Textual Analysis of Narratives in the Context of a Public Goods Experiment

Simon Bartke

1 Introduction

Narratives shape individual thinking and motivate action. Individuals understand themselves, others and their environment by telling narratives about events they experienced (Conway and Holmes, 2004; McAdams, 2001, 2006, see Holland and Kensinger, 2010 for a review). Narratives motivate behavior by offering cues for solving problems, offering heuristics derived from past experiences about dealing with new situations and presenting goals for future behavior (Pillemer, 2003). By means of a laboratory experiment, this study analyzes narratives within *i)* a public goods game context and *ii)* within the context of subjects' self-reported motivational states after having been exposed to their recalled narratives. To this end, we focus on narratives in the sense of stories that subjects wrote down from autobiographic memory. In particular, subjects were assigned to one of three conditions that determined what narratives they should recall. They either had to narrate memories focusing on compassionate and prosocial motivation (*Care*), or narratives focusing on angry and antisocial motivation (*Anger)* or write down a narrative about a motivationally unspecific experience *(Control)*. Based on motivation psychology, the concrete topics of the treatments have clear predictions in terms of how subjects exposed to them should behave in a subsequent public goods game. Likewise, motivation psychology delivers hypotheses how recalling the narratives in the different treatments impacts subjects' motivational states that we measure by means of subjects' ratings on motivational scales. Therefore, this experiment provides a causal investigation how narratives and which aspects of them affect behavior and motivation.

In order to analyze narratives, three approaches are compared. We investigate whether the presence of specific words (single marker words) by means of the dictionary method within narratives have a significant impact on motivational and behavioral reactions. Likewise, we analyze the relation between motivational content rating / assessment of motivational intensity within narratives by human assistants and subsequent behavioral and motivational responses. The results from the previous two approaches are compared with those form an unsupervised statistical learning algorithm, Latent Dirichlet Allocation (LDA, Blet et al., 2003) which synthesizes underlying topics in narratives. A controlled laboratory study allows to study this impact which is usually difficult to observe in everyday life. This study therefore adds to the literature in two domains: we demonstrate how the effect of narratives on behavior and motivation can be analyzed in a quantitative way within a controlled setting and make a methodological contribution by showcasing the potential of LDA for the purpose of analyzing narratives within experimental economics compared to other textual analysis methods. To the best of our knowledge, this paper is the first to apply text analysis by means of LDA to a context relevant to experimental economics.

This study uses response variables from two parts: one part in which subjects wrote narratives and subsequently indicated their motivational states and another part in which subjects also initially wrote down narratives and then made decisions in a public goods game. The first part serves to answer the question which dimensions of narratives affect the motivational state of subjects in a hypothesized direction. The second part yields insights which aspects of narratives are responsible for the behavioral impact of narratives. Furthermore, we can observe how different means of analyzing narratives compare based on different response variables. Specifically, we can compare the unique human skill to read between the lines, emotionally respond and emphasize with a narrative with the automated approaches of the dictionary method and LDA.

In the context of a text corpus of narratives, LDA posits that every single narrative from within the corpus is composed of K topics. The number of topics, K, that are found by LDA is selected as a parameter by the user of LDA. All K topics describe each document with a positive probability. A topic is defined as a distribution over the unique words of the population of all words that were used in the narratives. Individual words appear in different topics with different probabilities. Therefore, LDA yields quantitative variables that express how much of topic x was written in narrative y for every narrative in the corpus. Exemplarily, narratives investigated in this study may all consist to some degree of the motivationally unspecific topic of doing routine day-to-day activities, but the shares that this topic covers vary by narrative. In this example, words about taking a shower have high probability under the everyday activities topic and lower probability under a *Care*- or *Anger*- related topic.

A lot of decisive social problems challenging the global community (such as carbon emissions abatement and reduction of systemic financial risks) have a public good character. Public goods are a social dilemma: groups of individuals face a conflict between the maximization of individual gains and the collective interest. To the extent that beneficiaries of such global public goods do not contribute enough for their provision, a free-rider problem exists. It is therefore evident that different interest groups use and construct narratives to influence attitudes and actions towards global public goods. Hence, narratives are part of the context that influences whether provision of public goods does or does not occur. Exemplarily, some narratives regarding climate change (Pancost, 2017) seek to stress that climate change is not-existent, not man-made and that researchers themselves contradict each other when they discuss causes and consequences of climate change. A different narrative on the same topic might stress that research on causes and consequences of climate change is clear on the point that a reduction in CO_2 emissions is warranted, might stress that rise in atmospheric CO_2 and temperature correlates almost perfectly with the industrial age and is increasing ever since and outline horror scenarios such as biodiversity loss, melting of ice and flooding of whole nations to make their point. Such narratives are often strategically developed with great care, seen as an instrument of "soft power" in international relations and repeated by respective spokespersons at various venues and outlets (Ganz, 2011; Lowe et al., 2006; Nye Jr., 2013; Nye Jr., 2008; Roselle et al., 2014).

Narratives are naturally linked to text as data. Recent developments within economics use quantitative representation of text as data in their studies. Working with text as data has become feasible due to novel technological developments and received considerable interest by social scientists lately (Gentzkow et al., 2017; Einav and Levin, 2014). Within experimental economics, the focus of analyzing texts has treated mostly communication data and relied largely on coding schemes that human assistants were trained on to code communication content (exemplarily, Brandts and Cooper, 2007). While this approach seeks to analyze communication and narratives in a systematic way and uses the unique human ability of "getting a feeling" for the text corpus, it is subjective and often not blind to the hypotheses of the study, prone to errors due to human processing as well as resource intensive since assistants need to get paid and need a fair amount of time to rate texts.

Overall, topics within narratives as well as words used in them have a decisive impact on later motivational states as well as behavior, beliefs and perceptions in the public goods game. We find that both automated approaches, LDA and dictionary method, outperform human coding of motivation in revealing connections between narratives and behavior in the public goods game. We show that LDA can be used successfully within experimental economics for automated text analysis. LDA topics are systematically associated with both behavior in the public goods game as well as self-reported motivational states. In particular, we find that the identical LDA topics that are associated with increases in motivational states also predict behavior in the public goods game according to predictions from motivation psychology which indicates the presence of convergent validity. In line with predictions from motivation psychology we find that one topic in particular[1], topic 4: "helping others in public", has a significant and positive influence on subsequently self-reported states of compassion motivation and at the same time a positive effect on contributions and beliefs in the public goods game. We do not find this degree of convergent validity for the dictionary method. The only word category that indicates convergent validity between motivation and behavior given the experimental treatments is the anger words category. The use of anger words increases anger motivation and decreases attention that subjects pay to the payoff of the other group member in the public goods game. Several other word categories have an isolated impact on subsequent behavior which we discuss in the results below. For the dictionary method, we can only partly confirm the marker word hypothesis which states that one can infer the presence of motivation from the presence of single groups of words within texts. On this regard, we find that anger words do predict self-reported anger motivation but that no word category related to compassion is meaningfully associated with compassion motivation. The only reasonable result we find for human coding of motivations in narratives is that the higher the human assessment of anger within a narrative, the less likely subjects are to be a conditional cooperator in the public goods game. Our results indicate that insights from motivation psychology are essential for interpreting the results from the dictionary method and LDA methods and therefore to a quantitative analysis of the impact of narratives in general.

Section 2 of this paper briefly defines the concept of narrative used within this paper and introduces some relevant literature that most economists should not be acquainted to. Section 3 presents the design and behavioral as well as motivational results of the experiment. Section 4 proposes three approaches to analyze narratives within the texts and presents hypotheses as well as results for each approach. Section 5 investigates the robustness of the results and relates them to one another. Section 6 provides a concluding discussion.

2 Definition of Narrative Concept and Related Literature

2.1 Narratives and Autobiographic Memory

For this study, we define narratives as formulations of stories about personal episodic memories that are interpreted and created by the current self (Larsen, 1992).[2] One of narratives' core

[1] We find more LDA results that are shown and discussed below but limit the exposition in the introduction for brevity.

[2] It should be noted that psychologists consider narratives of this kind under the theory of „narrative identity“ (McAdams, 2001). Evidence has accumulated that the presence and

functions at the individual level is to construct a sense of self in a coherent way (Ross, 1989). Narratives help us to draw experiences from situations in the past to guide present and future reasoning and behavior (Bluck et al., 2005). Narratives and autobiographic memories are clearly interlinked due to their interdependent influence on each other (Pasupathi and Hoyt, 2009). It was found that narratives transform memories about situations that one experienced due to two reasons: Memories need to be narrated such that they fit into a coherent self-image and therefore the narrated memories have to follow a narrative format. Moreover, narrating memories occurs in a communication context. In order to make memories comprehensible for other persons, the representations of memories are changed such that others can understand them (Smorti and Fioretti, 2016). It was found that when humans are offered to tell personal stories, they implicitly convey details about their reasoning and choice process (O'Conner, 2000) which is of particular interest to this study. Narratives further allow us to better understand and empathize with others (Cohen, 1998). Furthermore, narratives serve a social purpose since they help engage and preserve social bonds by providing material for conversation. Ultimately, every known narrative, be it about the own identity, a political ideology or a scientific school of thought, once originated in the mind of one or a few individuals (Mokyr, 2016).

Evidently, the construction and communication of narratives is fundamental to the human nature which should make narratives a subject of study for economists. Many agents such as politicians, think tanks, historians or the media use narratives in order to influence opinions and debates (Turowski and Mikfeld, 2013; Berger, 1997; Alexander, 2017). Examples of such narratives are the "rags to riches", "American Dream" or "Space Race" narratives. Narratives are a powerful way of fixing and communicating ideas like the "market for lemons" narrative that George Akerlof used to illustrate the consequences of information asymmetries (Akerlof, 1978). Other disciplines in the social sciences such as history, sociology, anthropology and psychology have already put a bigger focus on studying narratives than the discipline of economics has; a situation which, as some economists argue, should change (Shiller, 2017; Wydick, 2015). Studying narratives could yield particularly interesting insights into what caused major changes in economic and political conditions like recessions, bursts of asset price bubbles or the Fall of the Wall in Germany in 1989.

---

organization of memories is essential such that an identity represented by a narrative can be formed (Habermas and Bluck, 2000).

2.2 Narratives in economics:

Within economics, few studies exist that study the ability of narratives to influence and even initiate changes in economic conditions (Shiller, 2017; Akerlof and Snower, 2016; Akerlof and Shiller, 2015). Rodrick et al. (2003) use growth theory and combine them with country experiences to shed light on unsolved riddles in the economic growth literature through the lens of country narratives. Falk and Tirole (2016) model narratives to be held up as justifications for immoral behavior. Agents that behave unethically use narratives to exaggerate the cost of behaving ethically or downplay the externalities of behaving unethically or showing themselves as not pivotal in the context of morally questionable decisions. Glaeser (2005) modeled political actors as spreading stories to their advantage. These stories do not gain credibility from being true but from mere repetition. Pöder (2010) provides an innovative example how analytical narratives can be constructed from historical studies to explain the provision of lighthouses in Estonia over time and changing incentives of public good provision. However, the quantitative study of the impact of narratives is still underrepresented. Generally, it is hard to determine how and when narratives impact decisions. Shiller (2017) suggests that textual analysis studies have great potential in studying effects of narratives quantitatively. Recently, this approach became more feasible due to the onset of big data and the availability of better methods of automated natural language processing like LDA.

2.3 Textual Analysis Studies in (Experimental) Economics

Given recent methodological developments in the area of studying text as data in a quantitative way, textual analysis has been increasingly applied to economics research questions lately. Exemplarily, Kuziemko and Washington (2015) perform a quantitative analysis featuring text data from newspapers, to investigate how racial views of whites in the southern USA are responsible for a dramatic change in voting behavior of this group since the 1960s. Through the analysis of Congressional speeches from 1873 to 2009 with machine learning methods focusing on classification, Gentzkow et al. (2016) show that political partisanship increased immensely since the 1990s. Boudoukh et al. (2013) use textual analysis to infer the type and tone of news. They show that once this information is accounted for, changes in news that reveal changes in fundamentals show a significant association with stock price movements (see also Nyman et al. (2018)). Based on the computational analysis of economic and financial newspaper articles, Shapiro et al. (2018) create time series of sentiments from 1980 to 2015 and analyze their correlations with the business cycle. They find that their sentiment indices improve forecasting performance and are significantly related to consumer sentiments.[3]

Within experimental economics, the object of textual analysis was almost exclusively communication data. For this purpose, the experimenter typically hires assistants who code or categorize the communication that occurred between participants in an experiment. The coding scheme / a priori categories are usually previously developed by the experimenter for the study. Xiao and Houser (2005) find that costly punishment and expression of emotion towards the sender in an ultimatum game in the case of unfair offers are substitutes. In order to evaluate the content of emotion expression messages, the authors hired student assistants who had to classify the messages into categories and where incentivized to coordinate on a concordant category for each message. In a principal-agent relationship, Brandts and Cooper (2007) investigate how communication by the principal can influence agents' coordination on providing high effort levels. High effort levels can be stimulated by explicitly demanding high efforts and highlighting the benefits for the agents and the principal when high efforts are provided. A very similar approach[4] is employed by Cooper and Kagel (2005) to analyze intra team communication in the context of studying cooperation in signaling game experiments. The authors use team communication to infer insights how players learn to play more strategical. Xiao and Houser (2011) compare different methods of classifying communication methods in the context of a coordination game. These authors compare a coordination-game message classification procedure with classification according to a coding scheme developed by the investigators and author's self-classification. They find that the coordination game for communication classification to be superior to the other two forms, but also note that it is resource intensive in terms of money and time and is likely to reach capacity limits when the amount of messages becomes large. Moellers et al. analyze how profit shares and efficiency varies in interactions between a monopolist and an oligopoly of downstream firms that compete when the possibility to communicate is altered. In this research, the authors let coders categorize the communication similar to Houser and Xiao (2011) and also mine the communication for keywords.

