility (e.g., the H factor of the Big Six personality inventory).

In all these texts, we identified terms that are used to represent the meaning of ‘intellectual humility’ or its relevant vices.

2. Four raters that have experience with the philosophical topic of intellectual humility assessed all terms collected in step 1 to determine whether they

could be used to express the concept of intellectual humility or a related vice. A term was kept on the list if three out of four raters agreed to do so. In this way, we identified 52 synonyms and 69 antonyms for ‘intellectual humility’. Each term was represented at least in noun form and usually in adjective form also, for example: tolerance, tolerant.

3. We identified all entries for each term generated in step 2 in the thesaurus.com database to generate word-bags for each synonym and antonym. For example, the word-bag for ‘tolerance’ included all entries on thesaurus.com for the term set tolerance, tolerant.
4. Next, we calculated the similarity in overlap between every pairwise combination of word-bags. For example, the word-bag of ‘tolerance’ contains 55 terms and the word-bag of ‘broadmindedness’ contains 40 terms. 12 terms are contained in both word-bags. Hence, the similarity between ‘tolerance’ and ‘broadmindedness’ is $12/40 = 0.3$. In this way, the similarity measures are always between 0 (no similarity) and 1 (one word-bag is completely contained in the other word-bag).
5. We checked for highly similar terms ($S > 0.5$).⁴ We collapsed the word-bags of these terms into a single word-bag to reduce the number of synonyms/antonyms. Conceptually, it’s unclear whether terms that share more than half of their semantic constellations represent genuinely distinct constructs. In this way, we reduced the number of synonyms from 52 to 39 and the number of antonyms from 69 to 46. When two terms were collapsed, our raters kept the term that in their estimation was better known. A new word-bag was created combining those of the two collapsed terms (and eliminating redundancies). In cases where the word-bag of a term t overlapped with two or more other terms by > 0.5 whose mutual overlap was, however, below the cutoff-value, the raters determined collapsing based on the highest mutual overlaps. This occurred 2 times for the synonyms and 8 times for of the antonyms. For all condensed word-bags, the similarities were re-calculated. Step 5 was not iterated.
6. The similarity measures obtained in this way were then used as inputs in a visualization algorithm called *superparamagnetic agent mapping*, which employs self-organizing agents governed by the dynamics of a clustering algorithm inspired by spin physics to generate denoised low-dimensional representations for which the characteristics of nonlinear data structures are preserved or even emphasized. To conceptualize this mapping, imagine each term as a particle that naturally repels all other particles. However, as overlap between two terms increases, they become more attracted to each other. Thus, superparamagnetic agent mapping typically produces clumping, where several particles clump together (connoting similarity) while collectively repelling a different cluster (connoting collective difference between the two clusters). It has been shown [9] that this method is superior to standard

⁴ This cut-off value was chosen based on a logarithmic count of the long-tailed distance distribution such that the tail was cut off before the beginning of the main mode of the distribution (i.e., the largest mode in a multi-modal distribution).

methods such as factor analysis, principal components analysis, and multidimensional scaling in preserving the topology of the data space with clustered data. Since such a map will never precisely display the real topology of the original, high-dimensional space, we calculated for each point on the map the sum of the differences between the point and all its neighbors both in the map and in the original space (normalized to the longest distance in either case). The lower this sum, the better the map displays the real distance distribution of a point from its neighbors in the original space, so this number is a proxy for the quality of the map.⁵ To increase the heuristic value of the maps, we rescaled the sizes of the points themselves so that larger points indicate greater topological certainty.

7. Finally, using the same clustering paradigm in an adapted version from [10], we identified clusters on the map generated in step 6.

Step 7 generates the maps below that are then used to inform our reasoning about intellectual humility.

