Delft University of Technology

Identifying Virtues and Values Through Obituary Data-Mining

Alfano, Mark; Higgins, Andrew; Levernier, Jacob

DOI
10.1007/s10790-017-9602-0

Publication date
2017

Document Version
Final published version

Published in
The Journal of Value Inquiry

Citation (APA)
Alfano, M., Higgins, A., & Levernier, J. (2017). Identifying Virtues and Values Through Obituary Data-Mining. *The Journal of Value Inquiry*, 1-21. https://doi.org/10.1007/s10790-017-9602-0

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

J Value Inquiry
DOI 10.1007/s10790-017-9602-0

Identifying Virtues and Values Through Obituary Data-Mining

Mark Alfano[1,2] · Andrew Higgins[3] · Jacob Levernier[4]

© The Author(s) 2017. This article is an open access publication

1 Introduction

"Call no one happy until they are dead." "Never speak ill of the dead." If we still heed the injunctions of Solon and Chilon of Sparta, then obituaries, which represent a prominent way of expressing the human universal of grief, are a resource for philosophical anthropology. Philosophers have emphasized that we can determine what counts as a virtue for a given type of person in a given cultural context by analyzing what people say about the dead [1]. Such judgments summarize the deceased's life and place it in a meaningful narrative context. This is why practitioners of Acceptance and Commitment Therapy ask patients to write their own obituaries: doing so helps people clarify what they value, how they want to live, and what their lives mean [2]. In popular culture, this attitude has recently cropped in the Black Obituary project, which was founded by Ja'han Jones in 2016 to enable Black Americans to write their own obituaries in case they become the victims of lethal police violence. He characterizes the texts as "agonizingly introspective works," saying, "These are Black lives, as told by those who experience them in all their glory—tragedy—frustration—triumph."[1] The ability —

[1] URL = < http://www.blackobituaryproject.com/>. Accessed 3 October 2016.

✉ Mark Alfano
mark.alfano@gmail.com

[1] Delft University of Technology, Delft, The Netherlands

[2] Australian Catholic University, Melbourne, VIC, Australia

[3] Illinois State University, Normal, IL, USA

[4] University of Pennsylvania, Philadelphia, PA, USA

Published online: 08 June 2017

one might even say the right and the duty — to tell the story of someone's life, whether that life is one's own or another's, is a capacity we dare not neglect as self-presenting creatures. Moreover, praising the dead is a way of performatively identifying with them, which is why Aristotle said of Plato after his death that he was a "man whom bad men have no right even to praise" [3]. And recent research on healing in social networks suggests that the death of one member often draws together those who survive [4].

Goethe allegedly observed that we die twice [5]: first when we stop breathing, and again when the last of those who knew and loved us die. Philosophers such as Blustein [6] have persuasively argued for the duty to remember and memorialize, which involves obligations to ensure accuracy, sincerity, emotional aptness, publicity, and longevity. We do a disservice to someone's legacy if we attribute characteristics that they did not embody, if we insincerely express our grief, if we erase or hide their characteristics, or if we let their memory fade too quickly. These sometimes-complementary, sometimes-competing injunctions have been recognized in the West as far back as Pericles' funeral oration and Plato's *Menexenus* [7]. In these mourning texts, Pericles and Plato compete to structure the meaning and value of Athenian citizenship by telling the story of recently-deceased soldiers in a way that emphasizes either their courage (in the case of Pericles) or their dedication to justice (in the case of Plato). Both understand that how we choose to remember the deceased expressively constitutes both individual and collective values, sentiments, and relationships. Pericles uses the occasion of patriotic grief to call for public recommitment to war and conflict with Sparta and her allies. Plato uses the same occasion to reorient Athens to justice, with militarism and empire side-lighted as merely sometimes instrumental to that higher value.

Thus, obituaries and poetic elegies reveal what counts as a value, virtue, or constituent of wellbeing for a particular type of person in a particular community. Speaking well of the dead enables us to tell the story both of individuals who have passed away and of the cultures that shaped them and to which they contributed. In so doing, we weave those we held dear into the fabric of human nature, human achievement, and human possibility. The present work aims to contribute to our understanding of moral psychology by extracting individual and communal values from of this rich and under-studied collection.

2 Big Data, Experimental Philosophy, and Moral Psychology

Traditionally, philosophers working in moral psychology have focused on designing and conducting surveys and experiments to elicit judgments that speak to laypeople's moral values and opinions. As with any study eliciting reactions from informed participants, this work is potentially limited in external validity. Studies employing vignettes from professional philosophy, such as stories involving mad scientists, runaway trolleys, and all-knowing supercomputers, run the additional risk of lacking ecological validity. While researchers have made significant progress investigating the judgments laypeople produce when prompted with extraordinary scenarios, our understanding of people's values would also be enhanced by

reviewing the values people express when prompted only by sincere feelings of love, respect, and admiration.

Because obituaries are succinct and explicitly intended to summarize their subjects' lives, they may be expected to include only the features that the author or authors find most salient, not only for themselves as relatives or friends of the deceased, but also to signal to others in the community the socially-recognized aspects of the deceased's character. Such posthumous descriptions have a summative character to them, so their positive or negative valence is presumably meant to evaluate the deceased's moral identity as a whole. To the extent that obituary-writers follow the social norm of only speaking well of the dead, we can expect these evaluations to highlight the deceased's greatest virtues.

The summative nature of death notices highlights our somewhat atypical (in the context of moral psychology and experimental philosophy) goal of studying aretaic phenomena like character and virtue. While most research on moral psychology by philosophers has focused on deontological vs. consequentialist judgments about hypothetical scenarios [8], our goal is instead to identify values associated with living a good life; these values may turn out to include character traits, personality traits, institutional affiliations, social roles, and personal attachments. Indeed, our work indicates that much of the current philosophical study of lay values is working with an impoverished conception of the normative domain.

This is not to say that character, virtue, and social identity have been ignored by moral psychologists. Several approaches have been developed to study lay values, including the models of Schwartz [9] and Graham et al. [10]. Schwartz's research suggests that values are embodied in complex, non-linear networks, an insight that we draw on in our analysis of funerary texts. Graham et al. [11] counted the number of value-laden words in sermons from politically liberal and conservative churches to support inferences about relative value structures within liberal and conservative groups more generally. They thus utilized a lexical approach to understanding evaluative language, according to which everyday language reflects folk conceptions of the nature of morals and values more generally. This approach starts from the thought that ideas that are particularly important to a community will naturally gain expression in that community's language, often as discrete descriptors (e.g., "loyal," "authoritarian," "nature-loving"). This type of lexical approach has a well-established history in the study of personality traits, a field that is also beginning to re-introduce evaluative terms (e.g., "wicked") to what it considers descriptors of traits [12].

We believe that obituaries are an especially promising resource for several pragmatic reasons, in addition to the theoretical considerations mentioned above. First, obituaries, unlike other value-laden corpora such as sermons, diaries, and letters, are published in newspapers and readily available online. Second, it is at least conceptually straightforward to use metadata to segment obituaries (and therefore populations and sub-populations) based on potentially moderating variables such as age, location, gender, and so on. Third, obituaries tend to focus primarily on moral values and virtues.

We expect the values expressed in these records to be diverse but also interconnected. Different communities and individuals are unlikely to unanimously

agree on a single set of norms or ideals, but it would be equally surprising to find no significant overlap. Thus, the ideal approach to analyzing these texts will be sensitive to the heterogeneity in the various ways that we praise the dead. This analysis could be conducted by developing an extensive list of correlations between terms and phrases, but human observers have little to gain from a sprawling list of correlations between hundreds of variables. Thus, where we aim to present a holistic picture of the virtues expressed by each community, we will do so in the form of network graphs simultaneously representing hundreds or thousands of relations in a single, easily digestible picture.

We report two approaches – one employing expert coding, the other employing machine coding – to the scientific study of virtue and value through obituaries. Studies 1 and 2 employed hand-coding, in which obituaries were carefully read and labeled. Study 3 further develops the results of studies 1 and 2 with a semi-automated, large-scale semantic analysis of several thousand obituaries. We conclude by reflecting on some philosophical implications of these studies, as well as prospects for further research.