[3] There are a few other recent studies that employ a similar strategy to a similar application: Tetlock (2007), Tetlock et al., 2008, Loughran and McDonald (2011), Heston and Sinha (2015), Fraiberger (2016) and Young and Soroka (2012).

[4] Likewise, Charness and Dufwenberg (2006, 2009), Chen and Chen (2011), Sutter and Strassmair (2009), Andreoni and Rao (2011) use own content coding regimes together with research assistants to analyze communication.

So far, studies in economics that used textual analysis have focused mostly on keyword retrieval or developed own content coding schemes. However, there is a further approach in quantitative textual analysis that is widely used in social sciences: counting words in texts from pre-specified lists (*dictionary method)*. Within this method, the LIWC dictionary (Pennebaker et al., 2001) has received considerable attention.[5] The LIWC dictionary stems from the field of psychology and was among other things previously used to detect motivations within texts (Schultheiss, 2013). Also, several studies within experimental economics have recently used the LIWC dictionary for quantitative textual analysis (Abatayo et al., 2018; Chen and Chen, 2011; Babin, 2016). Penczynski (2016) employs machine learning methods to show that human classification of experimental communication can be reliably replicated by machine learning methods. He points out that the adoption of his automated classification approach enables a fast and more objective classification of especially large-scale communication data at low costs.

2.3.1 LDA as a Textual Analysis Method in Economics

LDA is able to reveal underlying themes in unlabeled text data without receiving any structural information or evaluation criteria about the respective text data beforehand. LDA as a method should therefore enable many research endeavors that are of interest to economists. As an example, Hansen et al. (2017) study how transparency of central bank communication affects deliberation of monetary economists. The authors exploit data from a natural experiment and show that the release of minutes from meetings of central bankers has two effects: it disciplines the discussion but also leads to more conformity, with the latter effect dominating the former and hence an advantageous overall effect (in a similar vein, see Arango et al. (2017)). Thorsrud (2016) uses LDA for a nowcasting purpose. He uses LDA to infer latent topics from a business newspaper in order to nowcast quarterly GDP with these topics. Another application of LDA can be found in Budak et al. (2014) who study the effects of do-not-track policies on the Internet economy. They use LDA to tie retailers to specific market segments that they operate in and to link content providers into categories such as games, news, automotive and many more. Nimark and Pitschner (2016) make use of LDA to categorize news articles into overarching topics. They find that in general, heterogeneity exists in the topics that different newspapers choose to report on, but that the topic selection becomes more homogeneous after major events and link this observation to a theoretical model. Finally, Bandiera et al. (2017) make use of

[5] See section 4.1 for more details.

LDA to map typical behaviors to distinct CEO management styles. They show that especially in low/middle income countries there exists a discrepancy between the management style of the CEO and the necessities of the firm she manages that leads to a productivity loss accounting for 13% of the productivity gap between low- and high-income countries. At present, we are not aware of any other study within experimental economics that previously used LDA.

3 Experimental Data and previous Results

This section follows from sections 2.3 and 2.4 of chapter 2 of this dissertation text. Its purpose here is to introduce the data and previous results that the analysis of narratives departs from for the remainder of this study. Below, it is provided in a shorter exposition such that readers only interested in this study can comprehend the exposition below.

The recollection of the narratives was supposed to activate different motivations depending on the treatment. It was argued in chapter 2 of this dissertation that the activation of different motivation systems is associated with clear hypothesized behavioral patterns in the public goods game. In particular, the *Care* (*Anger*) treatment was supposed to lead to increases in self-reported compassion (anger) motivation as well as behavior, beliefs, attentional foci and perceptions in the public goods game that were significantly more prosocial (antisocial) than under *Control.* Based on the data that was collected with this experiment, we can analyze first, which parts of narratives are related to subsequent motivational state changes that subjects report. Second, we can link behavior, beliefs, attention and perceptions from the public goods game to previously written narratives in the context of a controlled experimental setting. Since subjects wrote narratives under the identical set of treatments in both parts of the experiment, the "motivational states part" and the "public goods game part", this allows to compare whether components of narratives that predict motivational state changes also have predictive power for the dependent variables in the public goods game. This increases the robustness in terms of external validity of the insights about the role of narratives in influencing motivations and behavior. Furthermore, we can compare how different methods of analyzing narrative content compare in the context of these two parts.

3.1 The "Motivational States Part"

The experiment consists of two parts. In the „motivational states part“ (henceforth ms-p), subjects first wrote down narratives based on own memories depending on their randomly assigned treatment. Subsequently, they rated the degree to which they felt a number of affect- and motivation-related adjectives described their current state. Within each treatment, subjects had to write down with pen and paper two narratives whose ordering was random for each session. In the *Care* treatment, subjects wrote about memories when they helped someone and when they felt feelings of compassion for another person. Under *Anger*, subjects wrote about memories when they were frustrated and when they were insulted or harassed. In the *Control* condition, subjects wrote about what they did yesterday and about the course of a typical day in their lives. Subjects received a fixed compensation of €4.50 for writing down the narratives and indicating their motivational states afterwards which together took around 30 minutes. Subjects indicated how much they felt different motivational and affective states with pen and paper. In particular, subjects made marks on a continuous scale ranging from "not at all" on the one side to "very much" on the other, "to which degree they feel like one of the following motivations and emotions in this very moment".[6] Each subject rated themselves in this way for 22 adjectives. These adjectives comprised words related to anger motives (5 words), compassion motives (5), fear motives (5), achievement motives (5), as well as the two feelings happy and sad. A complete list of words can be obtained upon request and was developed to be maximally representative of the indicated motivations (Chierchia et al., 2018).[7]

3.1.1 Hypotheses for the „Motivational States Part"[8]

Hypothesis 0.1: The *Care* treatment increases ratings of compassion motives significantly compared to the *Control* condition.

Hypothesis 0.2: The *Anger* treatment increases ratings of anger motives significantly compared to the *Control* condition.

3.1.2 Brief Exposition of Results for "Motivational States Part"[9]

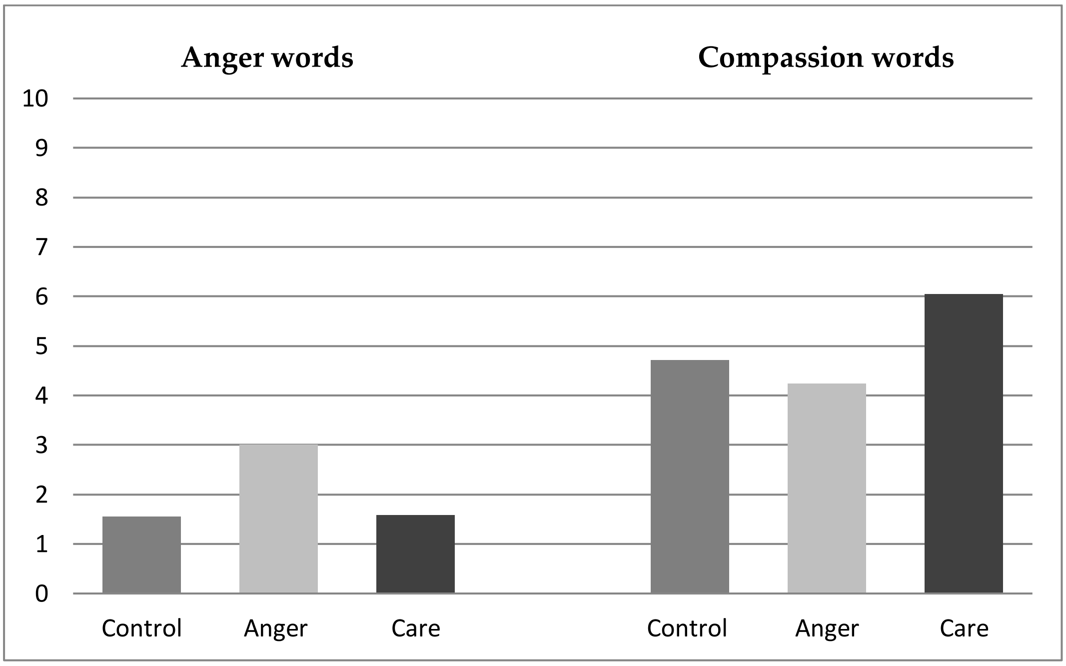
In total, 133 subjects[10] from Kiel University subject pool participated in the ms-p. Of these, 45 participated in *Control*, 41 in *Anger*, and 44 in the *Care* treatment. We compare the mean ratings of words that belong to either the compassion or anger motivational category between treatments. The *Care* treatment increases self-reported ratings of compassion words compared

[6] The rationale for choosing exactly these narrative topics is motivated by insights from motivation psychology and laid down in chapters 2.2.2.1, 2.2.2.2 and 2.3.1 of this manuscript.
[7] The procedure thus follows chapter 2.3.2 of this manuscript.
[8] Please refer to chapter 2.3.2 of this manuscript for a detailed discussion of the hypotheses for the "motivational states" part.
[9] Please refer to chapter 2.3.2 and Table 1 for a more detailed exposure.
[10] Three subjects were excluded due to a lack of proficiency in the German language.

to *Control* at p = .005. The *Anger* treatment increases self-reported ratings on anger words compared to *Control* at p < .001.[11]

Results 0.1 & 0.2: We find support for hypotheses 1 and 2: Writing down narratives targeted at compassion (anger) increases self-reported motivational states of compassion (anger) significantly compared to the control condition.

Figure A1: Results of "motivational states part"



3.2 The „Public Goods Game Part"

The second part of this experiment is the „Public Goods Game Part" (henceforth pgg-p), which was carried out in different sessions and with different subjects than the ms-p. The pgg-p consisted of two sessions that were run back-to-back. A first session, in which subjects wrote down narratives which was identical to the respective task in the ms-p and a second, in which we collected response variables from the public goods game context. We recruited subjects separately for each session. See chapter A.1 for the procedure and instructions used for the experiment.

The group size per public good was two subjects. During the experimental procedure, subjects learned at no point the identity of their other group member. The marginal per capita return (MPCR) was 0.75 for both subjects in a group. The initial private account (endowment) of each

[11] We used Somers' D, which is a variant of the ranksum test accounting for clustering at the subject level

subject i was €10. From this endowment, they had to decide how much $x_i \in \{0, \ldots, 10\}$ in whole Euro amounts to contribute to the public good. The monetary payoff for subject i was

$$\pi_i = 10 - x_i + 0.75 \times (x_i + x_j) \tag{X.1}$$

Where x_j is the contribution of $i's$ other group member j to the public good.

Subjects were exposed to example calculations and their comprehension of the public goods game was checked by means of comprehension questions to ensure that they understood the strategic incentives of the game. During this comprehension check, we report how much attention subjects devote to their own relative to their other group member's hypothetical payoff. Subsequently, the the public good contribution decision was collected, followed by the statement what subjects believed was contributed by their other group member. After that, subjects stated their perception of whether the game exhibits strategic substitutes or complements and finally decided how much to contribute to the public good in a conditional way which was elicited by means of the strategy method (Selten, 1967).[12]

3.2.1 Hypotheses for the „Public Goods Game Part“[13]

Hypothesis 0.3 – Unconditional contributions: Subjects under *Care* contribute significantly more to the public good than subjects under *Anger*, while contributions under *Control* lie between the two other treatments.

Hypothesis 0.4 – Attention: Both *Care* and *Anger* increase subjects' attentions to the others' payoffs relative to *Control*.

Hypothesis 0.5 – Perception*s: Care* subjects perceive the nature of the public goods game as significantly more cooperative than *Control* subjects. Subjects under *Anger* perceive the game as significantly more competitive than both *Control* and *Care* subjects.

Hypothesis 0.6 – Descriptive beliefs: *Care* subjects believe that their other group member has contributed more to the public good relative to this belief under *Control*.

Hypothesis 0.7 – Conditional contributions: Subjects motivated by *Care* display a higher conditional contribution for a fixed contribution of the other team member than both *Control* and *Anger* subjects.

[12] Details can be found in Appendix A.
[13] Please refer to chapter 2.3.3 of this manuscript for a detailed discussion of the hypotheses for the pgg-p.

3.2.2 Brief Exposition of Results for "Public Goods Game Part"[14]

The 184 subjects for this part came from the University of Kiel subject pool but it was made sure that they had not participated in the ms-p. 57 subjects participated in *Control*, 62 in the *Anger* treatment, and 65 in the *Care* treatment. The recruiting was carried out with the software hroot (Bock et al., 2014) and the experimental interface was programmed with z-Tree (Fischbacher, 2007). On average, subjects were paid €18.20 for their work in this part which took around 90 minutes.

The comprehension check revealed that 97 of the 184 subjects answered all comprehension questions correctly ("comprehension sample"). Strategies that subjects seek to realize are observable with least noise for those subjects that have a perfect understanding of the game's incentives. In the following, we will therefore report results for the full and the comprehension sample.

Result 0.3: Hypothesis 0.3 is confirmed in the comprehension sample only; subjects under *Care* contribute significantly more than subjects under *Anger*, while *Control* subjects contribute between the two.

