**Dieses Dokument ist eine Zweitveröffentlichung (Verlagsversion) /
This is a self-archiving document (published version):**

Daniel Leising, Joachim Scharloth, Oliver Lohse, Dustin Wood

What Types of Terms Do People Use When Describing an Individual's Personality?

Erstveröffentlichung in / First published in:

Psychological Science. 2014, 25(9), S. 1787 – 1794 [Zugriff am: 19.08.2019]. SAGE journals. ISSN 1467-9280.

DOI: <https://doi.org/10.1177/0956797614541285>

Diese Version ist verfügbar / This version is available on:

<https://nbn-resolving.org/urn:nbn:de:bsz:14-qucosa2-353833>

„Dieser Beitrag ist mit Zustimmung des Rechteinhabers aufgrund einer (DFGgeförderten) Allianz- bzw. Nationallizenz frei zugänglich.“

This publication is openly accessible with the permission of the copyright owner. The permission is granted within a nationwide license, supported by the German Research Foundation (abbr. in German DFG).

www.nationallizenzen.de/

What Types of Terms Do People Use When Describing an Individual's Personality?

Psychological Science
2014, Vol. 25(9) 1787–1794
© The Author(s) 2014
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0956797614541285
pss.sagepub.com

**Daniel Leising¹, Joachim Scharloth², Oliver Lohse³, and
Dustin Wood⁴**

¹Department of Psychology, Technische Universität Dresden; ²Applied Linguistics,
Technische Universität Dresden; ³Department of Psychology, University of Zurich; and

⁴Department of Psychology, Wake Forest University

Abstract

An important yet untested assumption within personality psychology is that more important person characteristics are more densely reflected in language. We investigated how ratings of importance and other term properties are associated with one another and with a term's frequency of use. Research participants were asked to provide terms that described individuals they knew, which resulted in a set of 624 adjectives. These terms were independently rated for importance, social desirability, observability, stateness versus traitness, level of abstraction, and base rate. Terms rated as describing more important person characteristics were in fact used more often by the participants in the sample and in a large corpus of online communications (close to 500 million words). More frequently used terms and more positive terms were also rated as being more abstract, more traitlike, and more widely applicable (i.e., having a greater base rate). We discuss the implications of these findings with regard to person perception in general.

Keywords

social cognition, social perception, psycholinguistics, personality

Received 4/3/14; Revision accepted 6/3/14

Psychologists have been interested in natural person-descriptive language for decades (Allport & Odbert, 1936; Goldberg, 1982, 1993; John, 1990; John, Angleitner, & Ostendorf, 1988; Norman, 1967; Tupes & Christal, 1961). By studying such language, psychologists hope to learn something about the mechanisms of person perception and personality structure. For instance, the *psycholexical hypothesis*, which holds that more important person characteristics are more densely reflected in language (Goldberg, 1993; Saucier & Goldberg, 2001), underlies the development of lexical trait structures such as the Big Five (Goldberg, 1993). Researchers have also fruitfully examined how basic properties of person descriptors such as observability, breadth, and social desirability are predictive of between-trait differences in interrater agreement, accuracy, and longitudinal stability (Edwards, 1953; Funder & Dobroth, 1987; John & Robins, 1993; Paunonen, 1989; Wood & Wortman, 2012).

The current study contributes to this literature as follows. Most previous studies of trait properties used sets of person-descriptive words that were compiled by very

small numbers of researchers using relatively idiosyncratic criteria (Peabody, 1987), and these researchers did not have the technical tools that are available to researchers today (e.g., Roivainen, 2013). In the present study, to improve on representativeness, we used a free-response format for generating a set of adjectives that laypeople, not researchers, considered appropriate descriptors of personality. We then asked participants to rate these adjectives for some well-established properties (e.g., observability) and some previously overlooked properties (e.g., traitness) and used these properties to predict language use in the same participant sample and—more important—in an extremely large corpus of online communications. Doing so enabled us to answer several research questions of theoretical importance.

