Title: Into The Wild: Big Data Analytics in Moral Psychology

Authors: Joseph Hoover[1], Morteza Dehghani[1, 2, 3], Kate Johnson[1], Rumen Iliev[4], Jesse Graham[1]

Affiliations:

[1]Psychology Department, University of Southern California

[2]Brain and Creativity Institute, University of Southern California

[3]Computer Science Department, University of Southern California

[4]Ford School of Public Policy, University of Michigan

For inclusion in K. Gray and J. Graham (Eds.), *The Atlas of Moral Psychology*.

Abstract: Moral values are culturally variable entities that emerge from dynamic, hierarchical interactions between individual- and group-level phenomena. Human-generated unstructured data, from sources such as social media, offer an unprecedented opportunity to observe these phenomena in a natural habitat, which is an essential vantage point for understanding moral values and their role in moral judgment and behavior.

1. Introduction.

The interdisciplinary science of morality has blossomed in the last decade, with insights from social psychology, neuroscience, behavioral economics, experimental philosophy, developmental science, sociology, consumer behavior, and anthropology informing one another and inspiring further interdisciplinary collaborations (Haidt, 2007). This proliferation of research has led to substantive theoretical advances, as well as a number of notable disagreements (e.g. see Graham & Iyer, 2012; Gray, Schein, & Ward, 2014; Janoff-Bulman & Carnes, 2013; Rai & Fiske, 2011, on moral values, or Cushman & Greene, 2012; Kahane, 2015, on moral decision-making). Nonetheless, several points of convergence have emerged within the field. It is generally accepted that morality is a fundamentally social evolutionary adaptation and that it arises dynamically through interactions between native, inter-individual mechanisms and socio-cultural factors (e.g. Graham et al., 2013; Haidt, 2012; Mikhail, 2007; Rai & Fiske, 2011; Fiske & Rai, 2015). However, while many contemporary approaches to morality are premised on some iteration of the social-functional evolutionary model, we believe that the majority of the research methodologies used to substantiate these theories are, ironically, not able to adequately account for the dynamic social functioning of morality that they prioritize. Most research on morality is conducted with undergraduates in decontextualized laboratory settings, and much more often than not, morally relevant variables are measured using self-reports. While there is nothing immediately wrong with these methods, we doubt they can fully capture, for example, the highly variable and subjective nature of individual moral values (Graham, 2014; Meindl & Graham, 2014) or group-level moral processes (Ginges, Atran, Sachdeva, & Medin, 2011).

Accordingly, we believe it is vital that researchers supplement traditional methodologies with alternative approaches that have greater ecological and external validity and that are better able to capture the full social-functional range of morality. In this chapter, we argue that a range of computationally intensive methods, drawn predominantly from Computer Science and Computational Linguistics, can help researchers do just this. By mining psychologically relevant information from large-scale, human-generated, online data such as blog posts, news articles, tweets, Facebook status updates, and social-network structures–collectively referred to as ‘big data’–researchers can use these methods to investigate morally relevant phenomena in the real world. These methods enable researchers to investigate large-scale, diachronic moral phenomena such as the diffusion of moral values through populations and the moralization of specific topics (Graham, Haidt, & Nosek, 2009; Sagi & Dehghani, 2014). They also offer researchers new opportunities to investigate the relationship between moral values and moral behavior (Dehghani et al., 2015; Boyd et al., 2015), which is both notoriously difficult to study in the laboratory and deeply important for understanding how morality functions (Graham, Meindl, & Beall, 2012; Graham et al. 2013). Of course, we are not suggesting that these methods–which we refer to as ‘big data analytics’–can or should replace traditional approaches. Big data analytics have their own weaknesses (Ruths & Pfeffer, 2014) and they cannot match all of the strengths of conventional methodologies. Fortunately, researchers do not have to choose one or the other.

Indeed, our view is that researchers will benefit the most by developing rigorous multi-method approaches that counterbalance the weaknesses of traditional methods with the strengths of big data analytics and vice versa (Dehghani et al., 2015).

We believe big data analytics offer sufficient advantages for basic psychological research to warrant their inclusion in social scientists' toolkits. However, the behaviors responsible for generating morality-relevant big data–such as participation in online social networks–have become increasingly prominent across social groups and cultures, which marks them as targets of study in their own right. As of January 2015, more than 25% of the global population was using social media (Kemp, 2015) and these platforms are increasingly being used as loud-speakers when morally relevant events take place. Four of many possible examples are the hashtags ("#"): #blacklivesmatter and #baltimore, which have been used widely in protests against recent incidents involving police brutality across the United States; #governmentshutdown, which was prominent in discussions of the 2013 U.S. Government shutdown; and #AllEyesOnISIS, which has been used by ISIS in disseminations of propaganda and by individuals to show support for the extremist organization. Clearly, this level of global connectivity is unprecedented, and it likely has important effects on processes related to morality. As social media communications become more deeply woven into the fabric of society, understanding trends in dynamic morally-relevant phenomena may increasingly require understanding the psychological role that social media plays in contemporary societies. By incorporating big data analytics into the study of morality, researchers will gain a new way to gather information in natural settings about the structure of moral visions (Graham & Haidt, 2012), large-scale moral behavioral patterns, and the relation between the two. However, they will also be able to explore the specific effects that today's communication technologies have on relevant phenomena. This methodological development could potentially transform the study of morality, improving the ecological and external validity of a field that has relied almost exclusively on self-reports sampled from predominantly WEIRD (Western, Educated, Industrialized, Rich, Democratic; Henrich, Heine & Norenzayan, 2010) populations.