3 Results

We produced three maps to convey our results. Figure 1 is the synonym map, showing the degree of overlap among intellectual humility’s 39 synonyms. The terms predominantly cluster into three groups. The first group (displayed in green) we have labeled *the Sensible Self* and is exemplified by terms such as ‘comprehension,’ ‘responsiveness,’ and ‘mindfulness’. We take this cluster to be representative of the notion that an intellectually humble person will be open and responsive to new ideas and information. Interestingly, ‘sympathy’ is also in this cluster, indicating that there at least can be an emotional component to this sort of intellectual humility. The second (pink) cluster we call *the Inquisitive Self*; it is illustrated by terms such as ‘curiosity,’ ‘exploration,’ and ‘learning’. The difference between the Sensible Self and the Inquisitive Self indicates that there is some difference between seeking new information or ideas and being open to them when they are presented. Third, we have named the blue cluster *the Discreet Self*, which is typified by ‘humility,’ ‘decency,’ and ‘unpretentiousness’. It is worth noting that the discreet cluster is further removed from the other two than the sensible and inquisitive clusters are from each other and that this part of the map most accurately displays the original topology of the high dimensional space. This would seem to indicate that intuitiveness and sensibility are more closely conceptually related to each other than either is to discreteness. Finally, some terms (shown in black) have intermediate positions among these groups (e.g., ‘flexibility’ and ‘tolerance’) and do not fit neatly within any cluster.

Figure 2 shows the results of the antonym map, displaying the degree of overlap between intellectual humility’s 46 antonyms. The first result to notice is

⁵ This map aberration index, defined as $\sum_i \sum_j \left| \frac{d_{orig}(x_i, x_j)}{\max_j \{d_{orig}(x_i, x_j)\}} - \frac{d_{map}(x_i, x_j)}{\max_j \{d_{map}(x_i, x_j)\}} \right|$, also serves as quality measure when comparing different types of visualizations like MDA or PCA; see [9] for further details.

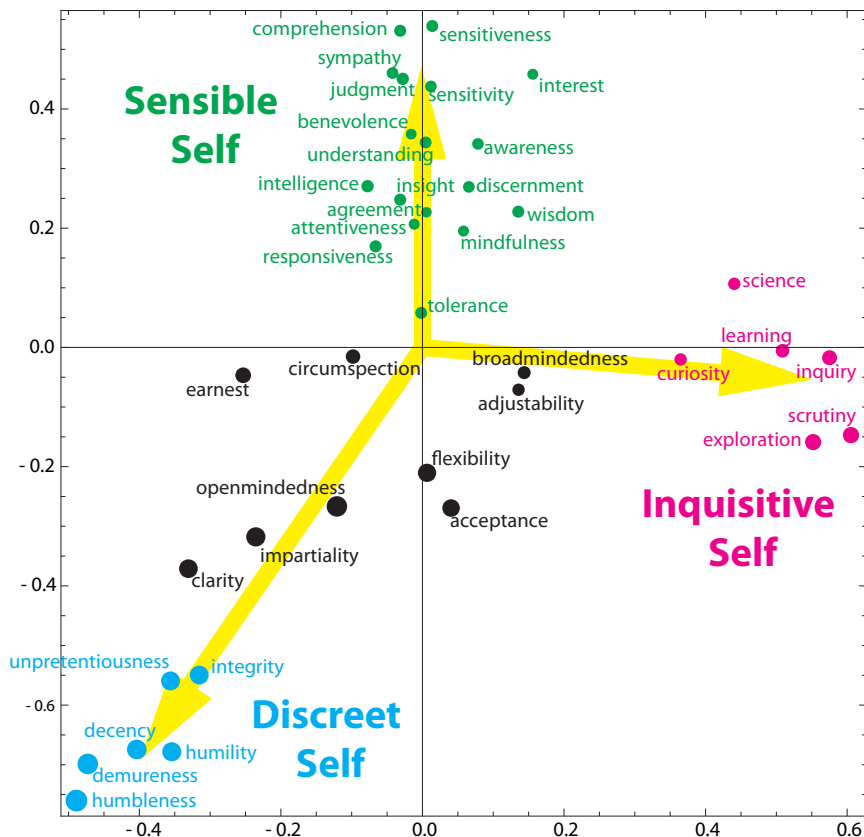


Fig. 1. IH Synonym map.