2.1 Study 1: Local Obituaries

Given our goal of using obituaries to better understand laypeople's values, we aimed to select obituaries that are broadly representative of the general public's values, insofar as this is possible. To this end, we selected local newspapers with obituaries of people from all walks of life instead of targeting obituaries of larger newspapers that selectively write about famous individuals. We realized that obituaries from newspapers such as *The New York Times* would be both more selective (only covering a few famous or infamous individuals) and more comprehensive (written with the goal of telling rich and captivating stories of those few individuals whose lives were deemed worthy of note). We also found that, while *Times* obituaries are authored by professional writers, the obituaries in local newspapers are typically composed by laypeople with personal relationships with the deceased. The stories of (in)famous individuals in the *Times* include a full range of virtues, vices, and value-neutral descriptions. Our initial goal, though, was to target obituaries as a means of developing a better understanding of *positively* (or at worst neutrally) valenced terms.

2.1.1 Methods

Obituaries published between November 2013 and January 2014 were collected from newspapers in four cities: *The Register Guard* (Eugene, Oregon), *The Mat-Su Valley Frontiersman* (Wasilla, Alaska), *The Flint Journal* (Flint, Michigan), and *The Hampshire Gazette* (Amherst, Massachusetts). These were read with an eye to agent-level descriptions of the deceased (e.g. 'hard-working', 'honest', 'generous'). General categories of traits were developed inductively as obituaries were read; put differently, trait terms were manually, iteratively recorded verbatim and then summarized as the reader found other terms that shared a theme. For example, 'hiker' and 'fisherman' could be jointly summarized in the bigram 'nature-lover'.

Given its exploratory and close-reading approach, this process utilized the expertise of the coder as a values-literate reader of American English rather than predefined rules for coding. Additional categories were added as new types of descriptions were found that did not match extant labels.

2.1.2 Results

We collected and analyzed 930 obituaries (52% female) in total across the four cities, as summarized in Table 1.

We were particularly interested in investigating trends in the co-occurrence of descriptions of the deceased within obituaries. To see this, the co-occurrence of traits X and Y was treated as an undirected edge in a network, with the weight of each edge equal to the total number of obituaries in which the deceased was described as both X and Y. The resulting networks for each city and the combined network of traits are represented visually in Figures 1, 2, 3, 4, 5, 6, 7, 8, 9. These visualizations were generated with *Gephi*, using a standard ForceAtlas layout, adjusting node positions to avoid overlapping labels, and, for more populated networks, attraction distribution to highlight clusters of traits.

Figure 1 displays the traits and values ascribed to the deceased in Eugene. This graph depicts trait frequency with size (e.g., obituary writers highlighted volunteering more often than integrity, represented by the fact that "volunteer" is larger than "integrity" in the graph). Gender differences are represented with label coloring, with terms colored red to the extent that they tended to be ascribed to women, blue to the extent that they tended to be ascribed to men, and black for terms that were roughly equal in frequency after adjusting for total word count of male and female obituaries. Finally, line thickness and color represent the frequency of co-occurrence of traits, with thick green lines indicating pairs of terms that were frequently ascribed to the same individual. A label's position is not meaningful on its own, but positions carry meaning in relation to the positions of every other node. In graphs such as this, node positions are determined by an iterated application of an algorithm with three primary factors: (1) gravity, pulling all nodes to the center, (2) attraction, pulling nodes together if they are connected by an edge, with the strength

Table 1 Descriptive statistics for local obituaries in Eugene, Wasilla, Flint, and Amherst

City	Total Obits	% Female	Trait Types[a]	Traits/Obit[b]
Eugene, OR	169	45%	76	2.69
Wasilla, AK	49	57%	40	1.92
Flint, MI	147	55%	74	2.16
Amherst, MA	565	53%	279	4.64
Total	930	52%	469	1.98

[a] "Trait Types" refers to the number of unique traits recorded across all obituaries from the city

[b] "Traits/Obit" refers to the average number of trait-words (regardless of their uniqueness) recorded in each obituary from the city

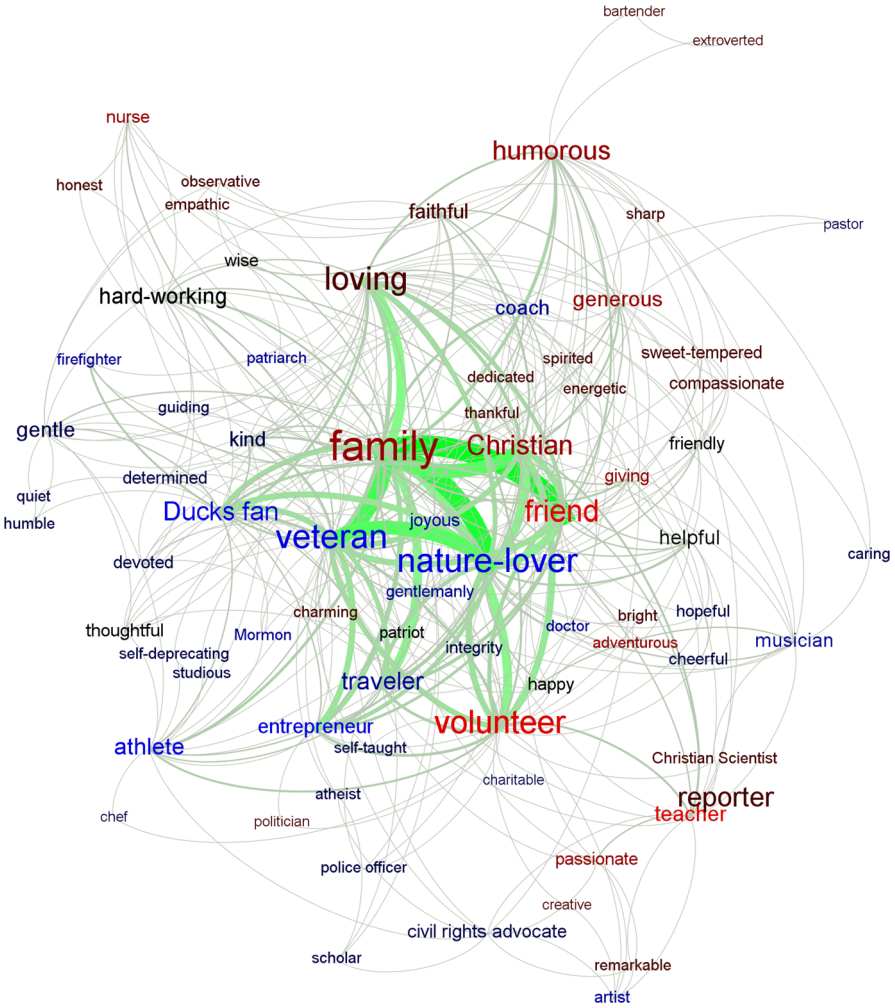

Figure 1 These terms represent descriptions of the deceased in the Eugene's The Register Guard. Node colors reflect sex (male = blue; female = red; mixed = black) and label size is based on node PageRank. The color and width of edge X–Y represents the frequency of co-occurrence between X and Y

of the attraction determined by the weight of the edge, and (3) repulsion, whereby all nodes are pushed away from all other nodes.

Similar analyses were conducted using the trait-terms gathered from obituaries from Wasilla, Alaska (Figure 2), Flint, Michigan (Figure 3), and Amherst, Massachusetts (Figure 4).

As with the obituaries from Eugene and Wasilla, gender differences were seen in the obituaries from Flint (above) and Amherst (below), but these graphs are colored to identify clustering of traits because the larger samples from these cities allowed for a more meaningful analysis of clusters of nodes, with clustering analyzed using Gephi's Modularity Measure.