2. Historical Context.

Mining massive sets of extant data for psychological information is a relatively new practice, and it has become possible only through constant increases in computational power, availability of new methods, and greater accessibility of human-generated data. Recently, two methods of analysis—Natural Language Processing and Social Network Analysis—have emerged as valuable tools for gleaning psychologically relevant information from online data. However, while these methods are gradually being incorporated into psychological research, psychologists still primarily rely on rudimentary and increasingly dated techniques. Further, until recently, these methods have remained almost completely neglected in moral psychology. Therefore, while a comprehensive review of these methodologies is beyond the scope of this chapter, a brief introduction to their aims and approaches is provided below, followed by a discussion of how they fit into contemporary models of morally-relevant phenomena.

Natural Language Processing (NLP) dates back to the 1950s (Nadkarni, Ohno-Machado, & Chapman, 2011; Jones, 1994; Dostert, 1955) and relies on a range of approaches to parse semantic information from unstructured text (Iliev, Dehghani, & Sagi, 2014). Initially developed in Linguistics and Computer Science, NLP has only recently been incorporated into psychological research. However, the notion that psychological information can be gleaned from language is hardly a new idea; for over a century, researchers have relied on language to make inferences about human psychology (Freud, 1901 ; Rorschach, 1964 [1921]; Murray, 1943 ; Van Dijk & Kintsch, 1977 ; Weber, Hsee, & Sokolowska, 1998; Braun & Clarke, 2006). The availability of digitized natural language corpora—drawn from sources such as blogs, Congressional transcripts, news publications, and social networking platforms like Facebook and Twitter—has allowed researchers to explore the relationship between natural language and psychology at an unprecedented scale.

How NLP is accomplished ranges considerably between methodologies. For conceptual clarity, Iliev, Dehghani, and Sagi (2014) separate NLP methods into three broad groups, which is the approach we take here. In the first group of methods, ‘user-defined dictionaries’ (UDD), researchers rely on expert-generated dictionaries, which specify words that are relevant to dimensions of interest. Popularized in psychology by James Pennebaker and colleagues (Pennebaker, 2011; Tausczik & Pennebaker, 2010), these methods aim to classify the semantic content of texts along a given dimension by summing the within-text occurrences of words specified by the UDD as related to the dimension. For example, sums of positive- and negative-affect word occurrences can be used to infer the overall sentiment of a text (Kahn, Tobin, Massey, & Anderson, 2007) and, further, such sentiment analyses can be used to make predictions about individual differences, such as depression (Rude, Gortner, & Pennebaker, 2004).

The methods in the second class, ‘feature extraction’ methods, forego UDDs, and rely on machine learning algorithms to extract features from texts that are predictive of variables of interest. In this case, a subset of texts pre-classified on a variable of interest (e.g. gender or religious affiliation) are used to ‘train’ an algorithm to detect the features that predict the target variable. After training, the algorithm is tested on an independent pre-classified set of texts, which allows researchers to obtain relatively stable estimates of the classifier’s error rate. The algorithm can then be used to classify un-labeled texts on the variable of interest through probabilistic estimation (though it should be noted that as target texts increase in difference between the training- and test-texts, accuracy has been shown to decrease, sometimes dramatically).

One shortcoming of both UDD and feature extraction methods, however, is that they rely on individual word occurrences and are not able to account for the context in which a word occurs. Because words do not occur in isolation, this leads to substantial information loss. The methods in the third class, ‘word co-occurrence’ methods, attempt to minimize this information loss by capturing the relations between words. In general, this is accomplished through several steps, though these steps vary between specific methods. For example, Latent Semantic Analysis (LSA; Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997; Dumais, 2004) involves first representing words and documents–any discrete set of texts, such as Tweets, blog posts, or entire novels–as vectors in high-dimensional space. In this space, words that tend to appear in the same documents are closer to each other, and documents that use similar words are closer to each other. This then permits analysts to the assess the semantic similarity between words and between documents by measuring the 'distance' between these entities. Other word co-occurence methods include, for example, Latent Dirichlet Analysis (Blei, Ng, & Jordan, 2003; Blei, 2012), new vector based methods (e.g. Mikolov, Sutskever, Chen, Corrada, & Dean, 2013; Sagi & Dehghani, 2014), and TopicMapping (Lancichinetti, Sirer, Wang, Acuna, Körding, & Amaral, 2015). While these methods are considerably more complex than UDD and feature extraction methods, they constitute much of the cutting edge in NLP. Accordingly, as morality researchers begin testing increasingly sophisticated hypotheses using large-scale text corpora, it will be essential that they incorporate these methods into their analyses.