Corresponding Author:

Daniel Leising, Department of Psychology, Technische Universität
Dresden, 01062 Dresden, Germany
E-mail: daniel.leising@tu-dresden.de

First, do people spontaneously tend to use person descriptors that refer to the things that are particularly important to know about someone (cf. Wood, 2014)? For example, whether someone is insidious or trustworthy might be seen as more important to know than whether someone is articulate or superstitious. The assumption that there are systematic differences between traits in this regard, and that these in part drive their density in language, lies at the core of the psycholexical hypothesis (e.g., Klages, 1926/1932, as noted by Saucier & Goldberg, 2001): "The degree of representation of an attribute in language has some correspondence with the general importance of the attribute" (p. 849). This hypothesis is a crucial conceptual foundation for the predominant personality taxonomy in mainstream psychology (i.e., the Big Five), but we are not aware of any empirical evidence that directly supports it. In the present study, we aimed to provide such evidence by investigating whether the frequency of term use in actual person descriptions may be predicted from ratings of trait importance.

Second, is it possible to clearly distinguish between state words (which describe the behavior of people at relatively specific instances) and trait words (which describe people's enduring behavioral tendencies over time)? The assumption that such a distinction is possible is also of great importance to the psycholexical tradition, because the samples of person-descriptive words used by such influential authors as Allport and Odbert (1936) and Norman (1967) consisted predominantly of words that these researchers considered to refer to traits, not states. Whether such a dichotomy actually exists, however, has never been tested, as far as we know. In the present study, we conducted such a test.

Third, do people prefer broad terms over narrow terms in describing themselves and other people? There is some evidence to this effect (John, Hampson, & Goldberg, 1991). For example, the term *nice* refers to a much greater range of behaviors than does the term *polite*. If people actually prefer to use the former kind of term over the latter, this would be theoretically significant because it would indicate a preference for descriptions that maximize representativeness, probably at the cost of specificity. There are several conceptually distinct types of broadness, however, and only some of these have been considered in previous research. Besides the level of abstractness (i.e., the number of specific behaviors that may be subsumed under a term; Hampson, John, & Goldberg, 1986), one may investigate base rate (i.e., the number of people to which a term is thought to apply) as well as the extent to which the term refers to a trait as opposed to a state (i.e., the proportion of a given target's behavior for which a term accounts), as noted earlier.

The situation is complicated further by evidence showing, for example, that terms that are more abstract are

judged as more positive (Hampson, Goldberg, & John, 1987) and that positive terms are generally used more often than negative terms (Augustine, Mehl, & Larsen, 2011; Boucher & Osgood, 1969). Obviously, the dimensions along which person-descriptive terms from the natural language may be distinguished from each other are interrelated. Thus, to obtain the most complete and accurate picture, it is important to assess all of these dimensions at once and to investigate both their relations with each other and their capacity to independently predict word-use frequency. This is what we did in the present study.

Method

Sample of terms

We used the complete adjective subsample ($n = 624$) of the set of terms ($N = 758$) generated by a convenience sample of 168 research participants (115 female, 53 male; mean age = 24.9 years, $SD = 6.2$) from a study by Leising, Ostrovski, and Borkenau (2012). The participants in that study were asked to come up with a number of terms (3–10) that they thought would best describe themselves and four other people: one person whom they liked and knew well, one whom they liked but did not know well, one whom they knew well but did not like, and one whom they neither knew well nor liked. Apart from these boundary conditions, the participants were free to choose whomever they wanted as targets.

Such free-response designs have been used only rarely in the past (e.g., Allen & Potkay, 1973; Chaplin & John, as cited in John, 1990; Church & Katigbak, 1989; Donahue, 1994; Kohnstamm, Halverson, Mervielde, & Havill, 1998). However, their distinct advantage is that they lead to samples of terms that people actually use in spontaneous descriptions of themselves and others, thereby improving on ecological validity compared with studies in which samples of terms are selected from dictionaries or by small teams of researchers. We used only adjectives to ensure maximum comparability between terms in the corpus analyses that we conducted.