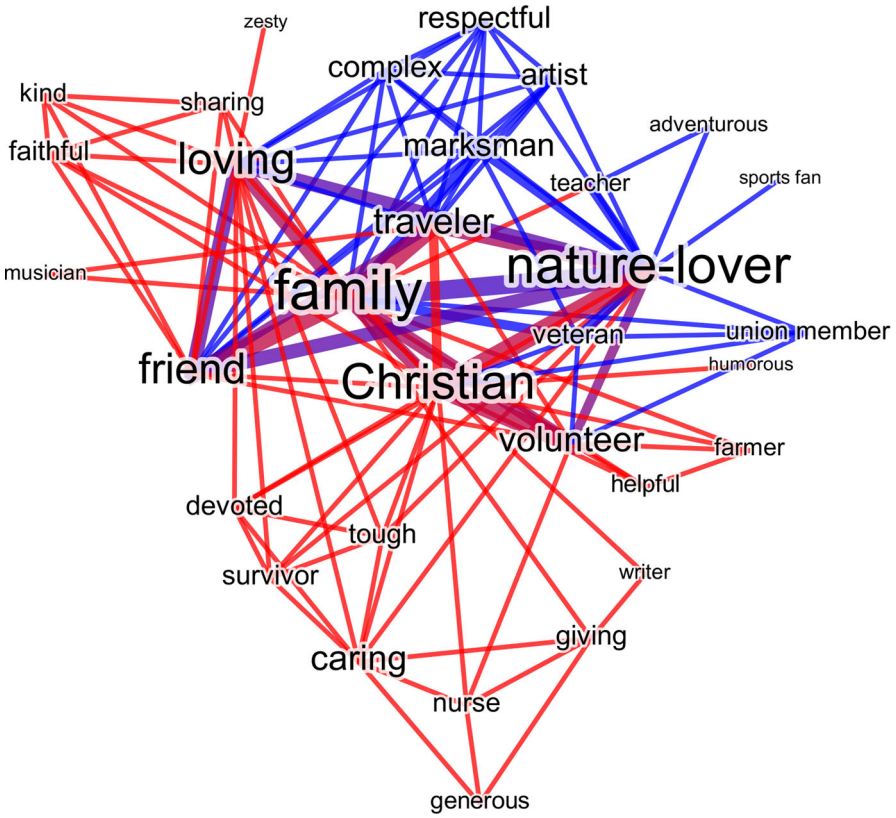

Figure 2 These terms represent descriptions of the deceased in the Wasilla's The Mat-Su Valley Frontiersman. Node positions were determined by the ForceAtlas2 layout algorithm with heightened edge weight influence and adjusting for label overlap. Edge colors reflect whether these traits co-occurred in obituaries for men (blue), women (red) or both (shades of purple). Label sizes are based on nodes' PageRank scores

In addition, we combined the data for all four towns to create a gender-contrast map for the full dataset (Figure 5).

To represent the combined data from local obituaries in more detail, nodes were assigned into four groups based on their edge weights using an algorithm introduced by Blondel et al. [13]. The four figures below show these four clusters of strongly interconnected descriptions.

2.1.3 Discussion

Naturally, these visualizations are only one perspective on the value-networks of the communities they represent. We believe, though, that they point to underutilized approaches to moral psychology. Experiments eliciting reactions to vignettes serve to identify causal relations. Studies employing neuroimaging may help us to identify the mechanisms at work, but our progress on these fronts will be limited by the conceptual frameworks employed, and naturalistic observation is needed to verify

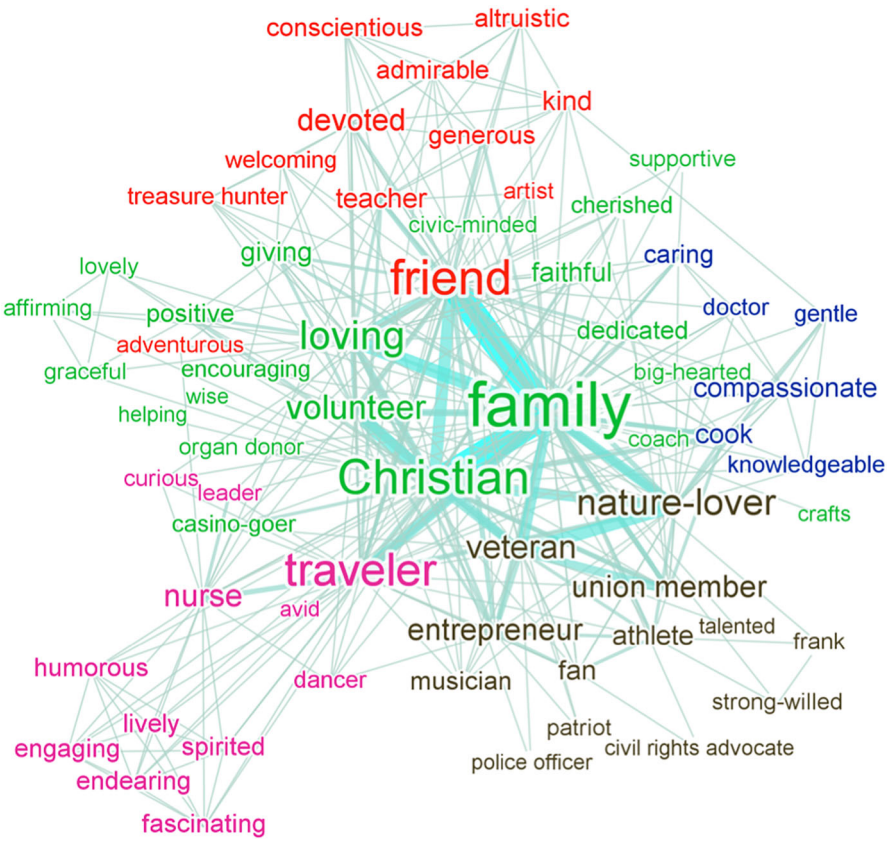

Figure 3 Nodes represent descriptions of the deceased in The Flint Journal. Label size reflects PageRank and label colors are based on group membership as defined by Modularity Measure (resolution = 1.0). One node (knowledgeable) was manually adjusted towards the center to permit a larger view of the network

that the conclusions drawn from controlled studies accurately reflect the lived values of study participants.

Consider, for example, the groundbreaking fMRI studies conducted by Greene and colleagues (Greene et al. [14]; Greene & Haidt [15]; Greene et al. [16]). These researchers presented evidence that reason and emotion each play a significant role in shaping people's moral judgments, and that characteristically consequentialist judgments are more likely to arise from cognitive deliberation whereas characteristically deontological judgments depend more on emotional reactions. While we acknowledge the value of Greene's work, it also serves as a case study highlighting the shortcomings of experimental work unconstrained by naturalistic observation. In a follow-up fMRI study, Kahane et al. [17] offer reasons to think Greene et al.'s results stem from a biased set of thought experiments, such that the distinction between consequentialist and deontological judgments is not the best explanation for the neurological differences in their subjects. More broadly, sociologists of science such as Abend [19] have questioned the narrow focus of much recent

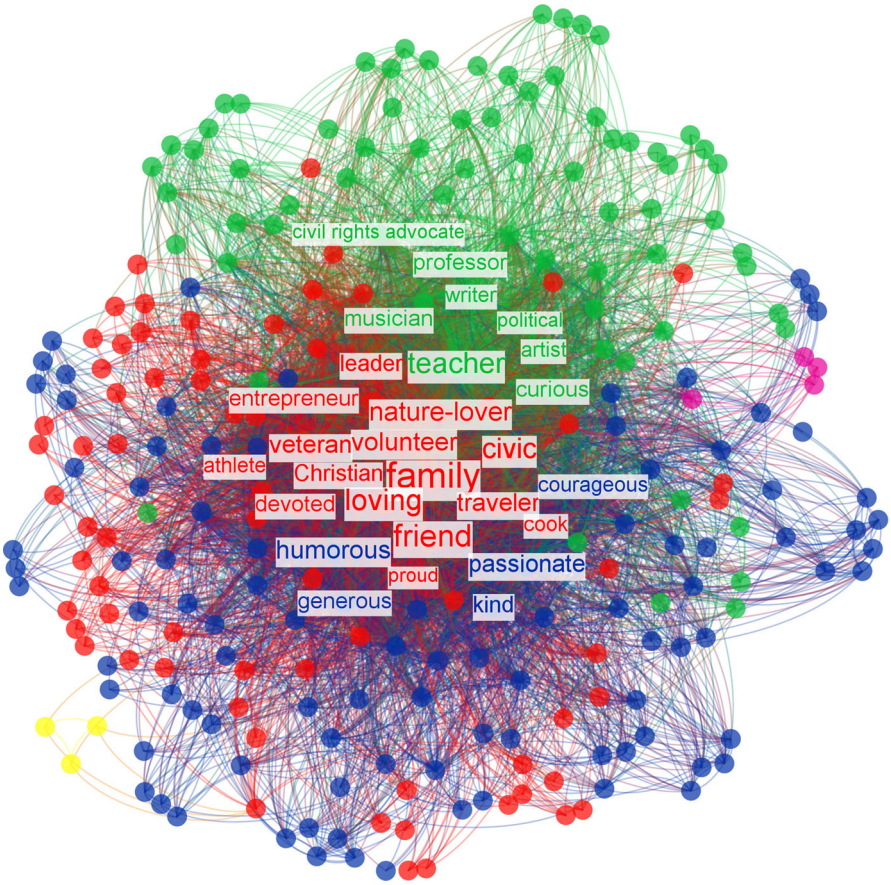

Figure 4 Nodes represent descriptions of deceased women in Amherst's The Hampshire Gazette. Label size reflects PageRank and label colors are based on group membership as defined by Modularity Measure (resolution = 1.0)