Figure X Average contributions

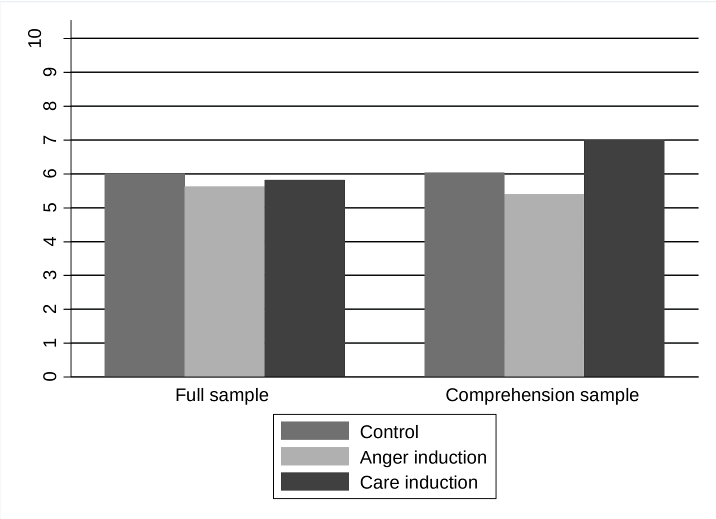


Figure X displays the mean contribution levels across the three treatments. In the full sample there are no significant differences. In the comprehension sample by contrast, average contributions are €6.03, €5.39, and €7.00 under *Control*, *Anger* and *Care* treatments,

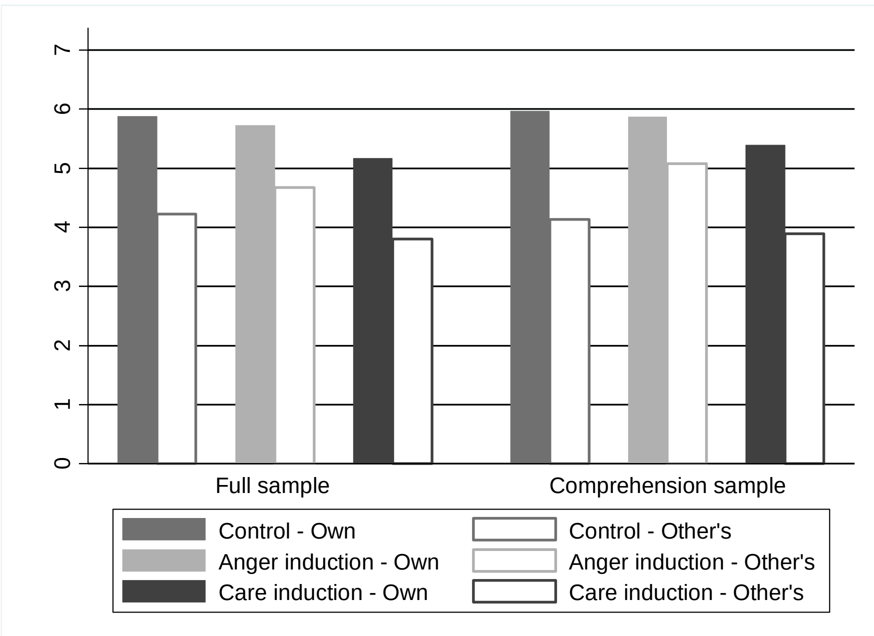
[14] Please refer to chapter 2.4 of this manuscript for a more detailed exposure.

respectively. Contributions in the care induction treatment are significantly higher than in the anger induction treatment at the p=.039 level according to a rank sum test.

Result 0.4: Different narrative topics lead to different attentional foci: Under *Care,* subjects show less self-focus, while subjects under *Anger* show more other-focus. We find support for hypothesis 0.4 in the *Anger* treatment.

Figure Z displays subject's attention to their own and other's hypothetical payoffs from the pre-game comprehension stage. The solid bars show the average number of views for subjects' own payoffs, while the outlined bars show the corresponding number of views for the other's payoff. In all treatments subjects pay more attention to their own payoffs. *Care* treatment subjects in the full sample view their own payoff significantly less often than those in the *Control* treatment (at p=.089 according to a rank sum test). No other differences are significant according to rank sum tests. Subjects under *Anger* pay the most attention to their partner's payoff, but there is a significant positive correlation (ρ=.28 in the full and ρ=.33 in the comprehension sample, both significant at 5%) between the difference in displayed payoffs for subjects in the anger induction treatment and subject's subsequent contributions, but not in either the control or care induction treatments.

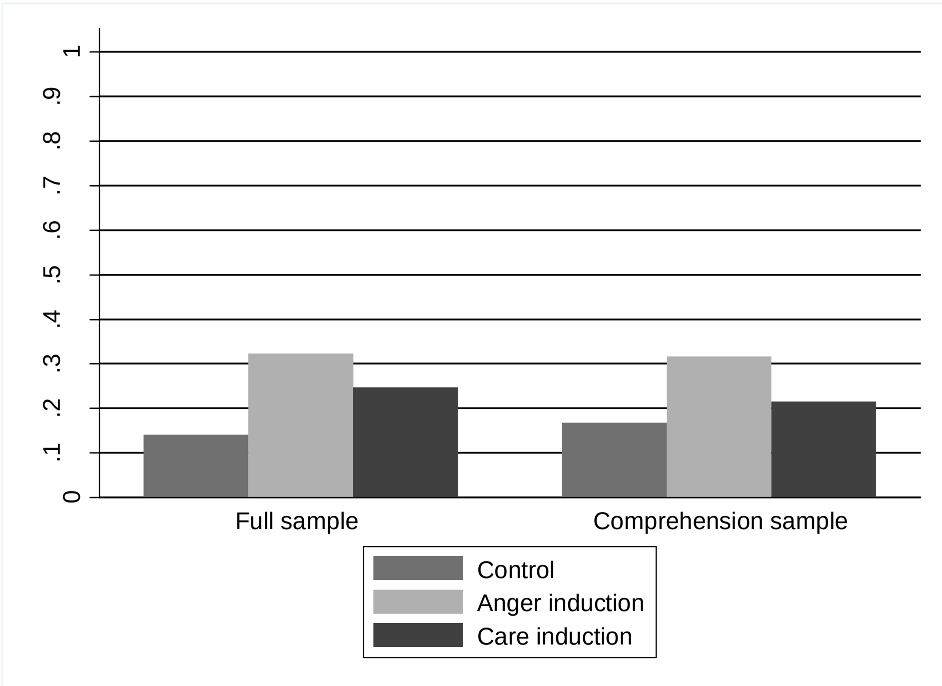
Figure Z: Average views of own and other's payoffs by treatment



Result 0.5: Evidence for hypothesis 0.5 in the *Anger* sample only: Writing narratives about *Anger* is associated with more competitive perceptions of game incentives.

Figure W displays the shares of subjects by treatment who think that the public goods game is more similar to a purely competitive (matching pennies) rather than a purely cooperative (pure coordination) game. The Anger treatment leads subjects the most often to consider the game to be more competitive than cooperative: 32% percent consider it to be more competitive in the full sample and also 32% in the comprehension sample. For the full sample, this difference is significant (at p=.02) compared to *Control*.

Figure W: Fraction of subjects reporting the public goods game to be more similar to a purely cooperative or purely competitive game



Result 0.6: Subjects' elicited descriptive and normative beliefs do not significantly differ across treatments, irrespective of the considered sample. We find a trend that subjects under *Care* have higher beliefs and slightly higher normative expectations than other subjects especially in the comprehension sample.

Figure K: Beliefs and norms

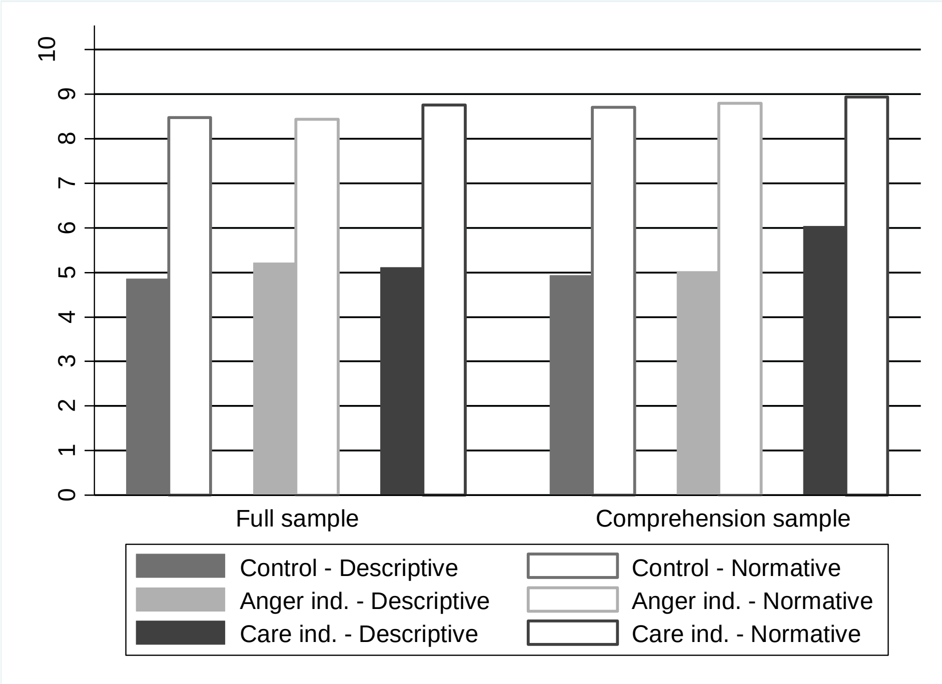
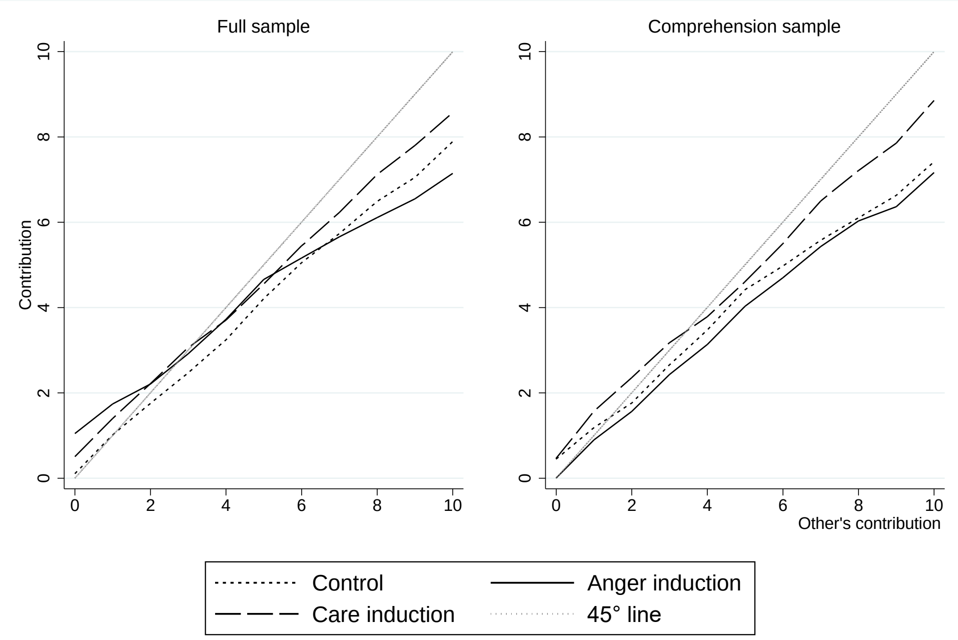


Figure K displays elicited beliefs about the actual contribution of subjects' partners (descriptive expectations, solid bars) as well as subject's perceptions of the normative contribution levels.

Result 0.7: Subjects' conditional contribution schedules – a proxy for their preferences - differ significantly across treatments with *Care* schedules being higher than those from *Control* and *Anger*.

Figure H presents subjects' contribution schedules as averages over all subjects within each treatment for all treatments. In the full sample on the left, contributions are slightly higher in the *Care* than in the *Control* treatment at all hypothesized partner contributions, while both lines show a similar slope. Subjects under *Anger* give slightly more than those in the other treatments when the other group member contributes little. At the same time, they give less at higher contributions of their group member (i.e. the slope is shallower). In the comprehension sample, *Care* subjects give the most for all possible contributions of a partner, followed by *Control* subjects and then by *Anger* subjects. There are significant pairwise differences (at the 10% level) between *Care* and *Control* subjects at hypothetical partner contributions of €0, €6, €7, €8, €9 and €10 in the comprehension sample according to ranksum tests. Also, there is a significant difference between *Care* and *Anger* subjects in the full sample for a hypothetical partner contribution of €10 (at p=.05) and between the *Anger* and *Control* subjects in the full sample for a hypothetical partner contribution of €0 (p=.09).

Figure H: Conditional contribution schedules



Especially when we focus at the results for the full sample we only find partial evidence for hypotheses 0.4 and 0.7. This suggest that a deeper look into the narrative texts that subjects wrote is warranted. Which we will turn to for the remainder of this study.

4 Narrative analysis

4.1 Description of the Text Corpus

The narrative texts corpus consists of two narratives that each subject from the ms-p (133) and the pgg-p (184) wrote down.[15] The size of this corpus can be inferred from Table N.

Table N: Narrative texts corpus

	Total words	Average words per subject per narrative (2 per subject; standard deviation in parentheses)
Motivational states part (N = 133)	47274	177.7 (37.4)
Public goods game part (N = 184)	66135	179.7 (36.1)

4.2 Text Pre-processing

[15] In contrast to chapter 2 of this document whose results are summarized under section 3.2.2 above, the results from the dictionary method, LDA and human rating analyses will be shown for the full sample of the pgg-p only.