Property ratings of terms

The 624 adjectives were presented to four different teams of student raters: One group of 20 raters judged each term's *social desirability* (i.e., how positive or negative an evaluation of a target person the use of a term implies), *observability* (i.e., how easy it is for an observer to see whether a person possesses the characteristic), and *trait-ness* (i.e., the extent to which the term denotes a characteristic that is stable over time, as opposed to changing from situation to situation). Observability was included to

ensure comparability with previous studies, because it is one of the most extensively investigated term properties (e.g., Funder & Dobroth, 1987; John & Robins, 1993). Another team of 17 raters judged each term's *importance* (i.e., how important it would be for them to know whether a person has the characteristic). Yet another team of 17 raters judged each term's *abstractness* (i.e., how many different behaviors could be subsumed under the term; Hampson et al., 1986). Finally, a team of 16 raters judged each term's *base rate* (i.e., whether a term could appropriately be applied to few people or to many). The raters worked independently. All ratings were provided on a scale ranging from 1 to 10 (1 = *low*, 10 = *high*).

Frequency of term use

As a measure of term-use frequency, we calculated the number of participants in the original sample (Leising et al., 2012) who had used a term to describe at least one of the five targets. For example, if Pete used the term *freundlich* ("friendly") to describe two targets, Anna used the same term to describe one target, and no other participant used the term to describe any target, then the frequency count for *freundlich* was 2 (i.e., we counted the number of participants, not the number of uses). We also obtained another estimate of word-use frequency by calculating how often each adjective directly preceded the words *Mann* ("man"), *Frau* ("woman"), *Person* ("person"), *Persönlichkeit* ("personality"), *Mensch* ("human being"), *Individuum* ("individual"), and *Typ* ("type") in a large corpus derived from four different online communication Web sites (seniorentreff.de, old forum: > 25 million words; seniorentreff.de, new forum: > 57 million words; bfriends.britte.de: > 141 million words; politikforen.net: > 263 million words). The overall corpus comprised more than 488 million words. Of the 624 adjectives generated by the research participants in the study by Leising et al. (2012), 112 (17.9%) did not appear in the corpus in any of these contexts.

Results

Term properties: reliability, distributions, and interrelations

Table 1 displays the interrater reliabilities for the property ratings of the terms, as well as the means and standard deviations of those ratings. Reliabilities were high for all the term properties (all α s $\geq .84$), which indicates that the orderings on these properties (e.g., which terms were rated as most or least observable or abstract) should be expected to be highly similar to the orderings obtained in new samples of raters. Table 1 also provides the seven

terms with the highest and lowest average ratings on each of the dimensions, to illustrate their content. For the sake of simplicity, we present only the English translations of the German terms. The complete list of terms and ratings may be obtained from the first author. Note that a few of the German adjectives were best translated into English nouns or combinations of two or more English words (e.g., *tanzbegeistert* in German translates approximately to "dance enthusiast"). Nevertheless, the original list of German terms comprised only single adjectives.

As reported in Leising et al. (2012), the distribution of social desirability ratings was clearly bimodal (cf. Anderson, 1968; Dumas, Johnson, & Lynch, 2002). In contrast, the other five rating dimensions showed very symmetric and unimodal distributions. The finding of a unimodal and symmetric distribution is particularly intriguing with regard to the dimension of traitness (see Fig. 1), because in many psycholexical studies, it has been assumed that terms denote either states or traits. However, our data clearly show that the extent to which terms denote stable versus nonstable person characteristics varies continuously and that most person descriptors from the natural language fall somewhere in the middle on the continuum (e.g., "cool," "tacit," "impatient," and "funny" all had mean traitness ratings of 5.5).