research in moral psychology, calling for renewed efforts to understand virtues and values in addition to the differences between processes associated with consequentialist and deontological reasoning. Horne & Powell [18] argue in a similar vein that, while emotions bear on moral judgment, they do so in a more complex way than Greene and colleagues have suggested. Just as our work differs from Graham et al.'s lexical approach in not predefining the semantic framework, so it also differs from Haidt and Greene's work in neuropsychology by not predefining the set of relevant character traits or behaviors. More generally, the field of ethics-oriented experimental philosophy has been criticized for narrowly focusing on deontology vs. consequentialism at the expense of the larger and richer body of normative concepts applied by lay people [19, 20]. Insofar as the results presented here reflect ordinary moral conceptualization and practice, this criticism has some merit.

The practice of eliciting one-dimensional evaluations of bizarre stories also raises concerns about ecological validity. Casebeer & Churchland [21] observe that moral

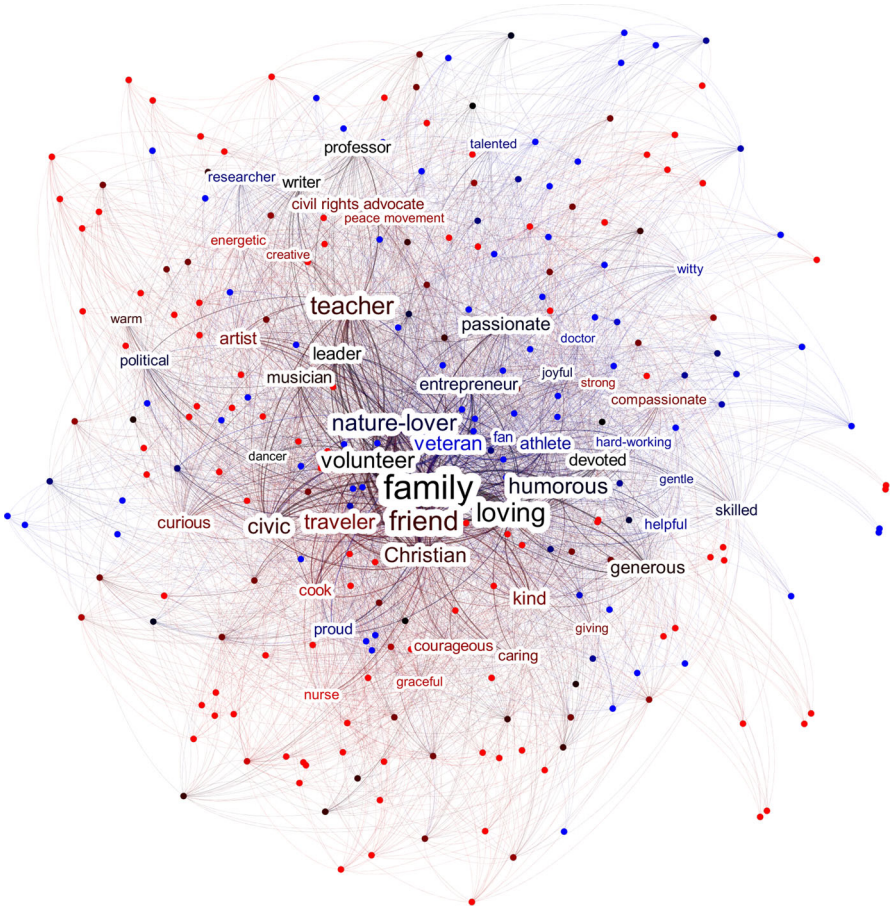

Figure 5 Nodes represent descriptions of the deceased in Amherst, Eugene, Flint, and Wasilla. Due to the large set of traits in this graph, only the top 50 nodes, as measured by PageRank, were labeled. Label, node, and edge colors represent the typical sex of the person thus characterized (red = female, blue = male, black = balanced). Edge width is directly proportional to edge weight

cognition is typically affective, social, organic, genuine, and based on a broad, distributed network of neural mechanisms. These features cannot be fully replicated in the lab, indicating that naturalistic observation is especially valuable in studying moral psychology. While our chosen datasets may have shortcomings, ecological validity is certainly not one of them. The information gathered here comes from real people describing those they knew intimately; as such, these texts are affectively rich and genuine, intentionally social, and free of any possible experimenter bias (though not, presumably, of other biases such as selection bias). In studying value-laden discourse "in the wild," we provide the field of moral psychology an important tool for assessing existing theoretical frameworks.

We will discuss these graphs further below, but it is worth noting already that the vast majority of terms identified by this method refer to virtues, values, or constituents of well-being. This is especially clear if the notion of a virtue is

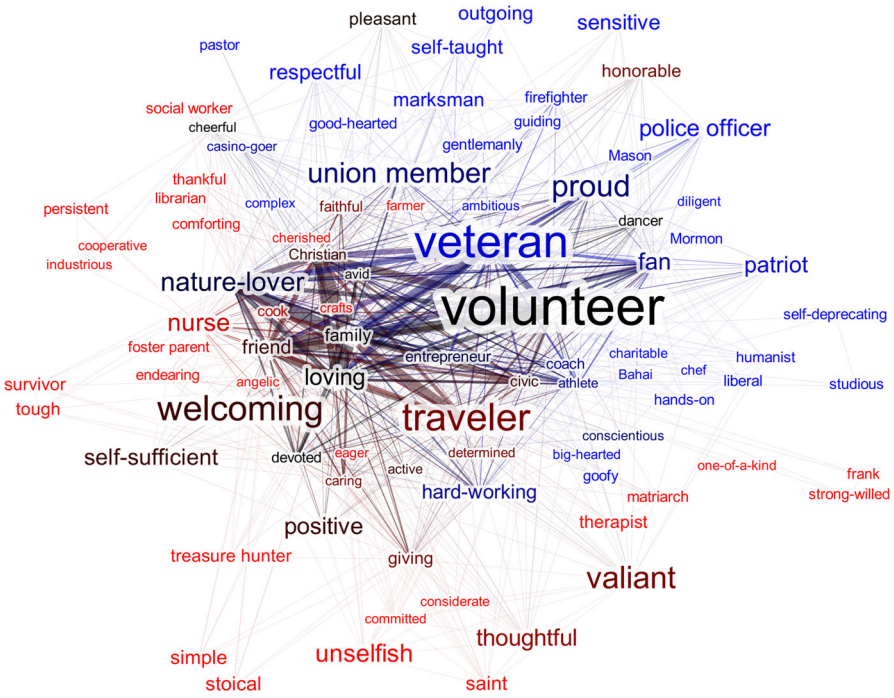

Figure 6 A subset of descriptions of the deceased in Amherst, Eugene, Flint, and Wasilla. Color and edge width were determined as in Figure 5. Label size is based on eigenvector centrality

construed broadly to include affiliations and social roles. As Adams [22] and Alfano [23–25] have argued, such affiliations and social roles may constitute virtues or vices under certain conditions. Another thing to notice is that, while there are some noticeable differences in the value networks of different towns, the main source of differences is gender. There are two interesting aspects to these gender differences. First, women's obituaries tend to focus more on domestic and care-related affairs (e.g., 'nurse', 'volunteer'), whereas men's tend to focus more on public and political matters (e.g., 'sports fan', 'veteran'). This is not particularly surprising, as these are highly stereotyped behaviors. We therefore regard this finding primarily as a proof of concept. However, to the extent that such stereotypical values are celebrated (and other values are not), women and men both presumably face social pressure to live up to different normative standards — which may burden both groups in different ways. This indicates that studying obituaries may help us to identify gender-based cultural constraints on autonomy. Second, some of the dispositions that are differentially associated with women (e.g., 'tough', 'survivor') bear a remarkable resemblance to the "burdened virtues" that Lisa Tessman [26] defines as "traits that make a contribution to human flourishing […] *only* because they enable survival of or resistance to oppression." If this is on the right track, the study of obituaries can unearth both the normative expectations that constrain the virtues and values someone is taken to be capable of embodying (via stereotyping) and the burdened virtues she may end up cultivating in the face of systematic oppression.