Before applying the dictionary and LDA methods to the corpus of narrative texts, the texts are pre-processed. This step increases the retrieval of meaningful information from both methods and ensures comparability between methods. In particular, all tokens in the corpus are converted to lower case. Subsequently, high-frequency words that are occurring regularly in any type of text, so called stop-words, are removed from the text[16] which is a standard procedure in natural language processing. Examples are "the", "to", "also", which bear little meaningful content and whose presence are not helpful in distinguishing different types of narratives based on their content. After this step, the corpus is tokenized[17] which means that any token that is not an alphabetical letter is removed from the texts. Finally, the words are reduced to their linguistic root, a process known as "stemming".[18] Exemplarily, the words "running" and "runs" both become "run" through stemming. Table X presents how these pre-processing steps have reduced the dimensionality in the data.

Table X: Dimensionality reduction of data through pre-processing

Size of corpus before pre-processing (both parts)	Size of corpus after pre-processing (stop-word removal, tokenization)	Unique tokens after pre-processing (stemming)
113409 words	51880 tokens	22666 stems

4.3 Latent Dirichlet Allocation (LDA)

LDA (Blei et al., 2003) is a probabilistic, unsupervised statistical learning algorithm. Unsupervised means that neither is the information which narrative was written under which treatment fed to the algorithm, nor are other dependent variables from the experiment, pre-defined lists of words that supposedly belong to a topic, or some kind of training set used to give the algorithm prior information. LDA works on a quantitative representation of text data and uses the insight that the meaning of a document can be assessed by the words that the document uses. It is supposed to produce K meaningful topics from a collection of documents (corpus) and to represent each document in the corpus in terms of these topics. In the context of this study, a document is one narrative text and the corpus is the collection of all 634

[16] The German stopwords corpus from Python's Natural Language Toolkit (NLTK) is used for this step.
[17] NLTK's tokenize package was used for this.
[18] The SnowballStemmer from Python's NLTK package was used for this.

narratives ((133 + 184) $\times$ 2). A topic will be a set of words directed towards an identical underlying matter which needs to be interpreted by a human after it was found. As can be seen below, the estimated topics can be interpreted in the light of the topics that subjects were supposed to write about according to the experimental instructions. Due to our experimental setting, this reduces the degrees of freedom researchers have in topic interpretation which makes our results more scientifically robust. LDA topics can be conceptualized as latent variables that give values how intensely a subject wrote about the LDA-estimated topics in her narratives. Words will be part of one topic with a certain probability. Single words can be parts of many topics with different probabilities. Therefore, LDA is a tool that can reduce the dimensionality of sparsely distributed variables such as count data of words that appear in a document considerably. The main motivation why LDA is used in the context of this experiment is that it results in a quantitative representation of the share of single topics within a narrative. These additional quantitative variables were originally absent, since we could only observe in which treatment a subject participated in the experiment. Therefore, the topic shares allow to infer how much a subject complied with the motive treatment induction in an approximative way which gives an additional perspective on treatment effects of induced motives that manifest in the narratives. At the same time, this quantified topic content is based on topics that are identical between narratives from both parts of the experiment: ms-p and pgg-p which allows to investigate identical LDA topics stemming from identical instructions in the context of different outcome variables.

4.3.1 LDA Statistical Model and Statistical Learning Algorithm

The following introduces the LDA algorithm in a detailed, yet non-technical way.[19] LDA makes the following assumptions about the generative process behind each word / document. For each document, a vector with the dimensionality of the amount of topics that should be found is drawn from the Dirichlet distribution. This means that for each document m, a vector θ_m of dimension K (the amount of topics to be found) is drawn. For each document, θ_m gives the shares of each topic that a document comprises of. Initially, topics were drawn from the corpus, where a topic is itself a probability vector over all the unique words in the corpus (denoted β_{k_i}). Following this step, LDA generates each word in every document by assigning every word a topic according to θ_m. Once a topic was assigned to a word, this word is generated by this topic probability vector β_{k_i}.

[19] Please refer to the Appendix for a more technical description.

Concrete example:

In order to fix ideas let's assume that we want the algorithm to find two latent topics (k_1 and k_2) for a document m_1. Initially, let's assume that topic k_1 contains word x with 50% and word y with 50%. Likewise, assume that topic k_2 contains x with 90% and y with 10%.

- θ_1 is assigned to m_1. Initially, θ_1 assigns m_1 to be 40% about topic 1 and 60% about topic 2; meaning that every word in the document is assigned with 40% probability topic 1 and with 60% probability topic 2.
- Word 1 in m_1 gets assigned k_1 according to θ_1.
- Word 1 is generated by k_1 to be y.

LDA assumes this generative model for the production of the corpus. Based on this, the algorithm finds topics that are likely to have produced the realization of the documents in the corpus. For this purpose, a statistical learning method is used that is called collapsed Gibbs sampling. This learning method first randomly assigns a topic to every word in the corpus. Subsequently, it goes over each document and computes for every word in each document:

- The proportion of words in a document that are currently assigned to a certain topic, say k_1: $p(topic = k_1|document = m_1)$.
- The proportion how often a certain word is assigned to k_1 in the whole corpus: $p(word = w|topic = k_1)$.
- Word w is assigned a new topic where the probability that w is assigned to k_2 is computed by $p(topic = k_1|document = m_1) * p(word = w|topic = k_1)$.

The last step is repeated very often for every word in every document. After a cut-off value of repetitions that is set by the user, the word – topic assignments should have converged to a stationary state. Based on this outcome, the topic mixture of every document and the words associated to each topic is inferred. This informs about the share each topic occupies within a narrative. Moreover, one can infer the assignment of each word to a topic and its relative weight in this topic.

4.3.2 LDA: Parameter Selection

As hyperparameter values, α was chosen to be $\frac{5}{K}$, K was set to 6 and β was 0.01. Given the experimental design and the resulting corpus of narratives, the parameter calibration for K, the number of topics to be estimated, is straightforward. Across the treatments, subjects were specifically asked to write about six different narrative themes (*Control*: course of a typical day / what they did yesterday; *Anger*: situations that frustrated them / situations in which they were insulted; *Care*: situations in which they helped another person / situations in which they felt compassion for someone). The LDA algorithm was therefore instructed to estimate six topics in the corpus. This facilitates the subjective interpretation of the topics that LDA produced against the context of the experimental treatments. Furthermore, the latent topics that were found have a real representation and should actually be present in the text. α gives a prior tuning possibility about the amount of topics you expect to be present in the documents of your corpus. If you expect that each document contains most of the topics, α should be set relatively higher. Similarly, a high β value expresses that most of the words from the whole corpus are likely to be a part of every topic. Expecting that only a limited list of words is relevant for the different topics informs you to set β to a relatively lower value. The chosen hyperparameter calibration is the default setting within the applied LDA model that was used to estimate topics[20]. These algorithm parameters can be seen to be rather large which allows for convergence of the word topic assignment with a high likelihood and is feasible given the present corpus as it is rather small. A burn-in period of the Gibbs sampling of 5000 iterations was carried out. Following this, a sample of the topic assignment for each of the words in the corpus is drawn every 50 iterations until a total of 5000 iterations have passed which results in a sample of 100 draws. Putting the hyperparameter calibration into the perspective of the underlying narrative text corpus, one can say that α has a relatively large value and β a moderately small value (compare for instance Griffiths and Steyvers, 2004), despite being the Mallet-LDA default settings. However, this serves the underlying corpus and a priori expectations well: Roughly one third of the narratives corpus consists of *Control* narratives. It is likely that all narratives to some degree will contain elements that could be assigned to the description of yesterday's or typical daily activities. On the other side, the relatively low β value attempts to increase the specificity of the estimated topics which supports the insight that certain words are associated with specific motives (Chierchia et al., 2017, mimeo).

4.3.3 LDA: Empirical Design and Hypotheses

LDA and the statistical learning algorithm with collapsed Gibbs sampling was applied to all narratives from the ms-p and pgg-p using the calibration discussed above. For every narrative from both samples, this resulted in six figures, how much of topic k_i is present in narrative m_j.

[20] Mallet, the LDA implementation that was used comes from http://mallet.cs.umass.edu/index.php.
It is a statistical package that can estimate topics by means of LDA implemented in Java.

What constitutes a topic is however identical between narratives from both parts. Moreover, the six topics are each represented by a weighted list of words, where the respective weight that is mapped onto a word within a topic can be interpreted as the importance of this word within this topic. The topic share distributions of the two narratives that every subject wrote were averaged at the subject level.[21] For every subject, this gives six new variables that indicate how much he wrote in her narratives about the six LDA topics. This enables to use topic shares per subject within regression models which investigate outcome variables from both experimental parts. In particular, this allows to assess whether a topic has predictive power in explaining motivational state changes from the ms-p, but at the same time, whether this identical topic is relevant for explaining behavior, beliefs or perceptions in the pgg-p. The idea is that LDA finds topics that can be associated with either the *Care*, *Anger* or *Control* treatment. Therefore, a topic share at the individual level represents how much a subject complied with the typical narrative content[22] of a given treatment. If the narrative of a subject is relatively congruent with a single LDA topic then this means that her narrative contains the words of this one LDA topic to a large extent.

Hypothesis 1 – Topics that can be related to representing the *Care* treatment predict motivational state ratings of compassion motivation.

Hypothesis 1 expects that there are LDA topics that can credibly be associated with the *Care* treatment and that the closer a subject's narrative resembles this LDA topic, the higher compassion motivation he subsequently reports.

Hypothesis 2 – Topics that can be related to representing the *Anger* treatment predict motivational state ratings of anger motivation.

Analogous to the previous discussion, we hypothesize that the closer a subject's narrative represents LDA topics that are clearly attributable to the *Anger* treatment, the higher anger motivation he subsequently reveals.

In addition, we hypothesize that LDA topics are significantly associated with outcomes from the pgg-p. Given the mediocre full sample results of the treatment dummies on predicting behavior in the public goods game depicted in Table F, this sheds light on whether a quantifiable (vs. dichotomous) treatment dimension by means of topic shares predicts behavior,

[21] Like in the dictionary method this is feasible since the treatment specific narrative themes were designed such that they activate the identical motivation system within a treatment.
[22] As measured by LDA.

beliefs, attention and perception in the public goods game. The reason for this is that only accounting for the treatment assignment ignores potentially important dimensions how narratives and which components of them influence behavior.

Table F

Dependent variable:	Public good contribution	Belief about other's contribution	Perception of the game (1 = competitive)	Attention to own payoff	Attention to other's payoff	Probability to be conditional contributor
Model	F1	F2	F3	F4	F5	F6
Care treatment dummy	-0.18 (0.56)	0.25 (0.51)	0.69 (0.48)	-0.71 (0.53)	-0.43 (0.45)	0.37 (0.46)
Anger treatment dummy	-0.37 (0.56)	0.35 (0.52)	1.07** (0.47)	-0.15 (0.53)	0.45 (0.45)	-0.55 (0.41)
Constant	6.00*** (0.41)	4.86*** (0.37)	-1.81*** (0.38)	5.88*** (0.38)	4.23*** (0.33)	1.23* (0.32)
N	184	184	184	184	184	184
Regression	OLS	OLS	logit	OLS	OLS	Logit
R^2/ *AIC*	<0.01	<0.01	AIC: 202.76	0.01	0.02	AIC: 205.7

*** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

By means of the weighted list of words that comprises a topic, this also helps to assess which concepts in particular are associated with changes in outcome variables. Based on insights from motivation psychology (Goetz et al., 2010; Lerner and Tiedens, 2006) we hypothesize the following.

Hypothesis 3 – In accordance with predictions that are informed by motivation psychology about compassion and anger motives: Topics that can be related to the *Care* (*Anger*) treatment lead to more prosocial (antisocial) behavior, beliefs, perceptions and attentional patters in the public goods game.

4.3.4 LDA: Results and Discussion

Table X in the appendix gives an account of the top 25 words for each of the six LDA topics. Based on this list of words, the topics can be interpreted. To illustrate the LDA topics and their key words, we provide a visualization with word clouds. Here, the size of a word is approximately proportional to its probability in the topic. In order to increase accessibility of the word clouds, only the 20 - 25 words with the highest probability within each topic are shown. Figure X presents the word clouds.

- *Topic 1* prominently contain words about university examination situations, frustration, as well as work and money related terms. This topic fits well the narrative theme that subjects were instructed to write about situations in which they were frustrated. Apparently, it was common to associate frustration with examination situations at university which makes sense given the nature of our subject pool.
- *Topic 2* puts emphasis on words like insult and bad-mouth, angry, class, teacher and football. We interpret this topic to be about narratives that recall situations in which subjects were insulted and harassed. It appears that these situations occurred in school or while playing football and were emotionally loaded as the world "feels" indicates.
- *Topics 3 and 5* describe the activities that subjects did yesterday / their everyday routine, respectively. For differentiating these two topics, note that topic 3 ("yesterday") circles around activities with friends but is otherwise unspecific and sparse which mirrors that there is considerable heterogeneity in yesterday's activities. Topic 5 on the other side is clearly focused on university and additionally contains the words "morning", "noon" and "afternoon" as well as activities that are probably associated with these activities like "breakfast", "lecture" and "sports", terms that are absent in such prominence in topic 3.
- *Topic 4* can be interpreted to contain narratives that describe how the subject noticed someone that needed help and subsequently helped this person. It appears that these situations frequently occurred in the public like on the street or in public transport, involved elderly citizens and was followed by gratitude and feelings of joy.
- *Topic 6* centers around terms such as "friends", "child", "parents", "father" and "mother". It further contains words of closeness, tenderness and deep emotions such as "feeling", "conversation", "notice", "eye", "love", "fear", "sadness" and "happiness" [23]. We therefore interpret this topic to describe situations in which subjects narrated experiences of compassionate and empathic states directed towards family and friends.