We first investigated associations between term properties by inspecting scatter plots. All of the 15 possible associations ($6 \times 5/2$) were examined. There was only one case in which the association (if there was one) was not obviously linear: The terms rated highest and lowest for desirability were rated as more important than terms with intermediate (neutral) desirability. This curvilinear relationship is displayed in Figure 2. The correlation between rated importance and *evaluateness* (i.e., the absolute difference between a term's average desirability rating and the scale mean; John & Robins, 1993) was significant, $r(621) = .63, p < .001$; the rated importance of a term increased as its positive or negative evaluative connotations increased. In all other analyses of associations between term properties, quadratic components could easily be ignored.

Table 2 displays the zero-order correlations between the six term properties (along with 95% confidence intervals) as well as partial correlations between the properties (i.e., pairwise correlations between rating dimensions when holding the other four dimensions constant). We focus here on associations that remained significant when controlling for the other dimensions: Terms that were judged to imply more positive evaluations (desirability) were also judged to refer to more stable person characteristics (traitness), partial $r = .41$; to refer to a greater range of behaviors (abstractness), partial $r = .23$; and to be applicable to a greater number of persons (base rate),

Table 1. Interrater Reliability, Means, and Standard Deviations for the Properties Rated, and Words With High and Low Average Ratings

Property	Reliability ^a	Mean	SD	Terms with the highest average ratings	Terms with the lowest average ratings
Desirability	.99	4.86	2.39	Intelligent, likeable, honest, smart, fun-loving, attractive, authentic	Antisocial, hate-filled, inhumane, malicious, dishonest, despicable, misogynistic
Observability	.89	5.54	1.12	Attractive, extraverted, sporty, macho, charismatic, super-friendly, aggressive	Deep, unfaithful, semieducated, false, without prospects, disingenuous, despicable
Importance	.84	5.35	1.18	Honest, authentic, trusting, trustworthy, sincere, likeable, tolerant	Scatterbrained, parent-oriented, dance enthusiast, footloose, cute, prosaic, messy
Traitiness	.88	5.77	1.16	Extraverted, intelligent, musical, animal-loving, introverted, fond of children, family guy	Desperate, bored, annoyed, uninformed, excited, disinterested, happy
Abstraction	.87	4.87	1.22	Different, simple, false, active, odd, difficult, correct	Drug-addicted, animal-loving, dance enthusiast, career-obsessed, parent-oriented, fond of children, musical
Base rate	.87	4.45	1.08	Humane, nice, vulnerable, sociable, life-affirming, productive, in need of love	Hate-filled, drug-addicted, inhumane, misanthropic, unprejudiced, coldhearted, despicable

Note: Terms with the highest average ratings are listed in order beginning with the highest rated term; those with the lowest average ratings are listed beginning with the lowest-rated term.

^aOur index of reliability is the intraclass correlation coefficient, ICC(2,*k*).

partial $r = .47$. Thus, more positive terms were broader in various ways: They were seen as pertaining to more enduring tendencies of a person, to more different types of behavior, and to more people.

There was a small but significant negative association, partial $r = -.10$, between importance and observability; the raters considered those person characteristics that were less directly observable to be somewhat more valuable to know about other people ("It's the inner qualities that matter"). A term's rated level of importance also

showed small positive associations with the degree to which it was considered traitlike (partial $r = .11$) and with the number of people to which it applied (partial $r = .10$). Finally, when controlling for all other rating dimensions, there was a significant negative association, partial $r = -.20$, between traitiness and level of abstraction, probably reflecting a sort of bandwidth-fidelity dilemma (i.e., assessing the stability of a given person characteristic may require a relatively precise conceptualization of that characteristic).

Fig. 1. Frequency distribution of the traitiness ratings of the 624 adjectives. The rating scale ranged from 1 (*low*) to 10 (*high*). **Fig. 2.** Scatter plot (with best-fitting regression line) showing the relationship between the rated desirability and importance of the terms used in the study.