While NLP focuses on quantifying natural language generated by individuals, Social Network Analysis (SNA; Marin & Wellman, 2011) aims to understand human behavior in terms of group-level systems of relational patterns. SNA represents social groups as relationships ('edges') between individuals ('nodes') in order to quantify complex group-level phenomena. As SNA was originally developed by sociologists, social network research tends to prioritize network-based explanations of phenomena and, in some instances, rejects outright the notion that social norms and individual-level psychological characteristics play an important causal role in network outcomes (Marin & Wellman, 2011). However, network analysts are increasingly recognizing the role of individual differences—particularly individual moral differences—in network composition (Vaisey & Lizardo 2010; Hitlin & Vaisey, 2013). While traditional SNA treats networks as exogenous factors that determine social behavior, recent research suggests non-network factors can also affect network formation. For example, Clifton, Turkheimer, and Oltmanns (2009) identified reliable relationships between psychopathological characteristics of military personnel and their social network positions. Further, Vaisey and Lizardo (2010) found that moral disposition is a better predictor of network composition than network composition is of moral disposition, suggesting that networks might be better conceptualized as endogenous factors with reciprocal, hierarchical relations to their nodes.

In tandem, NLP and SNA allow researchers to quantify individual-level natural language expressions and model complex group-level network dynamics. These methods have only recently begun to be applied to research on morally relevant phenomena (e.g. Dehghani, Sagae, Sachdeva, Gratch, 2014; Vaisey & Miles, 2014; Graham, Haidt, & Nosek, 2009); however, we believe that they offer a valuable complement to the methods traditionally used to investigate moral phenomena, which rely almost exclusively on self-report measures and highly controlled experimental paradigms. By incorporating these methods into their research programs, scientists—regardless of their theoretical framework—can begin to provide stronger tests for hypotheses by making predictions about real-world phenomena. While these methods are relatively new, they offer possibilities that have been sought for decades by psychologists—access to relevant phenomena untainted by the biases that accompany laboratory-based research (Barker, 1968; Gibson, 1977).

3. Methodological Stance.

Over time, researchers have come to recognize that morality is constituted by a network of components, including values, judgment, intuition, reasoning, and behavior. While exactly how these phenomena fit together is not fully agreed upon, this general view has been supported by research employing a wide range of methodologies, including laboratory experiments, cross-cultural surveys, online questionnaires, implicit social cognition measures, and neuro-physiological measurements, among others. Despite this methodological diversity, however, the vast majority of studies have relied on artificial paradigms and self-reports to approximate access to real-world morally relevant phenomena. While these methods have proven immensely useful, widespread reliance on them has motivated concern about the external validity of morality research. For example, Bauman et al. (2015) question the degree to which responses to moral judgment measures that use extreme scenarios actually correspond to real-world moral functioning; and other research suggests that conventional measures of moral utilitarianism (e.g. Greene et al., 2001; Greene, Morelli, Lowenberg, Nystrom, & Cohen, 2008) might actually be measuring non-moral or even immoral dimensions, rather than genuine utilitarian moral concerns (Bartels & Pizarro, 2011; Kahane et al., 2015).

Despite these criticisms, we believe that artificial paradigms and self-reports have been and will continue to be valuable tools for probing moral phenomena. However, we also believe that their value should not obscure their shortcomings. As has been widely observed, theories based on self-report measures–particularly those characterized by low ecological validity–need to be carefully vetted for external validity (Cronbach, 1949; Cronbach & Meehl, 1955; Messick, 1995; Allen & Yen, 2001). Unfortunately, rigorous external validity tests of moral theories have been infrequent, likely due to the considerable difficulty of accessing moral phenomena through alternative methods (Graham, 2014; Ginges, Atran, Sachdeva, & Medin, 2011; Hoffman et al., 2014). Thus, while we are not advocating for researchers to stop using traditional measurement methods to study morality, we believe the general absence of alternative methods that can counterbalance the weaknesses of traditional measures is problematic. Such a counterbalance can be at least partially provided by big data analytics, which we believe can help validate traditional measures and theories.