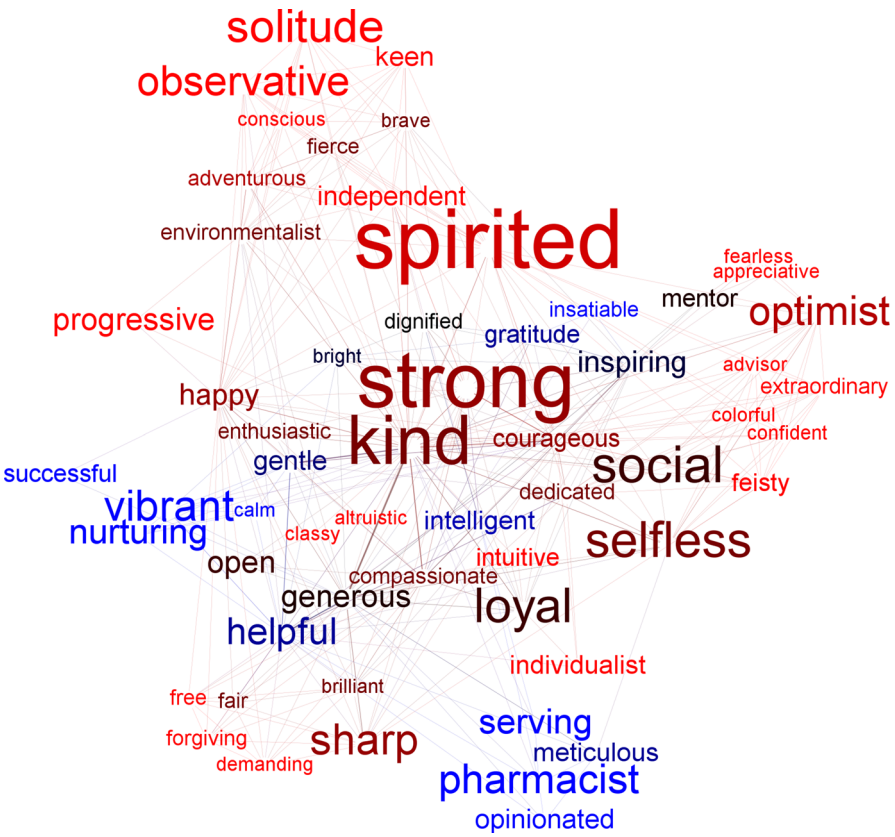

Figure 7 A subset of descriptions of the deceased in Amherst, Eugene, Flint, and Wasilla. Color and edge width were determined as in Figure 5. Label size is based on eigenvector centrality

2.2 Study 2: *New York Times* Obituaries

We anticipated that local newspapers would be a useful place to look for variability in expressions of commonly shared values. However, we also wished to investigate the obituaries of significant figures because these might give us a better picture of the dispositions of people who are commonly seen as great or otherwise noteworthy. It might be that people generally take certain traits, such as honesty and dedication, to be virtuous for friends and loved ones, while evaluating famous individuals using different normative criteria, perhaps differentially valuing traits such as leadership and decisiveness. Following Zagzebski [27] on exemplarist virtue theory and Alfano [28] on its Nietzschean variant, we consider someone a candidate moral exemplar if they evoke strong emotions such as admiration. There is no established index of admiration, so this definition must be operationalized in an empirically tractable way. We do so by assuming that individuals who are memorialized in the *New York Times* are candidate exemplars. To explore the differences between lay values and elite values, as well as the values of candidate exemplars, we read

Figure 8 A subset of descriptions of the deceased in Amherst, Eugene, Flint, and Wasilla. Color and edge width were determined as in Figure 5. Label size is based on eigenvector centrality

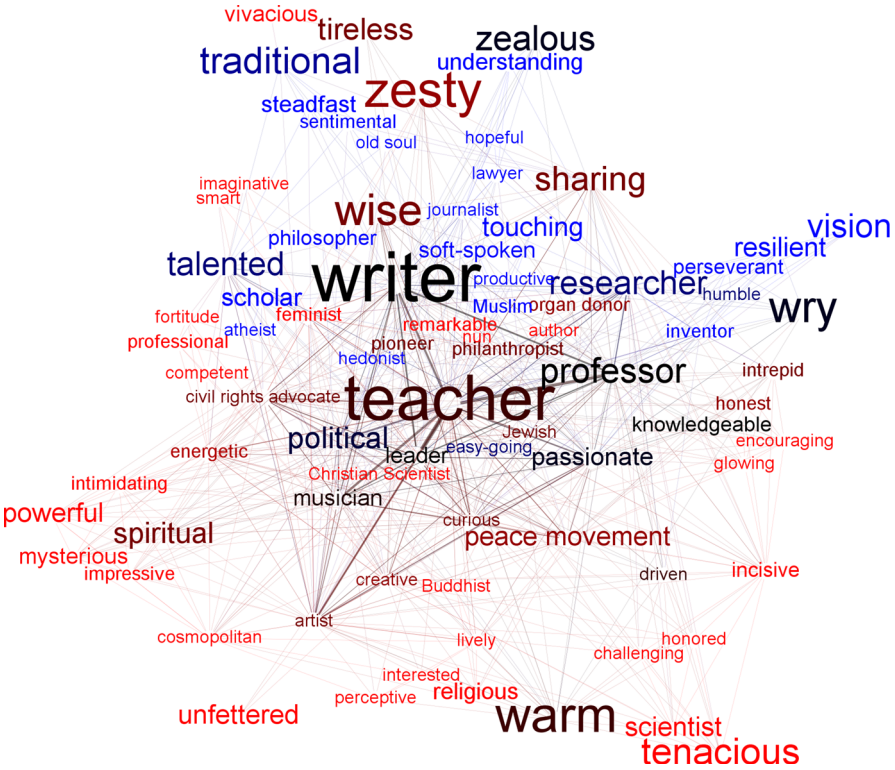

Figure 9 A subset of descriptions of the deceased in Amherst, Eugene, Flint, and Wasilla. Color and edge width were determined as in Figure 5. Label size is based on eigenvector centrality

and coded obituaries from *The New York Times*, following a similar set of methods as in study 1.

2.2.1 Methods

All obituaries published by *The New York Times* from 1 October 2013 to 1 February 2014 were read and analyzed. As with study 1, information was gathered on the age at death, gender, and traits of the deceased.

2.2.2 Results

A total of 74 obituaries (13% female) from *The New York Times* were read and labeled based on trait-ascriptions found in the obituaries. Three-hundred thirty-seven trait types were included in the sample, with an average of 8.5 traits per obituary. These data were then translated into a network format, treating co-occurrence of a trait within an obituary as an undirected edge between trait nodes, resulting in 3,939 edges. The resulting network is displayed in Figure 6.

As in study 1, the resulting sets of traits were analyzed from a network perspective. Where someone was described as both X and Y, an undirected edge was created linking X to Y, with edge weight based on the total number of people described as both X and Y (Figure 10).