Table 2. Correlations of Ratings of the Six Term Properties

Rating	Desirability		Observability		Importance		Traitness		Abstraction		Base rate	
	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>
Desirability			.04	.00	.07	-.04	.47*	.41*	.24*	.23*	.56*	.47*
			[-.04, .12]		[-.01, .15]		[.41, .53]		[.17, .31]		[.50, .61]	
Observability					-.09*	-.10*	.04	.03	.00	-.01	.05	.04
					[-.17, -.01]		[-.04, .12]		[-.08, .08]		[-.03, .13]	
Importance							.13*	.11*	.00	.00	.13*	.10*
							[.05, .21]		[-.08, .08]		[.05, .21]	
Traitness									-.06	-.20*	.31*	.06
									[-.14, .02]		[.24, .38]	
Abstraction											.18*	.07
											[.10, .26]	

Note: Values in brackets are 95% confidence intervals.

* $p < .05$.

Predicting word-use frequencies

We determined how often each adjective was used by participants in the study by Leising et al. (2012) and in the corpus. As is common in linguistic research, both word-frequency variables had an extreme positive skew; that is, most words in the sample were used very infrequently and few were used very frequently. To deal with the extreme skew of both word-usage variables, we log-transformed these data. After this transformation, the two variables showed a moderate positive correlation, $r(510) = .42$, indicating that terms that were spontaneously used by more research participants were also used more often to describe people within the corpus.

Table 3 displays the associations we found between ratings of term properties and the frequencies with which the terms were used. To determine the overall power of the rating variables as predictors of word-use frequency, we entered them into multiple regressions using the stepwise algorithm in IBM SPSS Statistics (Version 22). The model for word-use frequency among the participants in the study by Leising et al. (2012) yielded an impressive multiple correlation (R) of .52, $F(5, 617) = 45.84$, $p < .001$. Five property ratings were found to make independent contributions in the prediction of word-use frequency:

importance ($\beta = 0.28$), traitness ($\beta = 0.14$), abstraction ($\beta = 0.16$), base rate ($\beta = 0.22$), and observability ($\beta = 0.23$). Similar results were obtained for word-use frequency in the corpus, $R = .49$, $F(4, 506) = 39.95$, $p < .001$. In this model, significant independent contributions were made by importance ($\beta = 0.12$), traitness ($\beta = 0.15$), abstraction ($\beta = 0.41$), and base rate ($\beta = 0.11$).

Thus, across both measures of word-use frequency, terms that were used more often were deemed to denote more important characteristics that also are more stable (traitness), apply to more people (base rate), and refer to a wider range of behaviors (abstraction). All of these term properties could predict a term's frequency of use independently of one another. Observability was a significant independent predictor of word use only in the participant sample, not in the corpus.

Discussion

In the present study, we used a large natural-language sample of person-descriptive adjectives that was created by asking research participants to describe targets in a free-response format. By doing so, we aimed to overcome an important limitation of previous psycholexical studies, in which samples of terms were usually compiled by

Table 3. Correlations of Property Ratings With Word-Use Frequency

Sample	Desirability		Observability		Importance		Traitness		Abstraction		Base rate	
	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>	<i>r</i>	Partial <i>r</i>
Participants	.21*	-.07	.23*	.26*	.30*	.30*	.24*	.16*	.19*	.19*	.33*	.24*
	[.13, .28]		[.16, .30]		[.23, .37]		[.17, .31]		[.11, .26]		[.26, .40]	
Online corpus	.27*	.07	.05	.06	.14*	.14*	.18*	.12*	.42*	.40*	.25*	.07
	[.19, .35]		[-.04, .14]		[.05, .22]		[.10, .26]		[.35, .49]		[.17, .33]	

Note: Values in brackets are 95% confidence intervals.

* $p < .05$.

small numbers of researchers. As a consequence of the strategy we adopted, our sample of terms may be regarded as being quite representative of words that people in Germany actually use when describing someone's personality. Using this sample of terms, we investigated associations between several dimensions that systematically distinguish person descriptors, as well as associations between these dimensions and frequency of use. We now discuss our findings in more detail.