However, big data analytics are useful for much more than validation. They can also provide researchers ways to access dimensions of moral phenomena that traditional methods cannot reach at scale. For example, despite many notable differences, many contemporary psychological theories of morality converge on the view that morality emerges from a complex, recursive network of individual- and group-level influences (Haidt, 2007; Graham et al., 2013; Rai & Fiske, 2011; Fiske & Rai, 2015). While individual-level moral phenomena are generated from the moral components mentioned above, these phenomena are also influenced by social and cultural factors (Lakoff, 2002; Marietta, 2008; Dehghani et al., 2009; Koleva, Graham, Haidt, Iyer, & Ditto, 2012; Baumard, André, Sperber, 2013; Shariff et al., 2015), which in turn are influenced by individual-level factors. However, the extent to which traditional research methods can capture cross-level interactions of moral phenomena is limited. These interactions tend to occur at scales larger than can be accommodated by laboratory methods, and their temporal dynamism further complicates conventional psychological investigation. However, morality research employing computational methods such as NLP and SNA suggests that these obstacles for laboratory research can be at least partially circumnavigated via big data analytics. For example, Sagi and Dehghani (2014a) were able to measure dynamic changes in group-level moral concerns regarding The World Trade Center, the Ground Zero Mosque, and Abortion by analyzing text collected from the New York Times, the blogosphere, and transcriptions of U.S. Senate speeches, respectively. Additionally, combining NLP and SNA, Dehghani et al. (2015) demonstrate how individual moral concerns can influence group-level phenomena like social network structures.

While the application of big data analytics to morality research is in its infancy, it already seems clear that these methods can make a substantial contribution to the field. Researchers can use these methods to test established theories on data that is generated by messy, uncontrolled human behavior, which is a valuable opportunity given the historical inaccessibility of real-world morality phenomena. By providing alternative measurement methods, big data analytics can also help researchers improve the external validity of their measures. Perhaps even more importantly, however, big data analytics can help researchers study otherwise inaccessible dimensions of morality, such as changes in moral values associated with environmental and socio-ecological factors, group-level moral phenomena, and the relationship between real-world moral values and behavior. While most theories of morality at least recognize the importance of these dimensions, there has been little research that has been able to directly target them. This has left considerable gaps in our understanding of human morality. If, as we believe, the goal of moral psychology is to understand moral functioning in the real world, then researchers must begin to to fill these gaps.

4. Evidence.

For big data analytics to be useful for morality research, at least two conditions must be satisfied. Big data must contain reliable traces of moral phenomena left by human behavior, and these moral traces must have sufficient informational richness to offer genuine insights into moral phenomena. While big data analytics have only recently begun being incorporated into morality research, there is already a growing body of evidence that these conditions are amply met. Additionally, the increasing use of big-data analytics on non-moral psychological phenomena corroborates the value of these methods for psychological research (Tausczik & Pennebaker, 2010; Park et al., 2014). For example, various NLP methods have uncovered word usage patterns that predict depression (Rude, Gortner, & Pennebaker, 2004), status (Kacewicz, Pennebaker, Davis, Jeon, & Graesser, 2013), motivation (Gill, Nowson, & Oberlander, 2009), cultural epistemological orientations (Dehghani et al., Bang, Medin, Marin, Leddon, & Waxman, 2013), academic success (Pennebaker et al., 2014), political affiliation (Diermeier, Godbout, Yu, & Kaufmann, 2011; Dehghani, Sagae, Sachdeva, & Gratch, 2014), personality (Oberlander & Nowson, 2006), sentiments (Dave et al., 2003), and mental disorders (Strous et al., 2009). Further, personality researchers–who were among the first psychologists to begin rigorously incorporating big data analytics into their work–have developed measurement approaches that provide powerful insights into the links between personality and language use (e.g. Schwartz et al., 2013; Back, et al., 2010) and that possess impressive psychometric qualities (e.g. Park et al., 2014; Mehl & Robbins, 2012).

Similarly, morality researchers have successfully used techniques from machine learning, NLP, and SNA to investigate morally relevant phenomena. In one of the earliest applications of NLP to morality research, Graham et al. (2009) developed a UDD of words and word stems associated with the constructs of Moral Foundations Theory (MFT). This Moral Foundations Dictionary (MFD) was then used with an NLP program called Linguistics Inquiry and Word Count (LIWC; Tausczik & Pennebaker, 2010) to explore variations in moral concerns between liberal and conservative congregations as expressed in a corpus of sermons. Notably, their results converged with previous MFT findings. Sermons delivered in liberal churches were more associated with harm and fairness concerns, compared to those delivered in conservative churches, and sermons delivered in conservative churches were more associated with purity and authority concerns, compared to those delivered in liberal churches. In another investigation of moral value differences between liberals and conservatives, Dehghani, Sagae, Sachdeva, and Gratch (2013) used an unsupervised hierarchical generative topic modeling technique based on Latent Dirichlet Allocation (LDA: Blie, Ng, & Jordan, 2003), which enabled them to extract topics from a corpus of liberal and conservative blogs. Notably, while conventional LDA techniques have no control over what topics are extracted, the method (Andrzejewski & Zhu, 2009) employed in Dehghani et al. used small sets of words from the MFD as seeds to favor the detection of topics associated with moral concerns. Using subsequent statistical analyses to compare differences between the moral topics extracted from the liberal and conservative blogs, they found that their results were consistent with previous research on moral psychology and political ideology.