In what follows, we associate topics 1 and 2 with the *Anger* treatment, topics 3 and 5 with the *Control* treatment and topics 4 and 6 with the *Care* treatment. Furthermore, we omit topics 3

[23] The precise narrative theme that we associate with this LDA topic asked the subjects: "Please write down a personal memory of a situation in which you felt feelings of compassion and warmth for another person, as well as felt the motivation to improve this person's wellbeing."

and 5 in the following regression models for the reason that they represent the hold out category of the *Control* treatment.

Result 1: Topics that can be related to representing the Care treatment predict motivational state ratings of compassion motivation which gives support for hypothesis 1.

Model H1 in Table H confirms the significantly positive influence that the *Care* treatment has on subsequent ratings of compassion motivation. Model H2 shows that topics 4 and 6 have a significantly positive relationship with reported compassion motivation, albeit the influence of topic 4 is stronger. Note that this result holds for the whole sample of narratives from the ms-p and not only for narratives stemming from *Care*, which confirms expectations derived from motivation psychology. In Model H3[24], neither the coefficient estimates on the *Care* treatment dummy nor those for topics 4 and 6 indicate a significant relationship, while all three show a positive relationship with the compassionate motivational state.[25]

Table H

[24] The models in Table H may also lead to speculate whether there exists collinearity between the LDA topics and the treatment dummies. Indeed, while there exists a variance inflation factor (VIF) between the two treatment dummies of 1.32 and the highest VIF for the regressors in model G2 is 1.09 for topic 4, these values multiply in model G3. Specifically, the VIFs for the two treatment dummies in this model are above 8 and the VIFs on the topic variables double to triple compared to the situation in model G2. We interpret this as suggestive evidence of multicollinearity between the treatment dummies and the LDA topics.

[25] A robustness check in the appendix that regresses the LDA topics from the models in Table H, the word categories from Table D as well as the *Care* and *Anger* treatment dummies confirms the significance that topics 4 and 6 have on subsequent compassion ratings. Moreover, when LDA topics, treatment dummies and word categories are together in one model to predict compassion motivation, no variable shows a significant relationship with compassion motivation except topics 4 and 6.

Dependent variable: Self-reported compassion motivation rating in "motivational states part"			
Model	H1	H2	H3
Care treatment dummy	1.22***	-	0.04
	(0.44)	(-)	(1.09)
Anger treatment dummy	-0.50	-	-0.83
	(0.45)	(-)	(1.12)
Topic 1: Frustration at uni / job	-	-1.73	-0.46
	(-)	(1.49)	(2.22)
Topic 2: Anger at school	-	-1.59	0.27
	(-)	(1.36)	(2.53)
Topic 4: Help others in public	-	3.51**	3.55
	(-)	(1.52)	(2.30)
Topic 6: Empathy / compassion towards close ones	-	2.72*	2.24
	(-)	(1.37)	(2.46)
Constant	4.65***	4.37***	1.67***
	(0.31)	(0.47)	(0.52)
N	133	133	133
R^2	0.11	0.13	0.13

OLS; *** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

Considering however the size of the coefficient estimates in model H3, it appears that especially topic 4 is responsible for the treatment effect of the *Care* treatment on self-reported compassion motivation in comparison to the treatment dummy itself.

Result 2: Topics that are associated with the Anger treatment influence subsequent motivational states of anger significantly. Hypothesis 2 is confirmed.

(a) Topic 1: "Frustration at school"

(b) Topic 2: "Insult and anger"

(c) Topic 3: "Yesterday"

(d) Topic 4: "Help (elderly in public)"

(e) Topic 5: "Everyday routine"

(f) Topic 6: „Compassion (towards friends and family)"

Figure X: Estimated LDA topics from corpus of pilot and experimental essays illustrated as word clouds

The results from model I2 in Table I indicate that the more subjects across treatments wrote about topic 1 or topic 2, the more anger motivation they reported.[26]

Table I

Dependent variable: Self-reported anger motivation rating in "motivational states part"			
Model	I1	I2	I3
Care treatment dummy	0.10 (0.38)	- (-)	0.85 (0.94)
Anger treatment dummy	1.43*** (0.38)	- (-)	1.44 (0.96)
Topic 1: Frustration at uni / job	- (-)	2.94** (1.29)	0.81 (1.91)
Topic 2: Anger at school	- (-)	3.08*** (1.17)	0.27 (2.18)
Topic 4: Help others in public	- (-)	0.20 (1.31)	-1.23 (1.98)
Topic 6: Empathy / compassion towards close ones	- (-)	-1.15 (1.19)	-2.32 (2.12)
Constant	1.50*** (0.27)	1.26*** (0.41)	1.67*** (0.52)
N	133	133	133
R^2	0.12	0.12	0.13

OLS; *** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

A similar picture to the one under model H3 emerges in model I3. Integrating the LDA topics and the treatment dummies into the same model in order to predict subjects' ratings of anger motivation leads to the findings that none of these variables shows a significant relationship with anger motivation. However, taking a closer look at the coefficient estimates under model I3 we see that compared to model I1 the size of the coefficient estimate on the *Anger* treatment dummy stays almost equal between the two models. On the other hand, comparing the coefficient estimates on the topic variables between models I2 and I3 shows that those for

[26] A robustness check in the appendix that regresses the LDA topics from the models in Table I, the word categories from Table E as well as the *Care* and *Anger* treatment dummies on anger motivation rating confirms that the *Anger* treatment dummy has a significant impact on anger ratings while the anger LDA topics do not. Topic 6 has a significant negative relationship with reported anger motivation, while the result for the anger word category to have a significant negative influence on anger ratings persists. Together with the previous footnote this suggests that compassion motivation ratings can be explained well by the respective LDA topics while this does not hold for anger motivation when control variables are present.

topics 1 and 2 shrink considerably. We take this as suggestive evidence that topics 1 and 2 represent certain dimensions of what constitutes the *Anger* treatment effect, but not all as the influence of the *Anger* treatment dummy stays relatively large even when these two topics are in the model. Note that the opposite insights emerge from Table H, where it seems like topics 4 and 6 approximate the treatment effect of *Care* on self-reported compassion motivation well.

Result 3: Topic 4 which treats helping others, has a significantly positive influence on behavior and beliefs in the public goods game. Both topics that are associated with Anger *have an influence on outcomes (preferences, perceptions and attentional patterns) from the public goods game whose direction is as hypothesized. We find support for hypothesis 3.*

The presentation of the following results should be considered under the light of the insights from Table F above. Except for perception of the public goods game to be either a competitive or a cooperative game, where *Anger* leads subjects to perceive the game to be significantly more competitive, we find no significant influence of the treatments on the full sample of the pgg-p by means of regression analysis.[27] The LDA topics allow to expand this analysis into two further dimensions. First, the topic distributions over the narratives that a subject wrote, or topic shares, represent a quantitative variable as opposed to the treatment dummies that are binary. Second, LDA synthesized specific topics. According to results 1 and 2, *Care* gets split up into helping especially the elderly in public and situations of empathy and compassionate feelings towards friends and family. *Anger* is divided up into frustration with exam situations and insult / conflict in the school or at sports. This allows to relate behavior, beliefs and perceptions in the public goods game to more specific topics, i.e. what subjects actually wrote about within treatments as synthesized by LDA. This enables to check the predictions stemming from motivation psychology against the actually realized narrative contents as synthesized by LDA.

Model J1 in Table J shows that there exists a significant relationship between the *Care* topic of helping others and contributions to the public good. Interestingly, the other topic that is related to the *Care* treatment, empathy towards close ones, has a negative influence on contributions. Evidence from motivation psychology may rationalize this finding. While empathy is a necessary first step such that compassionate goals and action tendencies materialize, it crucially needs to be distinguished from compassion. Empathy means to feel with a person and share the same feelings of this person. Topic 6 prominently contains the

[27] This picture changes however, when we consider the comprehension sample.

words fear and sad among its most defining terms. Therefore, to the extent that subjects recalled narratives in which other persons were sad, this might have made them sad again too.

Table J

Dependent variable:	Public good contribution	Belief about other's contribution	Perception of the game (1 = competitive)	Attention to own payoff	Attention to other's payoff	Probability to be conditional contributor
Model	J1	J2	J3	J4	J5	J6
Topic 1: Frustration at uni / job	0.35 (2.12)	0.22 (1.94)	1.13 (1.57)	4.46** (1.98)	1.01 (1.72)	0.75 (1.61)
Topic 2: Anger at school	-0.22 (1.87)	1.94 (1.71)	2.85** (1.36)	-2.86 (1.75)	0.37 (1.52)	-2.66** (1.34)
Topic 4: Help others in public	3.02* (1.74)	2.80* (1.60)	-0.29 (1.44)	0.56 (1.64)	0.32 (1.42)	0.97 (1.41)
Topic 6: Empathy / compassion towards close ones	-1.76 (1.56)	-1.27 (1.43)	1.46 (1.22)	-2.35 (1.47)	-1.93 (1.27)	0.17 (1.23)
Constant	5.64*** (0.62)	4.52*** (0.56)	-2.05*** (0.55)	5.73*** (0.58)	4.35*** (0.50)	1.22* (0.62)
N	184	184	184	184	184	184
Regression	OLS	OLS	logit	OLS	OLS	Logit
R^2/ AIC	0.02	0.03	AIC: 205.37	0.05	0.02	AIC: 210.01

OLS; *** = p < 0.01; ** = p < 0.05; * = p < 0.1

Compassion on the other hand is conceptualized as feeling for another person. You are interested in the wellbeing of the other and actively want to promote it which motivates to prosocial action tendencies towards the other person (Klimecki and Singer, 2012). Therefore, narratives about helping might have reactivated the experience of being compassion motivated and of becoming active and doing something for others, while topic 6 might have led to empathic distress or sadness which previously was associated with self-protective and avoidant choices (Batson et al., 1983). A very similar picture emerges for beliefs about the contribution of the other group member. Subjects that have a higher share of the helping topic in their narratives, have higher beliefs. No other topic is significantly associated with beliefs, while topic 6 again has a negative influence on beliefs. Topic 2, anger at school / sports, displays a significant positive relationship with a competitive perception of the game form of the public goods game. This finding is supported by motivation psychological evidence. Smith and Lazarus (1990) find that anger leads subjects to perceive a situation as unjust. It was also argued that anger leads to a more selective perception towards a more hostile one (Finucane, 2011).

High narrative shares of frustrating experiences significantly increase the subsequent attention to the own payoff. Van Kleef et al. (2008) summarize in a literature review that the anger motive leads subjects to pursue more competitive behavior, which potentially increasing their focus on their own payoff relative to the other's. Conditional contributors increase their contribution if they believe that the other group member does so as well (Fischbacher et al., 2001). As Van Lange et al. (2017) argue, this conditionally cooperative behavior relies on the belief of the strategy of the other group member and is therefore a genuine trust situation. Taking this into account, it was found by Dunn and Schweitzer (2005) that anger significantly decreases trust. This argumentation is supported by the evidence depicted in model J6. Topic 2 that is associated with *Anger* leads subjects significantly less often to adopt the strategy of conditional cooperation.

4.4 Dictionary Method

In the dictionary method, sometimes also lexical or "bag of words" method, pre-defined dictionaries are used that map a word to a certain concept or group of words like "negative emotion" or "expressions related to cars". Applied to the present research this approach implies that every word within a narrative text is assigned to a certain category as indicated by the used dictionary. Subsequently, the narrative text can be represented as a frequency distribution over the word categories of the dictionary. The dictionary we use to check the marker word hypothesis is the German version of LIWC (Pennebaker et al., 2001; Wolf et al., 2008) which we augmented by seven motivation word categories. These seven motivation word categories are: achievement, affiliation, anger, care, consumption seeking, fear and power.[28] The LIWC dictionary covers around 80% of the words that are used in everyday language, written or spoken and organizes these words within categories. Table O in the appendix presents the category counts from the narrative texts.

We augmented the LIWC categories by motivation specific categories which were obtained as follows. Chierchia et al. (2018) conducted a semantic categorization task to investigate how subjects typically conceptualize the seven motives of achievement, affiliation, anger, care, consumption seeking, fear and power semantically. For every motive, the authors identified a list of words (21 on average) that significantly differentiated the motive from the six other

[28] All LIWC category and motivation category words were also stemmed before their counts in the narratives were assessed since tokens in the narratives were also stemmed.

motives.[29] Subsequently, for each motive category, this list of words was augmented by finding German synonyms for every word. To this end, we wrote a script that returned for every specific motivation word the synonyms that can be found in the Deutscher Wortschatz project of Leipzig University[30] which contains text corpora in German. The synonyms were added to the initial list of motivation words for each of the seven motivations respectively.[31] Following this, the original LIWC categories were appended by these seven new synthesized motivation word categories.[32] For every subject, category counts over her two narrative texts were added. This is feasible since narrative themes were treatment specific and designed such that a common behavioral response could be expected after having recalled the two narratives per treatment. Implicitly, this also reduces the width of the data point per subject considerably and can therefore be interpreted as a measure to reduce dimensionality while preserving the original information.[33] Some category counts deviate substantially from a normal distribution, violating an assumption of OLS-regression analysis. Every category that did not appear in more than 50% of the subjects' narratives under a certain treatment was dichotomized. More info on this can be found under Figure T in appendix C.