First, terms that were deemed more important were used more frequently. This finding corroborates an important aspect of the psycholexical hypothesis: Things that are more important to know about people feature more prominently in language (Saucier & Goldberg, 2001). To our knowledge, the present study is the first to directly support this assumption, which is important because the value of lexically derived trait structures, such as the Big Five framework, hinges on this assumption to a considerable extent.

Second, ratings of the traitness of terms had a unimodal distribution, with a mean of 5.77 (on a scale ranging from 1 to 10). Thus, whether terms denote states or traits is not a binary decision but a matter of degree, and most person characteristics fall somewhere in the middle on that continuum. This finding is relevant for psycholexical research as well, because the sets of terms used by influential authors in this research tradition (Allport & Odbert, 1936; Norman, 1967) were compiled with the notion of a state/trait dichotomy in mind. However, we found no evidence for the existence of such a dichotomy in natural person-descriptive language, which perhaps calls into question the selection strategies applied by some of the forefathers of psycholexical research (cf. Fleenor, 2001).

Third, terms used more frequently were broader in three different respects: They were perceived to apply to more people (base rate), to reflect more stable person characteristics (traitness), and to pertain to a wider range of behaviors (abstraction). All three of these ratings made independent contributions in predicting word-use frequency. The first two associations provide evidence for the validity of our term ratings: Terms that were judged to apply to more persons actually *were* applied to more persons, and terms that were judged to refer to more stable characteristics were actually used more often to describe *persons* (i.e., what makes people recognizable across occasions).

However, the latter association, between frequency of use and abstraction, requires a little more theoretical reasoning: The finding that perceivers prefer terms that refer to a greater range of behaviors is consistent with a previous finding by John et al. (1991), who concluded that "people prefer the highest level of abstraction that is still descriptive of behavior (e.g., *kind*) over more descriptive

subordinate levels (e.g., *charitable* and *generous*) and over an even broader level devoid of descriptive meaning (e.g., *good*)" (p. 348). In fact, the participants in the study by Leising et al. (2012) never described targets using purely evaluative terms such as "good," "great," or "bad." Therefore, our results, drawn from a much stronger database, clearly corroborate the notion presented by John et al. (1991): People seem to prefer bandwidth over fidelity in that they favor person descriptions that maximize generalizability across a broad range of behaviors, at the likely cost of being less able to predict more specific behaviors.

Terms that apply to more people (higher base rate), refer to a greater range of behaviors (greater abstractness), and reflect more stable person characteristics (greater traitness) were also rated as characterizing targets more positively. The associations between base rate and positivity and between abstraction and positivity directly replicate previous findings with samples of English terms (Hampson et al., 1986, 1987), but we are not aware of any previous study that demonstrates an association between traitness and positivity. All three associations persisted when we controlled for the other variables, which suggests that each of the three different variants of term broadness was uniquely associated with positivity. Positivity's association with base rate and traitness may indicate either that most people tend to show behaviors that perceivers (a priori) regard as positive or that perceivers tend to judge (a posteriori) the behaviors that people show most often as positive. Space is too limited here to address these issues in more detail (cf. Biesanz & Human, 2010; Bowles & Gintis, 2011; Wood, Gosling, & Potter, 2007; Wood & Wortman, 2012).

There may also be a fundamental asymmetry in the level of resolution at which positive and negative person characteristics are described (cf. Unkelbach, Fiedler, Bayer, Stegmüller, & Danner, 2008). These three associations between broadness and positivity suggest that specificity is lower for more positive terms: The more positive the terms, the less they seem applicable to *just some* persons, behaviors, or situations. Generally speaking, it may be more necessary or useful for perceivers to specifically identify characteristics that interfere with their personal interests, compared with characteristics that conform to their personal interests (cf. Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). Such a theoretical explanation may help account for all three associations between broadness and positivity.