that almost all the terms are aligned along one dimension and cluster at each endpoint. We take this to represent the distinction between underrating and overrating. The larger, red cluster can be thought of as *the Overrated Self*, and includes terms such as ‘vanity’, ‘pride’, and ‘arrogance’. This cluster suggests that one way not to be intellectually humble is to be overly focused on one’s own high status. Overrating oneself is not, however, the only way to fail to be intellectually humble. The opposite endpoint has two closely related clusters that indicate two other ways. There is *the Underrated Other* in purple (typified by terms such as ‘bias’, ‘prejudice’, and ‘unfairness’) and *the Underrated Self* cluster in orange, which is similar in that it involves underrating, but the object of underrating is oneself. This cluster is characterized by terms such as ‘diffidence’, ‘timidity’, and ‘acquiescence’. This cluster suggests that there is such a thing as being too humble, such that one’s lack of pride ceases to have any positive value. It is worth noting how close the two (orange and purple) underrated clusters are relative to the (red) overrated cluster. This indicates that there is a higher degree of similarity based on the nature of the rating (over or under) than on

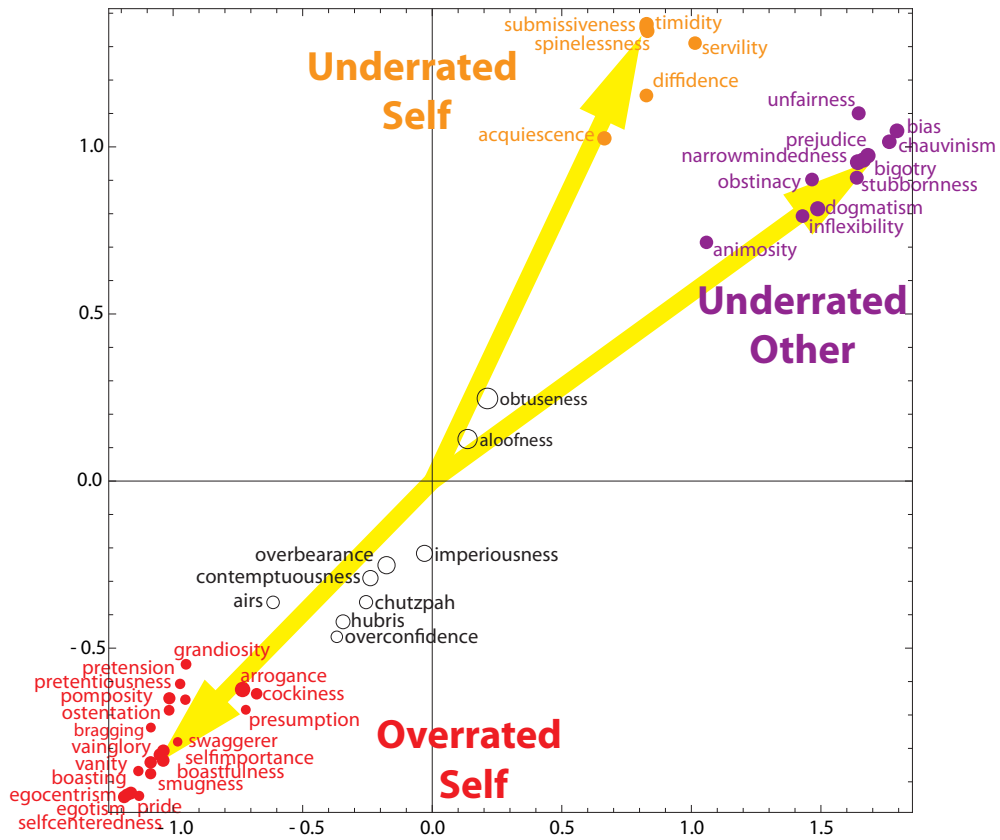


Fig. 2. IH Antonym map.

who is being evaluated (self or other). Note also that there is no overrated other cluster – which may either indicate that this disposition is conceptually unrelated to intellectual humility, or simply that our methodology didn’t uncover terms related to this disposition. Finally, we again see several terms (such as ‘hubris’, ‘chutzpah’, and ‘aloofness’) in white circles in the middle, indicating that these terms do not fit within any cluster.