Recent research has also demonstrated that NLP can be used to test sophisticated hypotheses about individual- and group-level moral phenomena. For example, in a series of three studies, Sagi and Dehghani (2014) showed that the ‘moral loading’ of specific topics can be estimated by calculating the semantic similarity between the contexts of keywords representing topics of interest and different moral concerns. In essence, this method allows researchers to measure the moralization of specific topics throughout an entire corpus and thereby produce group-level estimations of topic-specific moral concerns. Using these estimations, researchers can test hypotheses about longitudinal changes and between-group differences in the moral loadings of topics of interest. For instance, across three studies Sagi and Dehghani (2014) used this method to test hypotheses about the moral loadings of three different topics: the World Trade Center, the Ground Zero Mosque, and Abortion. In their first study, they used a corpus of 1.8 million New York Times articles dating from January 1987 to June 2007, to test the hypothesis that major events can precipitate lasting changes in moral rhetoric. More specifically, they predicted that the 9/11 attack on the World Trade Center led to significant increases in journalists' use of moral harm and ingroup rhetoric associated with the World Trade Center, but not with the Empire State Building, which was used as a control topic. Consistent with their hypothesis, they found that harm and loyalty concerns associated with the World Trade Center increased dramatically following 9/11, but that similar concerns associated with the Empire State Building remained relatively low. In their second study, Sagi and Dehghani predicted that moral concerns about the Cordoba Muslim Community Center in NYC—popularly referred to as the "Ground-Zero Mosque"—would increase sharply during the highly politicized debates that swept through the blogosphere in 2010, but that this moralization would decrease as the debates dwindled. Their results supported both predictions, indicating that NLP methods can be used to measure dynamic longitudinal patterns in moral rhetoric. In their final study, they explored differences between Democrat and Republican moralization of abortion by analyzing transcripts of nearly 230,000 U.S. Senate speeches. As predicted, they found that Republicans exhibited higher moral loadings than Democrats across all 5 MFD dimensions. Notably, their results converged with perceptions of both parties' stances on abortion: Democrats were most concerned about fairness, while Republicans were most concerned about purity. Unexpectedly, Sagi and Dehghani also found that harm concerns—which seem deeply incorporated into conservative stances on abortion—were only the third-highest loading moral dimension. However, noting that the purity dimension is represented by keywords such as abstinence, celibacy, and prostitution, Sagi and Dehghani propose that these results indicate that while Republicans endorse sanctity-of-life arguments when debating abortion, they are more often concerned with the relationship between abortion and sexual purity. In sum, these studies demonstrate that NLP can be used to test precise hypotheses about group-level moral phenomena, as well as uncover potentially counterintuitive patterns in moralization, such as the apparent primacy of purity concerns within Republican stances on abortion.

In addition to detecting patterns in group-level moral phenomena, big data analytics have been used to conduct novel explorations into the relationship between moral values and behavior. For example, Boyd et al. (2015) used a topic modeling technique called the meaning extraction method (MEM; Chung & Pennebaker, 2008) to investigate values and behaviors that emerge from natural language texts. Across two studies, Boyd et al. compared estimates of participants' values generated from the Schwartz Value Survey (SVS; Schwartz, 1992) and from MEM analyses of open-ended text produced during an online survey (Study 1) and over 130,000 Facebook status updates culled from myPersonality data (Study 2; Kosinski, Stillwell, and Graepel, 2013). While the results from the MEM analysis converged somewhat with the SVS measures, the correlations between values-relevant topics extracted by the MEM and the SVS dimensions were generally low, which Boyd et al. interpreted as suggesting that people's natural language expressions of their core values do not necessarily conform to the theory-driven set of values measured by the SVS. Finally, after comparing the SVS and MEM values measurements, Boyd et al. investigated the degree to which they could predict everyday behaviors. Notably, in both studies, they found that the MEM measurements showed greater predictive validity for participants' reported behaviors than did the SVS measurements, suggesting that the expressions of values contained in people's everyday language might actually provide more information about their behavior than traditional self-report methods.