2.2.3 Discussion

Value-maps like the one in Figure 6 represent a complex social structure of valuation. They show what journalists at the *Times* consider important about praiseworthy and noteworthy individuals.[2] These results show that the most common person-level descriptor in such obituaries is 'honored', an indication that these individuals inspire strong positive emotions. Unlike the obituaries from local papers, these hardly mention family, friends, and local community. Instead, they focus on larger communities (e.g., 'civil rights advocate'), intellectual accomplishments (e.g., 'author'), and leadership. Compared to the network of traits from ordinary obituaries, the *Times* network has a lower network density (.066 vs. .115), a slightly higher mean clustering coefficient (.875 vs. .806), and higher modularity (.454 vs. .121).[3] This is at least partially explained by the lower number of individuals represented and the wider range of terms used to describe these individuals, but the difference in network modularity is likely also explained by greater eccentricity of the people written about in the *Times*. Obituaries in the *Times*

[2] Another approach to the dispositions valued in famous individuals would be to mine the online comments on their obituaries. Still another would be to mine Twitter and other social media for posts with a '#RIP' or '#RestInPeace' hashtag. We conducted a pilot investigation of these prospects but found that the content was not sufficiently rich to support a scholarly analysis.

[3] We should note that, because *Gephi* does not support significance-tests, we are not in position to say whether any of these differences is statistically significant. One promising area for further research is to repeat our analyses, for example, in the recently-developed network analysis R package called IsingFit (Borkulo & Epskamp [29]; Epskamp et al. [30]; Borkulo [31].

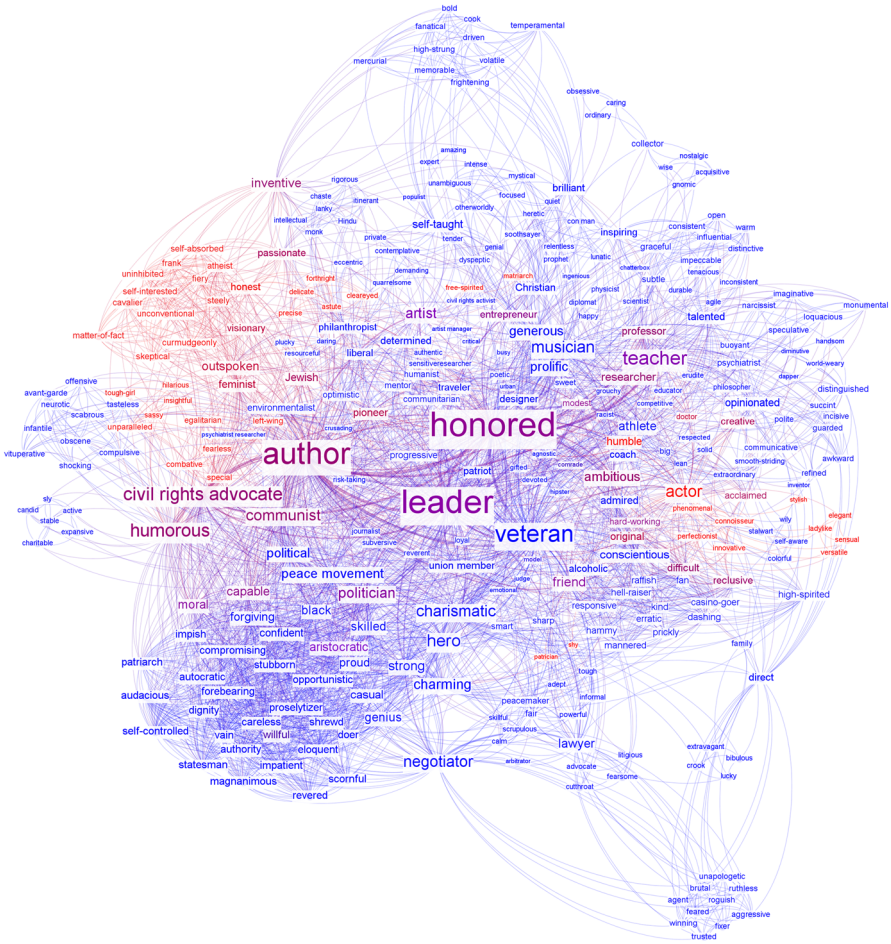

Figure 10 These terms represent descriptions of the deceased in The New York Times. Term and edge colors reflect whether these traits co-occurred in obituaries for men (blue), women (red) or both (shades of purple). Label sizes are based on nodes' PageRank scores. Unlike the obituaries from local papers, these hardly mention family, friends, and local community. Instead, they focus on larger communities (e.g., 'civil rights advocate'), intellectual accomplishments (e.g., 'author'), and leadership. These obituaries also include more negative terms than those in local newspapers. Terms printed in blue (e.g., 'veteran', 'hero', 'negotiator') are more associated with men; those in red (e.g., 'honest', 'outspoken', 'curmudgeonly') are more associated with women

also contain more negative terms than those in local papers. This is consistent with the fact that many exemplars inspire not simple admiration, as Zagzebski [27] suggests, but deep ambivalence. Many exemplars, such as Socrates, Jesus, Rosa Luxemburg, Malcolm X, Harvey Milk, and Dian Fossey, met violent ends that presumably stemmed from emotions other than admiration. By examining Figure 6, we can see that female exemplars in particular often inspire ambivalence rather than simple admiration. In addition, they often embody burdened virtues (e.g., 'outspoken', 'curmudgeonly').

2.3 Study 3: Large-Scale Data-Mining of Local Obituaries

After manually reading, coding, and analyzing over one thousand obituaries in studies 1 and 2, we were interested in developing methods for automatically encoding the traits ascribed in obituaries on a significantly larger scale. Hand-coding has clear advantages, especially with regards to our confidence that sampled obituaries are correctly parsed as ascribing traits to the deceased (rather than, say, commending caretakers for their devotion to the deceased). However, if an automated or semi-automated process should turn out to give similar results to the results of manual reading, we could be reasonably confident that the automated processes are not too compromised by misapplications of terms to the deceased (false positives) or missing content (false negatives) to render their results suspect. To gather more data and to test general reliability of a semi-automated data-mining process, we sought to test new methods on a batch of several thousand obituaries.[4]

2.3.1 Methods

Obituaries were collected from *ObituaryData.com* in collaboration with the Alumni Office of a medium-sized public university in the United States. *ObituaryData.com* is a data-warehousing company that maintains a subscription to the United States' Social Security Death Master File, allowing a wide and nearly complete sample of deaths and respective obituaries within the United States. The University Alumni Office's account included permission to search the full text of all archived obituaries for keywords including the university's name and most common abbreviation of the university. We conducted automated acquisition of the entire collection of obituaries matching these terms, totaling 13,209 records from March 2000 to May 2014 and containing over 3.9 million words.

Gathered records comprised the following information about the deceased: name, city and state of residence, date of birth, date of death (or, lacking this, date of obituary publication), and full-text obituary content. Age in years at death was calculated from date of birth and date of death. In addition, forenames of the deceased were used to guess gender (Female, Male, or Unknown). To accomplish this, we compiled the most popular 4,275 female and 1,219 male names in the US as of 2005, based on 1990 US Census Bureau Data and other sources in 2005. The female and male name-lists were each ordered by decreasing popularity of name. The forename of the deceased in each gathered obituary record was then compared with both lists and assigned the gender of whichever list in which it appeared higher (e.g., if the name of the deceased was given as "Bobby" in the obituary record, the record was listed as "Male," because "Bobby" was listed as a more popular male than female name in the Census Bureau name lists).