A limitation of the present study may be that we analyzed only German terms, so the findings remain to be replicated in other languages. However, in a recent study, Wood (2014) was able to corroborate some of the associations among trait properties that we found (e.g., the strong positive association between importance and

evaluativeness), using a sample of English terms. Thus, we are quite confident that the findings of the present study generalize well to languages other than German. Our methodological approach to generating an item sample (free-response format) holds promise for improving on the ecological validity of psycholexical research. This approach is promising because the terms are chosen by large groups of native-language users instead of by small groups of researchers. Furthermore, free response is the way people describe each other in real life: When people talk about themselves and each other, they usually do not complete rating sheets; rather, they choose a few specific terms very quickly from a vast universe of possibilities.

Future research on the factors (e.g., in the situation, in the perceiver, in the target, and in the term) that are involved in making a perceiver choose one term but not another to describe a person may use a combination of a free-response format with prior sampling of a set of potential descriptors (e.g., dictionary sampling). Thus, combining the unique advantages of our present approach with the approach traditionally used in psycholexical research may ultimately result in an even more complete understanding of how person perception actually works in real life.

Author Contributions

D. Leising and D. Wood wrote the manuscript. D. Leising performed the statistical analyses. J. Scharloth performed the corpus analyses. O. Lohse helped with collecting the term ratings.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

References

- Allen, B. P., & Potkay, C. R. (1973). Variability of self-description on a day-to-day basis: Longitudinal use of the adjective generation technique. *Journal of Personality, 41*, 638–652. doi:10.1111/j.1467-6494.1973.tb00115.x
- Allport, G. W., & Odbert, H. S. (1936). Trait names: A psycholexical study. *Psychological Monographs, 47*(1, Whole No. 211). doi:10.1037/h0093360
- Anderson, N. H. (1968). Likableness ratings of 555 personality trait-words. *Journal of Personality and Social Psychology, 9*, 272–279. doi:10.1037/h0025907
- Augustine, A. A., Mehl, M. R., & Larsen, R. J. (2011). A positivity bias in written and spoken English and its moderation by personality and gender. *Social Psychological & Personality Science, 2*, 508–515. doi:10.1177/1948550611399154
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323–370. doi:10.1037/1089-2680.5.4.323
- Biesanz, J. C., & Human, L. J. (2010). The cost of forming more accurate impressions: Accuracy-motivated perceivers see the personality of others more distinctively but less normatively than perceivers without an explicit goal. *Psychological Science, 21*, 589–594. doi:10.1177/0956797610364121
- Boucher, J., & Osgood, C. E. (1969). The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior, 8*, 1–8. doi:10.1016/S0022-5371(69)80002-2
- Bowles, S., & Gintis, H. (2011). *A cooperative species: Human reciprocity and its evolution*. Princeton, NJ: Princeton University Press.
- Church, A. T., & Katigbak, M. S. (1989). Internal, external, and self-report structure of personality in a non-western culture: An investigation of cross-language and cross-cultural generalizability. *Journal of Personality and Social Psychology, 57*, 857–872. doi:10.1037/0022-3514.57.5.857
- Donahue, E. M. (1994). Do children use the Big Five, too? Content and structural form in personality description. *Journal of Personality, 62*, 45–66. doi:10.1111/j.1467-6494.1994.tb00794.x
- Dumas, J. E., Johnson, M., & Lynch, A. M. (2002). Likableness, familiarity, and frequency of 844 person-descriptive words. *Personality and Individual Differences, 32*, 523–531. doi:10.1016/S0191-8869(01)00054-X
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology, 37*, 90–93. doi:10.1037/h0058073
- Fleeson, W. (2001). Toward a structure- and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology, 80*, 1011–1027. doi:10.1037/0022-3514.80.6.1011
- Funder, D. C., & Dobroth, K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology, 52*, 409–418. doi:10.1037/0022-3514.52.2.409
- Goldberg, L. R. (1982). From ace to zombie: Some explorations in the language of personality. In C. D. Spielberger & J. N. Butcher (Eds.), *Advances in personality assessment* (Vol. 1, pp. 203–234). Hillsdale, NJ: Erlbaum.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 26–34. doi:10.1037/0003-066X.48.1.26
- Hampson, S. E., Goldberg, L. R., & John, O. P. (1987). Category breadth and social-desirability values for 573 personality terms. *European Journal of Personality, 1*, 241–258. doi:10.1002/per.2410010405
- Hampson, S. E., John, O. P., & Goldberg, L. R. (1986). Category breadth and hierarchical structure in personality: Studies of asymmetries in judgments of trait implications. *Journal of Personality and Social Psychology, 51*, 37–54. doi:10.1037/0022-3514.51.1.37
- John, O. P. (1990). The “Big Five” factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin (Ed.), *Handbook of personality: Theory and research* (pp. 66–100). New York, NY: Guilford Press.
- John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2*, 171–203. doi:10.1002/per.2410020302