Dehghani et al. (2015) also used NLP measurements of moral values to predict behavior. Specifically, they investigated the idea that moral homophily (love of the same) plays a prominent role in the formation of network structures. Their hypothesis was that the distance between two people in a social network could be predicted by the differences in the moral purity loadings of their messages. To test this hypothesis, they used the same adapted LSA method applied in Sagi & Dehghani (2014) to estimate the moral loading of tweets collected from 188,467 Twitter users. They then generated a model of the network structure connecting these users and calculated the distance between them. Finally, using a series of statistical tests, they explored the degree to which differences in moral foundation-related concerns predict social network distance. Supporting their hypothesis, they found a strong association between purity difference and network distance and, importantly, they also found that purity loading difference was the most accurate predictor of network distance, compared to the loadings of other moral concerns. Dehghani et al. then replicated this finding experimentally by manipulating participants' perceptions of moral similarity and measuring the effect that this manipulation had on social distancing. As in their first study, moral purity difference predicted social distance preferences above and beyond all other moral foundation concerns. While the importance of moral homophily has been previously recognized by social scientists, these studies were, to our knowledge, the first to investigate *which* moral similarities drive this phenomenon.

So far, this chapter has focused primarily on the advantages of big data analytics for moral psychology research. However, while we believe these methods can be immensely useful for researchers, we also recognize that the full extent of this usefulness remains an open question–there is still much left to discover about both the value and the limitations of big data. Accordingly, in addition to revealing and exploiting the insights available through big data analytics, future research must also focus on uncovering the boundaries of this insight. Some specific goals should be developing a better understanding of how sampling biases affect big data and how social media platforms affect user behavior (Ruths & Pfeffer, 2014). It will also be vital for researchers to critically test assumptions about correspondences between social media behavior and real-world behavior. For example, Lewis, Gray, and Meierhenrich (2014) note that while there has been much speculation about the relationship between social media and civic engagement, there has been little empirical investigation of this relation. Further, they found that while social media is a powerful tool for forming groups around civic causes, group affiliation does not necessarily predict more meaningful civic behaviors, like making financial donations to causes. While big data offer an unprecedented window into human behavior, they are nonetheless vulnerable to many of the issues that distort the relation between other forms of data and the phenomena they purport to measure. This does not negate our contention that big data contain reliable and informationally rich traces of moral phenomena; however, it does highlight the importance of testing inferences drawn from big data, as well as the necessity of developing analytical protocols that can account for issues like population and selection biases.

5. Extension and Expansion.

Moral psychology holds that morality is a fundamental component of human psychology and that the social sphere is both permeated and partially structured by moral phenomena. However, there has been very little opportunity and, relatedly, very few attempts to investigate this directly. We know at least a little, and perhaps quite a lot, about moral functioning in the laboratory, and potentially much less about moral functioning in the world (Graham, 2014; Hofmann et al., 2014). Of course, the problem of investigating psychological phenomena in natural environments has been the Achilles heel of psychology since its inception. In part, this problem has been driven by a simple lack of data. While the availability of big data and the advent of big data analytics definitely does not resolve the problem, they do offer a partial solution. By complementing traditional methodologies with theoretically-driven big data analyses, researchers can dramatically increase the verisimilitude of theories about real-world moral functioning.

In addition to advancing basic research, coupling big data analytics with theories about the moral-psychological factors that influence social behavior will enable morality researchers to contribute substantive insights into real-world events. Social media analysis is already widely incorporated in predictive social and political forecasting models. These models have shown promising potential to predict crime (e.g. Wang, Gerber, & Brown, 2012; Gerber, 2014), electoral outcomes (e.g. Unankard, Li, Sharaf, Zhong, & Li, 2014; Franks & Scherr, 2015) and stock market trends (e.g. Bollen, Mao, & Zeng, 2011), for example. However, contemporary forecasting models generally do not attempt to account for the role of moral phenomena in human behavior. As morality researchers, we believe this is a grievous oversight. For instance, recent work in social and cognitive psychology suggests that sacred moral values are important motivators of political, social, and religious extremism and violence (Atran & Ginges, 2012; Dehghani et al., 2010), voting behavior (Caprara, Schwartz, Capanna, & Vecchione, 2006; Franks & Scherr, 2015; Johnson et al., 2014), and charitable giving (Aquino & Freeman, 2009) and that they can emerge from the use of moral rhetoric (Dehghani et al., 2010; Frimer, Aquino, Gebauer, Zhu, & Oakes, 2015). As Dehghani et al. (2014a) point out, researchers can use theoretically informed big data analytics to examine dynamic morality phenomena and thereby derive insights into the moralization of specific issues, as well as to help predict when “rational actors” become “devoted actors” (Atran, 2006). This kind of real-world predictive modeling will be doubly valuable for morality researchers. Not only can it help illuminate current events, but it also enables researchers to evaluate moral psychology theories based on the degree to which they can predict human behavior in the wild–the gold standard for psychological science.

Works Cited

Allen, M. J., & Yen, W. M. (2001). *Introduction to measurement theory*. Waveland Press.