4.4.1 Marker Word Hypothesis

The underlying hypothesis that is tested in our corpus of narrative texts by the LIWC-based dictionary method is the marker word hypothesis. It postulates that motives within a text can be assessed by identifying and counting so called "marker" words (Schultheiss, 2013). Put differently, does the appearance of certain motive specific words indicate the presence or absence of a motivational need? Traditionally, presence or absence of implicit motivation[34] in

---

[29] We wanted the motive category words to be maximally discriminative and typical for a motive, so we selected a significance level of below 5%. Put differently, the motive words we selected were exclusively associated with the target motive in Chierchia et al's study at the 95% confidence interval.

[30] http://wortschatz.uni-leipzig.de/de

[31] The complete lists of words can be obtained from the author upon request. On average, adding all synonyms for each word within a motivation category increased the size of these word categories by factor 25, which made them comparable in size to other LIWC categories.

[32] The original LIWC already contains categories that are relevant for identifying motivation words that we study. LIWC contains the sub-categories Anxiety, Anger and Achieve. Words from within these categories were merged with the three respective motivation word categories fear, anger and achievement.

[33] Also, the goal of the study is to analyze narratives that can be attributed to the same motive as we have specific hypotheses for this case.

[34] It should be noted that our design allows to test the marker word hypothesis rather in its explicit dimension as subjects are asked to recall specific memories from their own lives and

texts is evaluated by content coding of picture story exercises (PSE) (McClelland et al., 1989). This consists of subjects writing down imaginative narratives about a series of pictures that show stylized social situations. These narrative texts are subsequently analyzed by research assistants who use coding systems developed for this purpose to analyze the motivational intensity within the narrative. This process is resource intensive but still prevalent in current research. Several studies have suggested that human assessment of motivational needs that underlie a narrative can at least be approximated by taking the use of marker words into account (Pennebaker and King, 1999; Smith, 1968; Seidenstücker and Seidenstücker, 1974; Hogenraad, 2003, 2005).

4.4.2 Dictionary Method: Empirical Design and Hypotheses

The empirical strategy for the dictionary method seeks to find a regression model, i.e. a linear combination of relevant categories from the dictionary we consider, that consistently explains two different groups of dependent variables across treatments.[35] For one, we are looking for a combination of dictionary categories that predicts across all treatments self-reported ratings of compassion / anger motivation, the dependent variable from the ms-p. In a second step, we use this model that we validated on self- reported motivational ratings to also explain findings about behavior, beliefs and perception in the pgg-p. In terms of model selection, we rely on previous insights obtained from Schultheiss (2013) and Chierchia et al. (2018). The latter find that at the semantic level, compassion (care) and affiliation are closely related.[36] Based on this finding, we use the analysis of Schultheiss (2013) to select the relevant dictionary categories to predict self-reported ratings of compassion motivation (the compassion model) in our "motivational states part". Specifically, we use the LIWC dictionary categories that Schultheiss found to predict the affiliation motivation as variables to be informative about subjects' motivational states of compassion. This list of predictors was further augmented by the

subsequently rate their motivational state on a scale. Albeit, the data from the pgg-p allow also insights into more implicit motivational expressions.

[35] Initially, we investigate whether specific categories of motivation words are sensitive to experimental arousal. To this end, we investigated differences in the use of specific groups of motivational words between narrative treatments. For the ms-p and the pgg-p, rank sum tests[35] show that *Care* leads to significantly more care words than *Control* (at 1% significance level). A similar, yet more nuanced result emerges for the use of anger motivation words between treatments. Refer to appendix section C.2 for details.

[36] Both revolve around the common theme of communion (see also: Abele et al., 2008). This closeness of compassion and affiliation is further supported by means of principal component analysis that demonstrated an overlap in the factor loadings of care and affiliation words as well as an almost equal positioning in terms of valence and arousal (Chierchia et al. 2018).

motivation word categories of care and affiliation which we hypothesize have an impact on compassionate motivational states if the marker word hypothesis is true.[37]

Hypothesis 4 –Predominantly, the use of care words predicts self-reported motivational ratings of compassion motivation. Word categories that are related to compassion motivation are also relevant for predicting compassionate motivational states.

A similar approach was used for the variable selection of the model that predicts self-reported anger ratings (the anger model) from the ms-p. First, it was found by Chierchia et al. (2018) that power and anger motivation are semantically close to one another.[38] Therefore, we use the insights from Schultheiss (2013) about word categories that predict power motivation for the selection of our anger model and augment this by the anger motivation word category from our augmented dictionary. In addition, the compassion and anger models will also contain *Care* and *Anger* treatment dummies as a robustness check.

Hypothesis 5 - Predominantly, the use of anger words predicts self-reported motivational ratings of anger motivation. Word categories related to anger motivation are also relevant for predicting angry motivational states.

Subsequently, the explanatory variables from the compassion model and those from the anger model are combined to predict behavior, beliefs and perceptions in the public goods game. This procedure allows insights whether models that were tested on self-reported motivational state changes may also have the ability to predict response variables from the pgg-p. This outlined empirical design probes the marker word hypothesis with a focus on generalizability: a) Does the presence of specific marker words regardless of the narrative treatment predict motivational state changes? b) If so, do these predicting word categories also predict responses in a controlled laboratory public goods game experiment?

Hypothesis 6 – The marker word regression model will significantly predict behavioral response variables from the public goods game. Dictionary categories related to care predict prosocial behavior in the game while those related to anger predict antisocial behavior.

4.4.3 Dictionary Method: Results and Discussion

[37] In particular, we used the variables from Table 4 of Schultheiss (2013) and augmented them by the motivation word categories of care and affiliation.

[38] This manifests in a very similar valence of the two motivations as well as similar factor loadings as obtained from the principal component analysis that Chierchia et al. conducted on words that are specific to these two motivations. The two motivations also trigger similar behavioral responses.

Result 4: Care marker words are not linked up to motivational state ratings of compassion. Also, word categories related to care words have no systematic influence on compassion motivation. We find no evidence for hypothesis 4.

Table D contains three regression models that regress the self-reported compassion motivation after subjects wrote narratives across all three treatments on a list of variables. These variables follow from the empirical strategy outlined above and depending on the model comprise of the treatment dummies, care motivation and affiliation motivation word categories as well as categories related to these motivations as indicated by Schultheiss (2013). Model D1

Table D

Dependent variable: Self-reported compassion motivation rating in "motivational states part"			
Model	D1	D2	D3
Care treatment dummy	1.22*** (0.44)	- (-)	1.07 (0.70)
Anger treatment dummy	-0.50 (0.45)	- (-)	-0.60 (0.60)
"Other reference"	- (-)	0.14 (0.11)	0.07 (0.11)
"Positive emotion"	- (-)	0.05 (0.03)	0.05 (0.03)
"Care"	- (-)	0.07 (0.09)	0.03 (0.09)
"Affiliation"	- (-)	0.47 (0.40)	0.65 (0.40)
"Tentative"	- (-)	-0.13 (0.10)	-0.12 (0.10)
"Communication"	- (-)	-0.09* (0.05)	-0.05 (0.06)
"Friends"	- (-)	-0.07 (0.08)	-0.11 (0.08)
"Family"	- (-)	0.06 (0.09)	-0.04 (0.09)
"Humans"	- (-)	0.63 (0.45)	0.36 (0.53)
Constant	4.65*** (0.31)	4.25*** (0.55)	4.49*** (0.56)
N	133	133	133
R^2	0.11	0.11	0.17

OLS; *** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

demonstrates that treatment assignment to *Care* in the ms-p has a significantly positive influence on subsequent states of compassion motivation. Leaving the treatment dummies aside, model D2 indicates that while the influence of the use of care and affiliation motivation words in narratives has a positive influence on subsequent compassion motivation, this influence is not significant. Interestingly, the more subjects wrote about communication that lower compassionate states they reported afterwards. Model D3 however reveals that when you integrate the variables from models D1 and D2 into one model, the *Care* treatment dummy coefficient estimate loses its significance. This can be interpreted that the variables from model D2 contain explanatory power that was attributed to the treatment dummy in model D1 which indicates that taken the selected dictionary word categories together, they possess some explanatory power for predicting compassionate motivational states.

Result 5: The anger word category has significant explanatory power in predicting anger motivation, which supports hypothesis 5.

A clearer picture becomes evident for the motivational state of anger as the dependent variable. Table E depicts three regression models that regress either the treatment dummies, the word categories relevant for anger motivation or the combination between the two onto angry motivational states.

Table E

Dependent variable: Self-reported anger motivation rating in “motivational states part”			
Model	E1	E2	E3
Care treatment dummy	0.10 (0.38)	- (-)	0.39 (0.42)
Anger treatment dummy	1.43*** (0.38)	- (-)	1.38*** (0.41)
“Power”	- (-)	0.09 (0.06)	0.04 (0.06)
“Anger”	- (-)	0.14** (0.06)	0.12* (0.07)
“Tentative”	- (-)	-0.11 (0.09)	-0.12 (0.09)
Constant	1.50*** (0.27)	1.20*** (0.38)	1.04** (0.41)
N	133	133	133
R^2	0.12	0.07	0.15

OLS; *** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

It can be inferred from model E1 that participation in *Anger* has a strong positive influence on subsequently reported anger motivation. The significance on the coefficient estimate on the anger motivation words category in model E2 presents evidence that as hypothesized, the more anger words a subject uses within a narrative the higher are the states of anger motivation that he subsequently reports. Model E3 demonstrates that this result is robust to re-integrating the treatment dummies back into the regression model. This suggests that using anger words is not merely a substitute for having participated in the *Anger* treatment, but that the use of anger words leads to increased anger motivation also when used in other narratives.

Result 6: Neither the anger nor the care word category predicts behavior, beliefs or perception in the public goods game. Word categories related to the care category like the other reference, affiliation motivation, friends, family and humans categories are intuitively but unsteadily associated with outcome variables we measured in the public goods game.

Table G investigates whether the combination from the variables under models D2 and E2 have an influence on dependent variables from the public goods game. We find that subjects that use words related to friends or tentative words like "maybe" give more to the public good. Especially the former result is in line with expectations that the motive to care for or affiliate with friends leads to higher contributions to the public good. The result on the friends category carries over to beliefs about contributions of the other group member in the public goods game. Writing more about friends in the narrative might also be related to expect others to also care for oneself. Furthermore, two other sub-categories from the Social root category, the family and humans word categories are associated with significantly less attention to the own payoff from public good decisions. This result is not only intuitive, it mirrors again perceptional and attentional patterns that can be expected when the compassion motivation drives behavior towards supporting others (Crocker and Canevello, 2012). The use of words from the affiliation motivation word category predicts attention to other's payoff. Again, this is an intuitive result and rationalizable by insights from motivation psychology. Affiliation motivation describes the need to be liked or the feeling to belong (Heckhausen and Heckhausen, 2010). In a public goods game, the own actions influence the payoff of the other group member. Subjects motivated by affiliation may hope to be liked more by the other group member if they produced a higher payoff for this group member. Finally, the pgg-p revealed significant differences in conditionally cooperative behavior between treatments in the game (see result 0.7 above, Fischbacher et al., 2001). Model G6 shows that there exists a significant relationship between the use of words that refer to others and the probability to be a conditional cooperator.

Intuitively, a focus on other's behavior is necessary to be a conditional cooperator who pursues a strategy to provide towards a public good if others do so too.

Table G

Dependent variable:	Public good contribution	Belief about other's contribution	Perception of the game (1 = competitive)	Attention to own payoff	Attention to other's payoff	Probability to be conditional contributor
Model	G1	G2	G3	G4	G5	G6
"Power"	0.09	0.08	-0.05	0.04	0.02	-0.01
	(0.09)	(0.08)	(0.07)	(0.09)	(0.07)	(0.07)
"Anger"	0.12	0.08	0.02	-0.08	-0.11	-0.07
	(0.10)	(0.09)	(0.08)	(0.10)	(0.08)	(0.08)
"Other reference"	0.06	0.06	0.16	0.10	0.11	0.30**
	(0.15)	(0.14)	(0.11)	(0.14)	(0.12)	(0.14)
"Positive emotion"	-0.05	-0.03	0.05	-0.03	-0.03	0.00
	(0.04)	(0.04)	(0.03)	(0.04)	(0.04)	(0.03)
"Care"	-0.09	-0.09	-0.07	-0.01	-0.01	-0.04
	(0.11)	(0.10)	(0.09)	(0.10)	(0.09)	(0.09)
"Affiliation"	0.32	0.09	0.13	0.50	0.78*	0.19
	(0.50)	(0.46)	(0.39)	(0.47)	(0.40)	(0.40)
"Tentative"	0.24*	0.05	0.07	0.08	-0.07	-0.07
	(0.14)	(0.13)	(0.11)	(0.14)	(0.12)	(0.11)
"Communication"	-0.11	-0.14**	0.02	0.10	0.06	-0.02
	(0.07)	(0.06)	(0.05)	(0.07)	(0.06)	(0.05)
"Friends"	0.18*	0.19*	0.02	-0.05	0.13	0.02
	(0.10)	(0.10)	(0.08)	(0.10)	(0.08)	(0.08)
"Family"	0.03	-0.04	-0.06	-0.16**	-0.10	-0.01
	(0.08)	(0.07)	(0.07)	(0.08)	(0.06)	(0.07)
"Humans"	-0.07	0.84	0.42	-0.88*	-0.19	0.20
	(0.56)	(0.51)	(0.44)	(0.52)	(0.45)	(0.43)
Constant	5.35***	4.76***	-2.06***	6.05***	4.26***	1.22*
	(0.81)	(0.75)	(0.64)	(0.77)	(0.65)	(0.62)
N	184	184	184	184	184	184
Regression	OLS	OLS	logit	OLS	OLS	Logit
R^2/ AIC	0.06	0.06	AIC: 218.53	0.08	0.09	AIC: 220.31

*** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

While models G1 – G6 yield some intuitive results, we also observe relations that are not intuitive like the results on tentative and communication in models G1 and G2. An evident reason lies in the dictionary method itself which does not account for context and ambiguous meanings of words. Furthermore, generalizability of the dictionary method likely presents another hurdle. Smith (1968) pointed out that when marker-word dictionaries are derived, they are usually validated for a limited and specific scope of application. They may function well with sources that they were produced from and potentially validated on, but lose validity when applied to new and different applications. This is relevant to our study as the LIWC and the motivation dictionaries from Chierchia et al. that we used were developed from different sources than our narrative experiment.