Finally, in Figure 3, we mapped all synonyms and antonyms together. We preserved the colors from the two previous maps. The resulting map retains many of the structural features of the previous maps, but with a few significant changes. First, it reveals that for the antonyms the linear structure along the poles of the Overrated Self and the Underrated Other is mainly preserved, whereas the terms on the Underrated Self (orange) are in the same region as the terms for the Discreet Self (blue) from the synonym set. Additionally, the distinction between the terms for the Sensible Self (green) and Inquisitive Self (pink) is no longer discernible. This second merger merely indicates that the difference between the Inquisitive Self and the Sensible Self is large enough to be significant

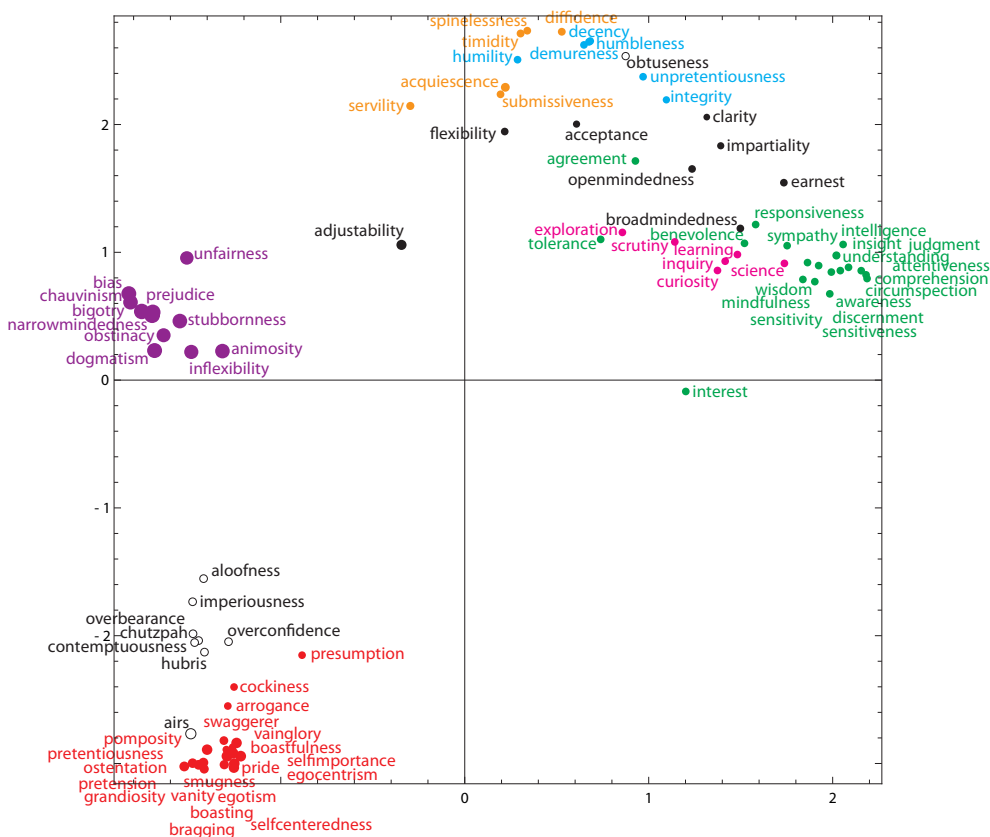


Fig. 3. Unified synonyms and antonyms map.

when compared to the Discreet Self, but small enough not to be significant when compared to intellectual humility’s antonyms.

4 Discussion and Conclusion

From these results, there are three points we wish to draw out for discussion. First, there is the matter of what the clusters represent. In the antonyms map, we take each cluster to represent a distinct vice, i.e., a different way one can fail to be intellectually humble. For the synonyms, however, two possibilities exist. It might be that each cluster represents a distinct trait, all three of which go by the same name of ‘intellectual humility’. Opposing this semantic diversity thesis is the alternate interpretation that sees each cluster representing a different facet of the single trait of intellectual humility.

Second, consider the merging of the synonym-based Discreet Self and antonym-based Underrated Self in the combined map. We see two possible interpretations.

It might be that the discreet aspect of intellectual humility is essentially akin to underrating oneself. Snow [11] and Taylor [12] both argue that humility essentially involves recognizing one's low status or personal faults. If this is right, then either the discreet aspect of humility is more of a vice than a virtue, or the underrated aspect of humility's antonyms is more of a virtue than a vice. Either way, the valence of one or both of these semantic clusters may need to change. Alternatively, there might be two different traits picked out by these clusters – one a virtue and the other a vice – that are behaviorally similar enough that they are easily conflated. Someone who underrates herself will behave very similarly to a discreet person. They will both not regularly speak up about controversial topics, in praise of themselves, or for their own rights and entitlements, making it difficult to differentiate them behaviorally. There could, however, be an underlying psychological difference that typically goes unobserved. The discreet person may not often attend to evaluating herself, but when she does so, she does it accurately. One who underrates herself, however, may pay significant attention to her own merits, but regularly devalue them. Further research on the behavioral and psychological aspects of intellectual humility and its contraries may help to answer this question.