Andrzejewski, D., & Zhu, X. (2009). Latent Dirichlet allocation with topic-in-set knowledge. In Proceedings of the NAACL 2009 Workshop on Semi-Supervised Learning for NLP (pp. 43–48). Stroudsburg, PA: Association for Computational Linguistics.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science, 21,* 372–374. http://dx.doi .org/10.1177/0956797609360756

Bartels, D. M., & Pizarro, D. A. (2011). The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition, 121*, 154–161.

Baumard N, André JB, & Sperber D (2013) A mutualistic theory of morality: The evolution of fairness by partner choice. *Behavioral and Brain Sciences 36,* 59–122

Blei, D. Probabilistic topic models. *Communications of the ACM, 55*(4):77–84, 2012.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *The Journal of Machine Learning research, 3,* 993–1022.

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, *2*(1), 1-8.

Boyd, R. L., Wilson, S. R., Pennebaker, J. W., Kosinski, M., Stillwell, D. J., Mihalcea, R. (2015) Values in Words: Using Language to Evaluate and Understand Personal Values. *International AAAI Conference on Weblogs and Social Media*. Retrieved from https://www.aaai.org/ocs/index.php/ICWSM/ICWSM15/paper/view/10482

Clifton, A., Turkheimer, E., & Oltmanns, T. F. (2009). Personality disorder in social networks: Network position as a marker of interpersonal dysfunction. *Social Networks*, *31*(1), 26–32. http://doi.org/10.1016/j.socnet.2008.08.003

Cronbach, L. J. (1949). *Essentials of psychological testing*.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests.*Psychological bulletin*, *52*(4), 281.

Cushman, F., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social neuroscience*, *7*(3), 269-279.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science, 41*, 391–407.

Dehghani, M., Johnson, K. M., Sagi, E., Garten, J., Parmar, N. J., Vaisey, S., Iliev, R., & Graham, J. (2015). Purity homophily in social networks. Manuscript submitted for publication.

Dehghani, M., Iliev, R., Sachdeva, S., Atran, S., Ginges, J., & Medin, D. (2009). Emerging sacred values: Iran’s nuclear program. *Judgment and Decision Making, 4*, 930–933.

Dehghani, M., Sagae, K., Sachdeva, S., Gratch, J. (2014). Analyzing Political Rhetoric in Conservative and Liberal Weblogs Related to the Construction of the "Ground Zero Mosque". Journal of Information Technology & Politics. Vol. 11 (1), pp. 1-14.

Dostert, L. E. (1955). The georgetown-ibm experiment. *1955). Machine translation of languages. John Wiley & Sons, New York*, 124-135.

Dumais, S. T. (2004). Latent semantic analysis. Ann. Rev. Info. Sci. Tech., 38: 188–230. doi: 10.1002/aris.1440380105

Fiske, A. P., & Rai, T. S. (2015). *Virtuous violence: Hurting and killing to create, sustain, end and honor social relationships.*

Gerber, M. S. (2014). Predicting crime using Twitter and kernel density estimation. *Decision Support Systems*, *61*, 115-125.

Gibson, J. J. (1977). The theory of affordances. *Hilldale, USA*.

Ginges, J., Atran, S., Sachdeva, S., & Medin, D. (2011). Psychology out of the laboratory: the challenge of violent extremism. *American Psychologist*, *66*(6), 507.

Graham, J. (2014). Morality beyond the lab. Science, 345(6202), 1242-1242.

Graham, J., & Iyer, R. (2012). The unbearable vagueness of “essence”: Forty-four clarification questions for Gray, Young, & Waytz. *Psychological Inquiry, 23*, 162-165.

Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S., & Ditto, P. H. (2013). Moral Foundations Theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology, 47,* 55-130.

Graham, J., Haidt, J.,& Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96,* 1029–1046.

Graham, J., Meindl, P., & Beall, E. (2012). Integrating the streams of morality research: The case of political ideology. *Current Directions in Psychological Science*, *21*, 373-377.

Gray, K., Schein, C., & Ward. A. F. (2014). The myth of harmless wrongs in moral cognition:Automatic dyadic completion from sin to suffering. *Journal of Experimental Psychology: General*, *143,* 1600-1615.

Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition, 107*, 1144–1154. doi:10.1016/j.cognition.2007.11.004

Haidt, J. (2007). The new synthesis in moral psychology. *Science*, *316*(5827), 998-1002.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. Vintage.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?

Behavioral and Brain Sciences, 33, 61-83.

Iliev, R., Dehghani, M., Sagi, E. (2014). Automated Text Analysis in Psychology: Methods, Applications, and Future Developments. *Language and Cognition*. 1-26.

Johnson, K. M., Iyer, R., Wojcik, S. P., Vaisey, S., Miles, A., Chu, V., & Graham, J. (2014). Ideology-specific patterns of moral indifference predict intentions not to vote. *Analyses of Social Issues and Public Policy, 14,* 61-77. [supplements]

Jones, K. S. (1994). Natural language processing: a historical review. In *Current issues in computational linguistics: in honour of Don Walker* (pp. 3-16). Springer Netherlands.