4.5 Human Coding of Motivation and Intensity

As outlined above, human coding represents the “gold standard” in quantifying narrative content in terms of accuracy and previous usage within economics. In our implementation of the human coding, two research assistants were provided with information about the two target motives compassion and anger. This information summarized the motives with respect to reactions, behavioral tendencies, thought patterns, appraisal tendencies, perception tendencies and accompanying emotions. Essentially, it was a condensed, yet accessible version of chapters 2.2.2.1 and 2.2.2.2 of this manuscript.[39] This information refrained from indicating to the RA’s which parts of narratives were typical for a specific motive thus preserving an unsupervised element of the human coding. After having studied the information, the two RA’s independently rated every narrative of the pgg-p[40] along three dimensions. These are: *How compassion-motivated were behavior, perceptions and sentiments of the narrating person in the situation it described?*; *How anger-motivated were behavior, perceptions and sentiments of the narrating person in the situation it described?*; *How high was the intensity in the narrated situation for you / How much did the narrated situation touch you / changed something in you?* RA’s had to indicate their rating on a scale from 0 (“Not at all”) to 11 (“Absolutely”). We made no effort to mediate agreement between coders since we wanted the ratings to be as independent as possible such that any potential errors are uncorrelated. The RA’s were completely unaware of the study’s hypotheses as they were only presented with the narratives and had no information about the structure or instruction materials of the experiment. Notably, the research assistants were unaware of which narrative was written under which treatment since they were given the narratives in random order and without any other additional information than the narrative text. Like in the previous two parts we use averages at the subject level. This means that RA’s coded narratives at the individual level, but that their ratings were

[39] Please find a translated version of this information in section E of the appendix.
[40] Due to resource constraints this human coding section could only consider one part of the experiment for the narrative analysis. We think that it is interesting to compare the two computerized approaches with a human assessment of motivational intensity and focused on the more interesting behavioral part, the pgg-p before having to leave out the human coding analysis completely.

averaged over the two narratives a subject wrote per treatment as well as over the two RA’s. For every subject this resulted in three new variables that ranged from 0 to 11 and which described how compassion-motivated, anger-motivated and intense the content of a subject’s narratives was. With these design choices we follow previous work by Brandts and Cooper (2007).

4.5.1 Human Coding of Motivation and Intensity: Empirical Design and Hypotheses

The human coding sought to meet a comparable goal like the dictionary method and LDA approaches. Narrative texts of each subject should be represented as a number of quantitative variables that are supposed to capture the degree of motivational intensity that the narratives contained. We hypothesize that human coders can assess the motivational content of a narrative due to their human interpretation skills of the meaning behind texts, as well as being able to empathize and putting themselves in the positions of the narrators. Therefore, we expect that human coders can assess how clear a motivation comes through in a narrative, thus adding a quantitative dimension how motivational content can be accounted for other than the binary treatment dummies. It is therefore insightful to account for a proxy of the research assistants’ abilities to discriminate between different motivations and assess their intensities. For this purpose, we look at the mean human motivation ratings of narratives between treatments and compare whether motivation scores that RA’s gave narratives from different treatments differ directionally:

Hypothesis 9 – Anger motivation ratings as assigned by the human coders are significantly higher in narrative texts from the *Anger* treatment than in narratives from *Control.*

Hypothesis 10 – Compassion motivation ratings as assigned by the human coders are significantly higher in narrative texts from the *Care* treatment than in narratives from *Control.*

If these two hypotheses are true, they jointly indicate with results 0.1 and 0.2 depicted in Figure A1 above that human coders can assess the degree of prevalence of a certain motivation within a subject that produced a narrative by assessing this narrative. It is then natural to link these assessments to behavior in the subsequently observed public goods game as follows:

Hypothesis 11 – Higher compassion motivation ratings as assessed by the human coders lead to more prosocial behavior, beliefs and perceptions in the public goods game, while higher anger motivation ratings analogously lead to more antisocial outcome variables. Higher motivational intensity leads to less social behavior, beliefs and perceptions.

We further hypothesize that higher intensity in a text leads to less social behavior, beliefs and perceptions. We conjecture that high intensity ratings as perceived by the RA's indicate a high arousal of the subject while writing up the narrative and therefore also while experiencing the original situation. Arousal describes a state of activeness in both mind and body that makes humans more vigilant. The degree of arousal can be scaled from low to high. It was previously observed that typically, discrete emotivational[41] states are connected to specific levels of arousal. According to Schachter and Singer (1962) an emotivational state is determined by the physiological arousal and how this arousal is cognitively evaluated. Goetz et al. (2010) find that compassion-inducing stimuli lead to low levels of arousal while sadness, fear and distress-inducing stimuli are associated with high levels of arousal. At the same time sadness, fear and distress are related to avoidant, antisocial and self-protecting tendencies (Gu et al., 2010) which should influence subsequent behavior in the public goods game if the narratives elicited such states in the subjects. To the contrary, compassion motivation is characterized by approaching behavioral tendencies as well as the urge to help and care for another person but at the same time low levels of arousal. Furthermore, there is ample evidence that anger is usually accompanied by states of high arousal and is characterized by antisocial decisions (Lerner and Tiedens, 2006). Therefore, high levels of arousal are indicative of a motivational state that is related to antisocial tendencies.

4.5.2 Human Coding of Motivation and Intensity: Results and Discussion

Result 9: Human coders assign narratives written under Anger a significantly higher human anger score than to narratives stemming from Control. Hypothesis 9 is supported.

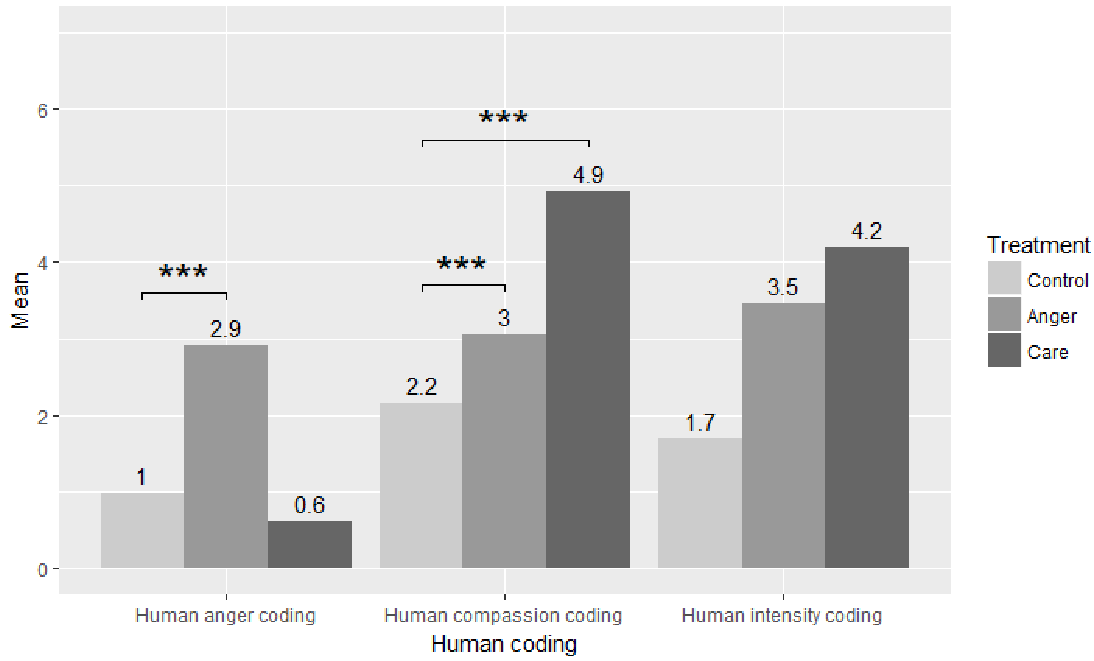
As can be seen by the left of the three groups in Figure Y, the human anger coding of narratives from *Anger* were coded on average with a 2.9 while the RA's on average coded anger motivation in *Control* narratives with an average of 1. By means of a Wilcoxon-Mann-Whitney-Test, this difference is statistically significant at the 1% significance level.

Result 10: Human ratings of compassion motivation are the highest for narratives under Care. The difference compared to Control is statistically significant which presents evidence for hypothesis 10.

[41] Anger was previously studied as a core emotion. However, emotions have a motivational dimension as they cause typical action tendencies. Roseman (2011) calls them therefore emotivations.

We find the result that narratives from *Care* are assigned a significantly higher (at the 1% level, Mann-Whitney-Test) compassion motivation rating than narratives from *Control*.

Figure Y



However, we find at the same time that narratives written under *Anger* are rated with a significantly higher compassion motivation rating than those from *Control* (also significant at the 1% level, Mann-Whitney-Test). While the result for human coding of anger motivation shows that research assistants perceive the degree of anger motivation in narratives stemming from *Care* as even lower compared to those from *Control*, a different picture emerges for the assessment of compassion motivation. Interestingly, the RA's assign narratives stemming from *Care* a higher mean intensity score than those from *Anger*. This is indicative of subjects under *Care* being more aroused than subjects under *Anger* which points towards subjects that wrote about sadness, fear and distress under *Care* as opposed to truly compassionate experiences.[42] Note that the differences between mean *Control* and *Anger* scores as well as between mean *Anger* and *Care* scores for human intensity coding are both statistically significant at the 1% level by means of a Mann-Whitney-Test. Recall also that RA's could rate narratives in the range of 0 – 11. The mean scores over the bars in Figure Y show that no variable on average received a higher rating than 4.9. This can be seen as evidence that the degree to which

[42] By means of analyzing Spearman correlation coefficients, we find that $\boldsymbol{r}_{\text{human compassion, human intensity}}$ to be 0.74 and significant at the 1% level, while the correlation between human anger rating and intensity rating, $\boldsymbol{r}_{\text{human anger, human intensity}}$ is 0.11 and not significant.

motivations and intensities were present in the text was not perceived as very pronounced by the RA's.

Result 11: Human ratings of Anger motivation predict antisocial behavior in the public goods game. No other hypothesized influence of human ratings on behavior, beliefs, attention and perception is detectable. Except for weak suggestive evidence for Anger*, hypothesis 11 is rejected.*

By using OLS regression specifications Table K shows six dependent variables from the public goods game and the influence that the human coding variables of motivation and intensity have in terms of explaining their variation. As can be seen, human assessment of intensity within narratives never has a significant influence on contributions in the public goods game. Strangely, human encoding of compassion motivation within narratives shows a positive influence on subjects' perceptions to judge the public goods game to be of competitive nature. This stands in contrast to hypothesis 0.5 above which states the expectation that *Care* leads subjects to perceive the game form to be significantly more often of a cooperative nature than subjects under *Anger* and *Control.* However, as already indicated by Figure W, subjects under *Care* perceive the public goods game slightly more often to be competitive than subjects under *Control*.

Table K

Dependent variable:	Public good contribution	Belief about other's contribution	Perception of the game (1 = competitive)	Attention to own payoff	Attention to other's payoff	Probability to be conditional contributor
Model	K1	K2	K3	K4	K5	K6
Human Care encoding	0.03 (0.26)	-0.11 (0.24)	0.40* (0.21)	-0.02 (0.25)	0.09 (0.21)	-0.09 (0.20)
Human Anger encoding	0.10 (0.21)	-0.05 (0.20)	0.27 (0.17)	0.22 (0.20)	0.28 (0.17)	-0.32* (0.16)
Human motivational intensity encoding	-0.15 (0.28)	0.06 (0.26)	-0.28 (0.23)	-0.07 (0.27)	-0.17 (0.23)	0.21 (0.23)
Constant	6.07*** (0.64)	5.32*** (0.59)	-2.09*** (0.54)	5.52*** (0.61)	4.04*** (0.52)	1.28*** (0.49)
N	184	184	184	184	184	184
Regression	OLS	OLS	logit	OLS	OLS	Logit
R^2/ AIC	<0.01	<0.01	AIC: 205.72	0.01	0.02	AIC: 207

*** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

However, model K3 in Table K fails to replicate the main finding in result 0.5 that writing narratives under *Anger* increases the perception that the public goods game is a competitive game. This makes the finding on human care coding in model K3 questionable. Potentially, the finding could be rationalized for those *Care* subjects that actually wrote narratives featuring sadness, fear or empathic distress.