Using *ConText*, a program developed for semantic network analysis, we performed the following manipulations of the text: (1) changed all letters to

[4] Another way to check the reliability and validity of automatic processing would be to use both hand-coding and machine-coding on the same corpus, then compare the results. We have not yet done this, but future research — either by ourselves or by others — should undertake it.

lowercase, (2) applied a generic stoplist to the texts, (3) identified bigrams, and (4) merged near-synonyms. After steps (1) and (2), the resulting text, comprising two million words, was used to generate a full list of terms used in all obituaries; two of the three authors then independently read through this full word list and selected each unigram or bigram that could serve as a description of the deceased. The remaining author then reviewed all terms in cases where the first two authors disagreed. Terms selected by at least two of the three authors were retained; all other words from the original texts were deleted. We then reviewed the list of terms again and identified cases of synonyms or near-synonyms. Synonyms were retained based on the following general rules:

1. Adjectives (e.g., "Honorable") were preferable to past participles (e.g., "Honored"), which were preferable to gerunds (e.g., "Honoring"), which were preferable to nouns (e.g., "Honor").
2. Words (e.g., "Adventist") were preferable to phrases (e.g., "Adventist Church").
3. Singular nouns (e.g., "Airplane," to describe a theme of a pilot's life) were preferable to plural nouns (e.g., "Airplanes").
4. Person-descriptors (e.g., "Pilot") were preferable to "themes" (e.g., "Airplane").

After two members of the research team had made independent judgments on conflicting terms, we used the *openNLP* (*open Natural Language Processing*) package in R [32] to automatically tag each suggestion by Part of Speech (POS), including its singular vs. plural status for count nouns. Conflicting judgments were then automatically resolved using the rules above. Conflicts unresolved in this automated way were then resolved manually by the remaining research team member. Through this process, for example, "accomplish," "accomplished," "accomplishing," "accomplishment," and "accomplishments"were all replaced with "accomplishment." In doing so, we ran the risk of conflating semantically distinct terms in some contexts; however, this approach was preferable to not identifying the semantic similarity of these terms in the majority of contexts where they might be used more or less interchangeably.

With the texts filtered and cleaned, we then constructed a semantic map of the obituaries, treating co-occurrence of terms X and Y within the same obituary as an undirected edge in a network. This resulted in a network of 910 nodes and 19,034 edges. Given the scale of this network, it would not be informative to present a visual representation of it in its entirety. Thus, we instead developed the simplified visualization shown in Figure 7, which collapses closely-connected nodes into a single group node. Using *Gephi's* modularity measure (resolution = 1.0), nodes were assigned to groups based on which terms tended to be clustered together. Groups of nodes were then treated as individual nodes.

2.3.2 Results

In Figure 7, nodes represent large sets of trait terms.

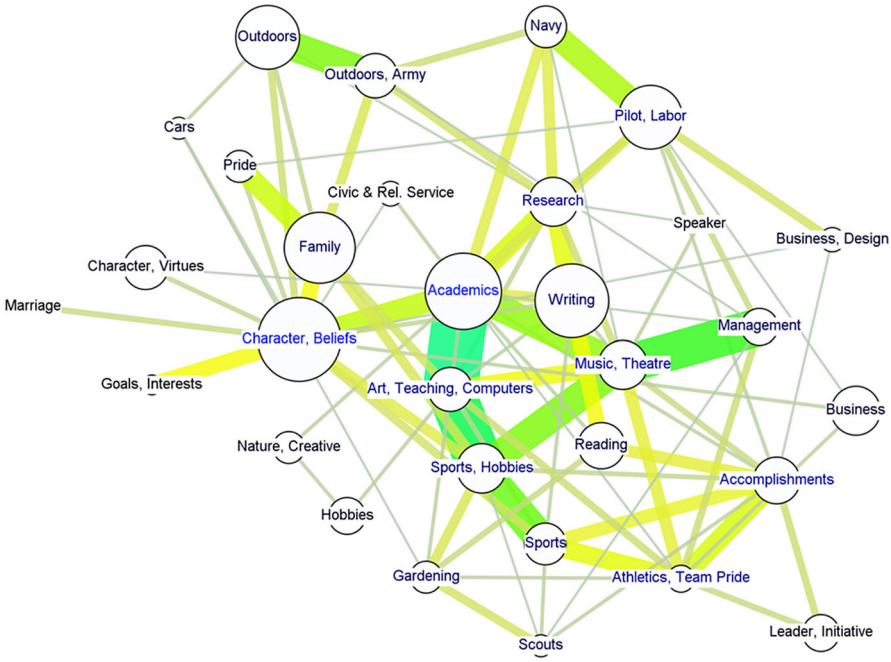

Figure 11 These obituaries were machine-coded. Trait terms identified in the set of 13,209 obituaries, grouped by modularity measure (resolution = 1.0), with groups of nodes treated as individual nodes. Label color reflects PageRank, node size is based on the number of trait terms subsumed under each group, and edge width and color are based on edge weight

Labels were assigned manually based on the common theme we identified in each cluster. Node size was determined by the number of terms subsumed within each group. Edge width was determined by edge weight, and node positions were determined by an iterated application of the ForceAtlas algorithm (Figure 11).

2.3.3 Discussion

Before we initiated this semi-automated analysis, it was an open question whether this process would produce results similar to the results of manual reading. The results suggest that semi-automated data processing results in a network of trait-terms similar to those identified by the authors in carefully reading individual obituaries, but the results diverge from the results of study 1 in some important respects.

First, our semi-automated approach with this new corpus found a substantially higher frequency of scholarly traits. Far more people were described as writers, scientists, researchers, students, teachers, etc. In the obituaries analyzed in study 1, family, friendship and faith were consistently at the core of the network; in the current study, these traits were still frequently mentioned but were less central to the network than academic attributes. This difference is naturally explained by

differences in the original data. The obituaries used for study 1 came from local newspapers where friends or relatives of any deceased person may write about the deceased, but for the present study, obituaries were filtered to include only those mentioning the university with which we were working. Thus, most people described in these obituaries were either alumni or university employees, making is unsurprising that they tend to be described with more academic language.

Second, we observed higher modularity and more distinct patterns in the clustering of traits in the present study. The network's modularity (.514) and mean clustering coefficient (.146) differ from what was found with local papers (.121 and .806) and the *Times* (.454 and .875). The higher modularity is partially explained by the larger sample size. With over 13,000 obituaries to draw from, we could more easily identify the patterns of connections and disregard less frequent or accidental connections.

Third, in addition to virtues and values, which were explicitly coded for in the manual studies, this automated study turned up many constituents of well-being. While it does not seem correct to say that accomplishments, athletics, gardening, or hobbies are *virtues* or *values*, they do seem to be part of what makes some lives enjoyable or worthwhile.

3 General Discussion

We believe this research has three main philosophical implications. First, it can be used to *identify under-theorized candidate virtues, values, and constituents of wellbeing*. In this paper, we understand these terms intuitively, but in general we are in agreement with Zagzebski [27] and Alfano [24] on how to conceptualize virtues, with Tiberius [33] on how to conceptualize values, and with Parfit [34] on how to conceptualize wellbeing. For example, studies 1 and 3 suggest that being a sports fan is commonly conceived of as a virtue or constituent of well-being, but contemporary ethics has almost nothing to say about fandom. Is there such a virtue (or vice)? If so, what is it? Does it differ from being loyal or a community member more generally? The deceased are often praised for their volunteerism. Is there such a virtue? If so, what is it, and how does it differ from benevolence? Likewise, a sense of humor seems to be viewed as a virtue or constituent of well-being, but almost no philosophical work explores what that virtue might be. This possibility has also been raised by Harry Frankfurt [35], John Lippett [36], Adam Morton [37], and Robert Roberts [38]. Roberts in particular argues that being able to laugh at oneself is essential to self-transcendence, which is involved in many virtues. Although Roberts approaches the question from a Christian point of view, he ends up agreeing with Nietzsche [39] that one of the most important targets of laughter is oneself – in the service of what Roberts calls self-transcendence and Nietzsche calls self-overcoming.

The search for under-theorized normative phenomena has so far been fruitful (e.g., Anthony Appiah's *Honor Code* [40]) but haphazardly carried out. Systematically mining obituaries promises a much more comprehensive perspective.

Second, the patterns of virtues, values, and constituents of well-being represented in these maps have implications for normative ethics and the study of oppression.

Whether a trait is considered a virtue for a person of your type influences whether you can see that trait as desirable or praiseworthy. By mapping virtues, values, and constituents of well-being, we *uncover potentially-oppressive constraints on people's virtues and values*. As mentioned above, one conspicuous set of constraints relates to gender. Further research may uncover similar constraints (e.g., the expectation that victims of injustice, such as Black Americans, always be willing to forgive — cf. [41].