Kahane, G. (2015). Sidetracked by trolleys: Why sacrificial moral dilemmas tell us little (or nothing) about utilitarian judgment. *Social neuroscience,* (ahead-of-print), 1-10.

Kahane, G., Everett, J. A., Earp, B. D., Farias, M., & Savulescu, J. (2015). 'Utilitarian'judgments in sacrificial moral dilemmas do not reflect impartial concern for the greater good. *Cognition*, *134*, 193-209.

Kemp, S. (2015, January 21). Digital, Social & Mobile Worldwide in 2015. [Blog post]. Retrieved from http://wearesocial.net/tag/sdmw/

Koleva, S., Graham, J., Haidt, J., Iyer, R., & Ditto, P. H. (2012). Tracing the threads: How five moral concerns (especially Purity) help explain culture war attitudes. *Journal of Research in Personality, 46*,184–194.

Lakoff, G. (2002). *Moral politics: How liberals and conservatives think*. Chicago, IL: University of Chicago Press. Marietta,

Lancichinetti, A., Sirer, M. I., Wang, J. X., Acuna, D., Körding, K., & Amaral, L. A. N. (2015). High-Reproducibility and High-Accuracy Method for Automated Topic Classification. *Physical Review X*, *5*(1), 011007.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211.

Lewis, K., Gray, K., & Meierhenrich, J. (2014). The Structure of Online Activism. *Sociological Science, 1*(February), 1–19. http://doi.org/10.15195/v1.a1

Marietta, M. (2008). From my cold, dead hands: Democratic consequences of sacred rhetoric. *Journal of Politics*, 70, 767–779.

Marin, A., & Wellman, B. (2011). Social network analysis: An introduction. *The SAGE handbook of social network analysis*, 11-25.

Meindl, P., & Graham, J. (2014). Know thy participant: The trouble with nomothetic assumptions in moral psychology. In H. Sarkissian and J. C. Wright (Eds.), *Advances in Experimental Moral Psychology* (pp. 233-252). London: Bloomsbury.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In*Advances in Neural Information Processing Systems* (pp. 3111-3119).

Nadkarni, P. M., Ohno-Machado, L., & Chapman, W. W. (2011). Natural language processing: an introduction. *Journal of the American Medical Informatics Association : JAMIA*, *18*(5), 544–551. doi:10.1136/amiajnl-2011-000464

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H., & Seligman, M. E. P. (2014). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology*.

Pennebaker , J. W. (2011). T*he secret life of pronouns: what our words say about us*. New York : Bloomsbury Press.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2003). Linguistic inquiry and word count: LIWC 2001 manual. Mahwah, NJ: Erlbaum.

Pennebaker JW, Chung CK, Frazee J, Lavergne GM, Beaver DI (2014) When Small Words Foretell Academic Success: The Case of College Admissions Essays. *PLoS ONE 9*(12): e115844. doi: 10.1371/journal.pone.0115844

Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: moral motives for unity, hierarchy, equality, and proportionality. *Psychological review*, *118*(1), 57.

Ruths, D., & Pfeffer, J. (2014). Social media for large studies of behavior.*Science*, *346*(6213), 1063-1064.

Sagi, E., Dehghani, M. (2014a). Measuring moral rhetoric in text. *Social Science Computer Review. Vol. 32* (2), pp. 132-144.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theory and empirical tests in 20 countries (pp. 1–65). In M. Zanna (Ed.), *Advances in experimental social psychology*. New York: Academic Press.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS one*, *8*(9), e73791.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Park, G., Sap, M., Stillwell, D., Kosinski, M., Ungar, L. H. (2014). Towards Assessing Changes in Degree of Depression through Facebook. ACL 2014 Workshop on Computational Linguistics and Clinical Psychology, 118.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology, 29*, 24–54.

Unankard, S., Li, X., Sharaf, M., Zhong, J., & Li, X. (2014). Predicting Elections from Social Networks Based on Sub-event Detection and Sentiment Analysis. In *Web Information Systems Engineering–WISE 2014* (pp. 1-16). Springer International Publishing.

Vaisey, S., & Lizardo, O. (2010). Can Cultural Worldviews Influence Network Composition? *Social Forces*, *88*, 1595–1618.

Vaisey, S., & Miles, A. (2014). Tools from moral psychology for measuring personal moral culture. *Theory and Society*, 311–332. http://doi.org/10.1007/s11186-014-9221-8

Wang, X., Gerber, M. S., & Brown, D. E. (2012). Automatic crime prediction using events extracted from twitter posts. In *Social Computing, Behavioral-Cultural Modeling and Prediction* (pp. 231-238). Springer Berlin Heidelberg.