The third point relates back to the Big Six personality inventory [4, 5]. As mentioned, the H factor is meant to represent facets of personality related to honesty and humility. The 100-item revised version measures the participant's humility (specifically her modesty) by having her indicate (dis)agreement with statements such as "I am an ordinary person who is no better than others." We worry that the Big Six therefore includes in its H dimension items that are better understood as contrary to humility, not allied with or constitutive of it.

Finally, we remind three shortcomings of our study. First, the nature of the data does not allow to decide, whether the similarities found only reflect purely pragmatic use of words within (American) English, or whether they actually represent relations between mental concepts. This requires additional empirical research that is currently in preparation. Second, the similarity measure used may misguide the analysis as we do not take into account differences in usage frequency of terms. For example, two pairs of terms may share an overlapping set of the same relative size – but in one case the overlap may consist of terms that are very frequently used as synonyms for the terms under investigation, whereas in the other case, the word bags share synonyms that are only very rarely used. It would be plausible to treat these two cases differently – however, this would again need additional empirical data on how frequent certain terms are used as synonyms. To diminish this problem, services like Google Ngram (<https://books.google.com/ngrams>) could be used for obtaining proxy data regarding term frequency in synonymy relations. Third, our current methodology requires manual coding for term collapsing, as there is no biunique way of merging word bags. Although manual coding was feasible in our case, for larger term sets this approach may reach its limits. However, in our case, manual coding was mainly used to decrease the number of synonyms/antonyms and to allow for visualizations that are easier to understand. One may skip this part to the cost

that the resulting maps are harder to interpret. In ongoing research where we aim for inter-lingual comparison (English and German), we skipped manual coding and found, that the qualitative results remained but that a richer structure was visible, which helps to better understand differences across languages.

Acknowledgments. We thank Daniel Lapsley, ACE Collegiate Professor and Chair of the Department of Psychology of the University of Notre Dame, and Paul C. Stey, Department of Psychology of the University of Notre Dame, for their input to this research. This work was made possible through a grant from the Thrive Center at Fuller Theological Seminary.

References

1. Galton, F.: *Measurement of Character* (1884)
2. Allport, G., Odbert, H.: *Trait-names: A Psycho-lexical Study* (1936)
3. Peabody, D., Goldberg, L.: Some determinants of factor structures from personality-trait descriptors. *Journal of Personality and Social Psychology* 57(3), 552–567 (1989)
4. Ashton, M., Lee, K., Perugini, M., Szarota, P., de Vries, R., Di Blas, L., Boies, K., De Raad, B.: A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology* 86(2), 356–366 (2004)
5. Saucier, G.: Effects of variable selection on the factor structure of person descriptors. *Journal of Personality and Social Psychology* 73(6), 1298–1312 (1997)
6. Roberts, R., Wood, J.: *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford: Oxford University Press (2007)
7. Spiegel, J.G.: Open-mindedness and intellectual humility. *Theory and Research in Education* 10, 27–38 (2012)
8. Aristotle. *Nicomachean Ethics*
9. Ott, T., Eggel, T., Christen, M.: Generating Low-Dimensional Denoised Embeddings of Nonlinear Data with Superparamagnetic Agents. In: *Proceedings of the 2014 International Symposium on Nonlinear Theory and its Applications (NOLTA)*, Lucerne, Switzerland, September 14-18, (2014)
10. Ott, T., Kern, A., Steeb, W.-H., Stoop, R.: Sequential Clustering: Tracking Down the Most Natural Clusters. *Journal of Statistical Mechanics: theory and experiment*, P11014 (2005)
11. Snow, N.: Humility. *The Journal of Value Inquiry* 29(2), 203–216 (1995)
12. Taylor, G.: *Pride, Shame and Guilt*. Oxford: Clarendon Press (1985)