Third, this project contributes to the study of descriptive cultural relativism. Although cultural anthropology of the early 20th century tended to support strong versions of relativism, cross-cultural psychological work by Schwartz [42] and philosophical interpretations of value-pluralism [43] suggest that there is at least as much variance in values within cultures as there is between them. This project thus *illuminates both universal human values and the kinds of values-related variability that may differentiate communities*.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Zagzebski, L. 1996. *Virtues of the Mind*. Cambridge: Cambridge University Press.
2. Hayes, S., K. Strosahl, and K. Wilson. 2011. *Acceptance and Commitment Therapy: The Process and Practice of Mindful Change*. New York: Guilford Press.
3. Jaeger, W. 1927. Aristotle's Verses in Praise of Plato. *The Classical Quarterly* 21 (1): 13–17.
4. Hobbs, W., and M. Burke. 2017. Connective Recovery in Social Networks After the Death of a Friend. *Nature: Human Behavior* 1 (92): 1–6.
5. Stokes, P. 2015. Deletion as Second Death: The Moral Status of Digital Remains. *Ethics and Information Technology* 17: 237–248.
6. Blustein, J. 2008. *The Moral Demands of Memory*. Cambridge: Cambridge University Press.
7. Collins, S., and D. Stauffer. 1999. *Plato's Menexenus and Pericles' Funeral Oration: Empire and the Ends of Politics*. Newburyport, MA: Focus Classical Library.
8. Liao, M. (ed.). 2016. *Moral Brains: The Neuroscience of Morality*. Oxford: Oxford University Press.
9. Schwartz, S. 1992. Universals in the Content and Structure of values: Theoretical Advances and Empirical Tests in 20 Countries. In *Advances in Experimental Social Psychology*, vol. 25, ed. M. Zanna, 1–65. New York: Academic Press.
10. Graham, J., B. Nosek, J. Haidt, R. Iyer, S. Koleva, and P. Ditto. 2011. Mapping the Moral Domain. *Journal of Personality and Social Psychology* 101 (2): 366–385.
11. Graham, J., J. Haidt, and B. Nosek. 2009. Liberals and Conservatives Rely on Different Sets of Moral Foundations. *Journal of Personality and Social Psychology* 96 (5): 1029–1046.
12. Saucier, G. 2009. Recurrent Personality Dimensions in Inclusive Lexical Studies: Indications for a Big Six Structure. *Journal of Personality* 77 (5): 1577–1614.
13. Blondel, V., J. Guillaume, R. Lambiotte, and E. Lefebvre. 2008. Fast Unfolding of Communities in Large Networks. *Journal of Statistical Mechanics* 10: P10008.
14. Greene, J., R. Sommerville, L. Nystrom, J. Darley, and J. Cohen. 2001. An fMRI investigation of emotional engagement in moral judgment. *Science* 293: 2105–08.
15. Greene, J. and J. Haidt. 2002. How (and where) does more judgment work? *Trends in Cognitive Sciences* 6 (12): 517–23.
16. Greene, J. D., L. E. Nystrom, A. D. Engell, J. M. Darley, and J. D. Cohen. 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44 (2): 389–400. doi:10.1016/j.neuron.2004.09.027.

17. Kahane, G., K. Wiech, N. Shackel, M. Farias, J. Savulescu and I. Tracey. 2012. The neural basis of intuitive and counterintuitive moral judgment. *Social Cognitive and Affective Neuroscience* 7 (4): 393–402. doi:10.1093/scan/nsr005.
18. Horne, Z., and D. Powell. 2016. How Large is the Role of Emotion in Judgments of Moral Dilemmas? *PLoS ONE* 11 (7): e0154780.
19. Abend, G. 2013. What the Science of Morality Doesn't Say About Morality. *Philosophy of the Social Sciences* 43 (2): 157–200.
20. Churchland, P.M. 1998. Toward a Cognitive Neurobiology of the Moral Virtues. *Topoi* 17 (2): 83–96.
21. Casebeer, W., and P.S. Churchland. 2003. The Neural Mechanisms of Moral Cognition: A Multiple-Aspect Approach to Moral Judgment and Decision-Making. *Biology and Philosophy* 18 (1): 169–194.
22. Adams, R. 2006. *A Theory of Virtue: Excellence in Being for the Good*. Oxford: Oxford University Press.
23. Alfano, M. 2015. Friendship and the Structure of Trust. In *From Personality to Virtue: Essays in the Psychology and Ethics of Character*, ed. J. Webber, and A. Masala. Oxford: Oxford University Press.
24. Alfano, M. 2015. An enchanting Abundance of Types: Nietzsche's Modest Unity of Virtue Thesis. *Journal of Value Inquiry* 49 (3): 417–435.
25. Alfano, M. 2013. *Character as Moral Fiction*. Cambridge: Cambridge University Press.
26. Tessman, L. 2005. *Burdened Virtues: Virtue Ethics for Liberatory Struggles*. Oxford: Oxford University Press.
27. Zagzebski, L. 2017. *Exemplarist Moral Theory*. Oxford: Oxford University Press.
28. Alfano, M. 2015. Ramsifying Virtue Theory. In *Current Controversies in Virtue Theory*, ed. M. Alfano, 123–135. Oxon: Routledge.
29. Borkulo, C., and Epskamp, S. 2016. Fitting Ising Models Using the ELasso Method. https://cran.r-project.org/web/packages/IsingFit/IsingFit.pdf. Accessed 3 Oct 2016.
30. Epskamp, S., Borsboom, D., and Fried, E. 2016. Estimating Psychological Networks and Their Accuracy: A tutorial paper. arXiv:1604.08462v3.
31. Borkulo, C. 2016. Statistical Comparison of Two Networks Based on Three Invariance Measures. https://cran.r-project.org/web/packages/NetworkComparisonTest/NetworkComparisonTest.pdf. Accessed 3 Oct 2016.
32. Hornik, K. 2016. Apache OpenNLP Tools Interface. https://cran.r-project.org/web/packages/openNLP/index.html.
33. Tiberius, V. 2008. *The Reflective Life*. Oxford: Oxford University Press.
34. Parfit, D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
35. Frankfurt, H. 2001. The Dear Self. *Philosophers' Imprint* 1: 1–14.
36. Lippitt, J. 2005. Is a Sense of Humor a Virtue? *The Monist* 88 (1): 72–92.
37. Morton, A. 2013. *Emotion and Imagination*. Cambridge: Polity.
38. Roberts, R. 1988. Humor and the Virtues. *Inquiry* 31 (2): 127–149.
39. Alfano, M. forthcoming. The epistemic function of contempt and laughter in Nietzsche. In *The Moral Psychology of Contemp*t, ed, M. Mason. Lanham, MD: Rowman & Littlefield.
40. Appiah, K.A. 2010. *The Honor Code: How Moral Revolutions Happen*. New York: Norton.
41. Cherry, M. forthcoming. Forgiveness Exemplars and the Oppressed. In *The Moral Psychology of Forgiveness*, ed. K. Norlock. Lanham, MD: Rowman & Littlefield.
42. Schwartz, S. 2012. An Overview of the Schwartz Theory of Basic Values. *Online Readings in Psychology and Culture* 2 (1): 11.
43. Wong, D. 2006. *Natural Moralities*. Oxford: Oxford University Press.
44. Andreason, A., and P. Kotler. 2008. *Strategic Marketing for Non-Profit Organizations*, 7th ed. Upper Saddle River, NJ: Prentice Hall.
45. Christen, M., M. Alfano, E. Bangerter, and D. Lapsley. 2013. Ethical Issues of 'Morality Mining'. In *Ethical Data Mining Applications for Socio-Economic Development*, ed. H. Rahman, and I. Ramos, 1–21. Hershey, PA: IGI Global.
46. Haidt, J., and C. Joseph. 2004. Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues. *Daedalus* 133 (4): 55–66.
47. Kotler, P., and G. Zaltman. 1971. Social Marketing: An Approach to Planned Social Change. *Journal of Marketing* 35 (3): 3–12.
48. Mill, J.S. 1861/1998. *Utilitarianism*, edited with an introduction by Roger Crisp. New York: Oxford University Press.