- John, O. P., Hampson, S. E., & Goldberg, L. R. (1991). The basic level in personality-trait hierarchies: Studies of trait use and accessibility in three different contexts. *Journal of Personality and Social Psychology*, 60, 348–360. doi:10.1037/0022-3514.60.3.348
- John, O. P., & Robins, R. W. (1993). Determinants of interjudge agreement on personality traits: The Big Five domains, observability, evaluativeness, and the unique perspective of the self. *Journal of Personality*, 61, 521–551. doi:10.1111/j.1467-6494.1993.tb00781.x
- Klages, L. (1932). *The science of character*. London, England: Allen & Unwin. (Original work published 1926)
- Kohnstamm, G. A., Halverson, C. F., Mervielde, I., & Havill, V. L. (1998). *Parental descriptions of child personality: Developmental antecedents of the Big Five?* Hillsdale, NJ: Erlbaum.
- Leising, D., Ostrovski, O., & Borkenau, P. (2012). Vocabulary for describing disliked persons is more differentiated than vocabulary for describing liked persons. *Journal of Research in Personality*, 46, 393–396. doi:10.1016/j.jrp.2012.03.006
- Norman, T. (1967). *2800 personality trait descriptors: Normative operating characteristics for a university population* (Report UM-08310-1-T). Ann Arbor: University of Michigan, Department of Psychology. Retrieved from <http://files.eric.ed.gov/fulltext/ED014738.pdf>
- Paunonen, S. V. (1989). Consensus in personality judgments: Moderating effects of target-rater acquaintanceship and behavior observability. *Journal of Personality and Social Psychology*, 56, 823–833. doi:10.1037/0022-3514.56.5.823
- Peabody, D. (1987). Selecting representative trait adjectives. *Journal of Personality and Social Psychology*, 52, 59–71. doi:10.1037/0022-3514.52.1.59
- Roivainen, E. (2013). Frequency of the use of English personality adjectives: Implications for personality theory. *Journal of Research in Personality*, 47, 417–420. doi:10.1016/j.jrp.2013.04.004
- Saucier, G., & Goldberg, L. R. (2001). Lexical studies of indigenous personality factors: Premises, products, and prospects. *Journal of Personality*, 69, 847–879. doi:10.1111/1467-6494.696167
- Tupes, E. C., & Christal, R. C. (1961). *Recurrent personality factors based on trait ratings* (U.S. Air Force Aeronautical Systems Division Technical Report No. 61-97). Lackland Air Force Base, TX: U.S. Air Force.
- Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: The density hypothesis. *Journal of Personality and Social Psychology*, 95, 36–49. doi:10.1037/0022-3514.95.1.36
- Wood, D. (2014). *Testing the psycholexical hypothesis: Are more important traits more densely reflected in the English lexicon?* Manuscript submitted for publication.
- Wood, D., Gosling, S. D., & Potter, J. (2007). Normality evaluations and their relation to personality traits and well-being. *Journal of Personality and Social Psychology*, 93, 861–879. doi:10.1037/0022-3514.93.5.861
- Wood, D., & Wortman, J. (2012). Trait means and desirabilities as artifactual and real sources of differential stability of personality traits. *Journal of Personality*, 80, 665–701. doi:10.1111/j.1467-6494.2011.00740.x