As we have seen in Figure H and result 0.7, subjects under Anger are significantly less conditionally cooperative in contributing to a public good. This result is mirrored in the human coding of anger motivation in model K6 in Table K. An increase in RAs' assessments of anger motivation within a narrative is associated with a significant decrease in the probability to be a conditional cooperator. While this might be driven by the raters' abilities to identify the *Anger* narratives, the anger motivation coding OLS coefficients in models K3 and K5 consistently demonstrate the positive influence that human anger coding has on the perception of the game and attention towards other's payoff respectively. It can therefore be summarized that human coding of motivation within narratives shows weak predictive ability in inferring behavior in the public goods game. While human coding of anger motivation shows a significant correlation with the probability to be a conditional contributor, confirming the previous result and intuition, the coefficient estimates for the human compassion motivation rating on dependent variables in the public goods game yields no meaningful results. Potentially, this is due to the content that the *Care* narratives contained which might have been partly driven by sadness, fear or empathic distress.

5 Robustness and Summary of Overall Results

So far, the effect of LDA topics, dictionary method words and human coding within narratives on motivational states and on the public goods game was considered in isolation. Section D in the appendix provides robustness checks of the results from section 4 above. Initially, we contrast the results of the dictionary method and LDA as well as treatment dummies on the ms-p with one another. This is achieved by adding the variables from these three parts into one regression model. This model either predicts compassionate or angry motivational states that subjects self-reported after having written down narratives in the ms-p. We continue to find that an increase in the presence of anger words from the dictionary robustly predicts motivational states of anger. Likewise, the positive influence of LDA topics 4 (help others in public) and 6 (empathy / compassion towards close ones) on compassionate motivational states remains robust after controlling for relevant dictionary word categories and treatment dummies. We further find that LDA topics 1's (frustration at uni / job) and 2's (anger / insult at school) positive impact on subsequent self-reported anger motivation ratings is robust to controlling for relevant word categories from the dictionary. However, the effect is not robust to including dictionary categories and treatment dummies. This indicates that by and large, the results on the ms-p are robust to controlling simultaneously for different factors that show an influence on self-reported motivational states.

Furthermore, the results from the pgg-p are subjected to a robustness analysis. Table K revealed that there is no strong or systematic relationship between how existent a motivation is perceived in a narrative by the human coders and the influence of this narrative on behavior in the public goods game. Since human motivation coding therefore yields no explanatory insights for behavior in the public goods game, we consider it only in so far in the robustness check that the only sensible result of the human coding will be reviewed.[43] Also, as Table F shows, treatment dummies give few insights on what drives behavior, beliefs and perceptions in the public goods game. Therefore, they are also not present in the robustness analysis. Accordingly, the robustness check for the results relevant for the public goods game contains the LDA topics and synthesized dictionary word categories from Table G in one regression model.[44] This procedure confirms that the "friends" word category has a positive and significant influence on contributions and beliefs in the public goods game. Moreover, when subjects use more words from the "family" and "humans" categories in their narratives, then this reduces the attentional focus of these subjects towards their own payoff. This dictionary method result does not survive the comparison with LDA topics within the same model. In terms of the robustness of other dictionary method results, we can confirm the robustness of the results on the "affiliation motivation" and "other references" word categories: The more affiliation words are used, the more attention to other's payoff; the more other references are used the higher was the likelihood to be a conditional cooperator. We can also confirm the stability of the result that LDA topic 4 (help others in public) has a positive and significant influence on contributions and beliefs in the public goods game. Likewise, the effect persists that a relatively larger share

[43] Table E .A2 in the appendix exposes the human coding result that an increase in human anger motivation coding is associated with a significant decrease in the likelihood to be a conditional cooperator to a robustness analysis. The analysis contains the human coding variables, LDA topic variables and dictionary word category variables and regresses them on dependent outcome variables in the public goods game.

[44] Table E .A2 offers the interested reader to observe robustness of previous LDA and dictionary method results when LDA topics, relevant dictionary categories and human motivation coding variables are all in one regression model.

of topic 2 (anger / insult at school) leads to more competitive perceptions of the game form of the public goods game. On the one hand, we cannot confirm the positive impact of a more pronounced presence of topic 1 (frustration at uni / job) on attention towards the own payoff. On the other hand, topic 2 (anger / insult at school) continues to have a negative and significant effect on the likelihood to be a conditional cooperator. Finally, even when LDA topics and dictionary categories are present, an increased human anger encoding is reliably associated with a subject being less likely to be a conditional contributor.

6 Concluding Discussion

This paper provided a study of the causal effects of narratives on behavior and motivation through the lens of psychological motives. To this end, conditional and unconditional contributions, beliefs as well as perceptions and attentional patterns were collected within a controlled laboratory setting after subjects recalled personal narratives. The goal of this study was further to present three different approaches to textual analysis in experimental economics and compare their relative performances in the context of studying narratives. One of these approaches is LDA for which we demonstrated the usefulness for the quantitative analysis of text as data within experimental economics. The LDA method shows convergent validity as topics for which exist clear hypotheses from motivation psychology predict motivational states and at the same time behavioral and psychological responses in the public goods game. An analysis of the robustness of the findings shows that in particular, topic 4 ("help others in public") and topic 2 ("anger at school") have this convergent validity and are the main drivers behind the public goods game results. These results are consistent with insights from motivation psychology. Topic 4 clearly centers around helping actions and doing something for others. This stands in contrast to topic 6 ("empathy / compassion towards close ones") which is prominently defined by words like "sad", "tried", "conversation" and "fear", indicating the presence of empathic distress. Psychologists argue that there exists a decisive difference between empathy and compassion and that empathy can either lead to empathic distress or compassion (Singer and Klimecki, 2014; Klimecki and Singer, 2012). Empathy describes the capacity to share the feeling of others. Empathic distress is a self-protective reaction to suffering of others that can manifest in escaping the situation that causes empathic distress. Importantly, this empathic distress does not lead to the behavioral response of caring for and being prosocial to others. This is however the main behavioral response for states of compassion. Active compassion motivation is therefore hypothesized to lead to prosocial behavior in the public goods game and increased levels of self-reported prosocial or caring motivation. To the extent that topic 4 expresses compassionate content in narratives we find that recalling compassionate narratives leads to compassion motivation and prosocial behavior in a public good context.

We find evidence that LDA topics and word category frequencies are no substitutes to each other. Both anger related topics (1 and 2) as well as anger word category counts have a positive influence on anger motivation when one controls for the influence of the respective other across all treatments. The same holds for behavior in the public goods game. The robustness analysis reveals that LDA topics and word categories rather complement each other in predicting behavior in the public goods game.

The results further indicate that detecting the influence of marker words on motivation and behavior is more delicate than finding these influences with LDA topics. This is particularly true for compassion motivation which suggests a heterogeneity in detectability, inadequate selection of words for the compassion motivation word category, or a larger complexity of the compassion motivation compared to anger motivation. This insight stems from the result that we do not find a significant relationship between word categories supposedly related to compassion and motivational states of compassion as well as prosocial responses in the public goods game. Interestingly, the positive effect of *Care* on subsequent self-reported states of compassion motivation is mostly captured by the topics 4 and 6. This picture however reverses under *Anger*, where it is rather the treatment dummy that drives the effect on motivation rather than the two anger related topics. Given that the *Anger* treatment did lead to increases in anger motivation, we interpret this as suggesting that the actual topics that subjects wrote about under *Anger* are broad and could not be approximatively and exhaustively captured by the two LDA topics 1 and 2. Under *Care* however, the actual narrative content was likely to be more homogeneous and could therefore be represented in the two LDA topics 4 and 6.

Exemplarily, topic 6 also highlights limitations of the LDA method. Topics need to be subjectively filled with an identity and meaning which is not an unambiguous task. This is connected to the feature that individual words within LDA topics may appear prominently in different topics which makes a selective attribution of meaning difficult and increases degrees of freedom that researchers have in the interpretation of LDA topics and their implications. The present study overcomes these concerns partly due to the design and context of a controlled laboratory experiment in which subjects were asked to recall six particular narratives that were picked up by LDA and easily interpretable given the instructions of the experiment.

Human encoding of motivations within narratives is not clearly associated with predicting behavior in the game. However, the encoding through assistants gave insights about intensity of motivations in narratives. By means of the dictionary method we find that *Anger* leads to more use of words from the care category than *Control*. This finding is mirrored by the human assessment of motivation within narratives. It reveals that compassion motivation was found to be significantly higher in anger narratives than in control narratives. The presence of anger and compassion motivation as well as overall motivational intensity was rated as not very high. This is in line with previous analyses on narrative recollection method to induce emotional or motivational states: the method works to induce target states in the right direction but the effect is not particularly strong (Angie et al. 2011). While letting subjects recollect own narratives has the advantage of featuring less potential for experimenter demand effects, this comes at the cost that control is lost about the content of the narratives. Through human encoding we found evidence that the general intensity within *Care* narratives was perceived higher as those in *Anger* narratives. This is suggestive of higher levels of arousal within subjects under *Care* which is not backed by psychological evidence about states of compassion that are associated with low physiological arousal (Goetz et al., 2010; Stellar et al., 2014). Not only for these reasons should future research attempt to investigate the causal influence of narratives that do not stem from the own biography and identity on behavior.

Digitalization and the evolution of the internet have led to an unprecedented increase in the amount of data being available for economic analyses. In particular, data from social media like users' posts and as well as texts that are put online or where digitized from sources like historic archives contain information that become newly available to economists (Gentzkow et al., 2017). The digital and often real-time availability of such data provides new opportunities to incorporate them into economists' quantitative empirical research targeted at studying narratives. Recent methodological developments have provided tools like LDA to quantify the information content of narratives with the help of statistical models. Much more research on the influence of narratives from parties such as companies, governments, think tanks and historical sources will therefore be conducted in the future as research methods advance and are taken up. This research should make use of latest developments in textual analysis methods. Exemplarily, part of speech and collocations recognition, tf-idf ranking and tone recognition are methodological catchwords in textual analysis that were not showcased in this study but show great promise for conducting textual analysis.

Research on narratives will yield insights on deliberation processes, feedback mechanisms between narrative exposure and behavior as well as storytelling and manipulations by pressure groups. Hopefully, this will lead to positive policy recommendations that take into account the influence of narratives. Building on insights from psychology, the present study attempted to analyze the motivational implications of narratives and how they manifest in behavior. Such "micro-level" narrative research could be applied to data that observes texts like diaries, private letters and therapists' therapy notes in combination with life decisions and events to study human deliberation processes in more depth. Our study provided an exposition of the methods that can be applied for such research endeavors and provided first insights into the weaknesses and strengths of these approaches.

Appendix

A LDA

Latent Dirichlet Allocation (LDA, Blei et al., 2003) is a generative statistical model that finds latent semantic topics in a set of documents (corpus) in an unsupervised way. Latent topics are found in the corpus by identifying groups of words which regularly appear together within documents.

Introduction of terminology and notation:

1. A token $t \epsilon [1, \ldots, V]$ is the smallest form of discrete data.
2. A document m is a sequence of N tokens denoted by $m = (t_1, t_2, \ldots, t_N)$, where t_n is the n^{th} token in m.
3. A corpus is a collection of M documents denoted by $\mathcal{M} = [m_1, m_2, \ldots, m_M]$. $\boldsymbol{\mathcal{M}}$ will be used as a term representing the corpus in the conditional parameter statements and learning algorithm below.
4. A topic $k \epsilon [1, \ldots, K]$ is a probability distribution over the vocabulary of V tokens.
5. α is a parameter of the Dirichlet distribution and a LDA hyperparameter. It gives a prior belief about topic sparsity/uniformity in the documents, i.e. does one document usually contain only a few or most of the topics.
6. β is a parameter of the Dirichlet distribution and a LDA hyperparameter. It gives a prior belief about token sparsity/uniformity in the topics, i.e. does one topic usually contain only a few or most of the words.

The model assumes the following data generating process. For each m in $\mathcal{M}$, a topic weight vector θ_m (modeled by a Dirichlet random variable α) of dimension K is drawn, that informs about which topics appear in m with which weight. Following this, each t in m is mapped onto exactly one topic according to θ_m. Now, each t in m is generated by using the topic it was mapped onto. This happens by drawing a token from the probability distribution over tokens of the assigned topic which is denoted by $\phi_{i,j}$. More formally:

For each m in $\mathcal{M}$:

1. Choose a K-dimensional topic weight vector θ_m from the distribution $p(\theta|\alpha) = Dirichlet(\alpha)$.
2. For each token indexed by $n \epsilon [1, \ldots, N]$ in a document m:
   (a) Choose a topic $k_j \epsilon [1, \ldots, K]$ from the multinomial distribution $p(k_j = k|\theta_m) = \theta_m^k$ resulting in $\theta_{n,m}^{k_j}$ the token topic assignment of token n in document m_m with topic k_j. $\boldsymbol{\theta_m^k}$ describes the current topic assignment for each of the V tokens in the corpus.
   (b) Given the chosen topic k_j, draw a token t_n from the probability distribution $p(t_n = i|k_j = j, \beta) = \phi_{i,j}$.

A.1 Learning

In order to fulfill the task of finding topic representations of each document and the tokens associated with each topic, posterior inference needs to be applied. For this, the above generative process needs to be reversed in order to learn the posterior distributions of the latent variables (topics, topic shares for each document) in the corpus. In LDA, this means solving the following equation:

$$p(\theta, \phi, \boldsymbol{\theta_m^k}|\boldsymbol{\mathcal{M}}, \alpha, \beta) = \frac{p(\theta, \phi, \boldsymbol{\theta_m^k}, \boldsymbol{\mathcal{M}}|\alpha, \beta)}{p(\boldsymbol{\mathcal{M}}|\alpha, \beta)} \tag{A 1}$$

This equation is however impossible to be computed since $p(\boldsymbol{\mathcal{M}}|\alpha, \beta)$ cannot be computed exactly. In order to infer k topics with the latent distributions θ and ϕ however, one can use approximative inference techniques such as collapsed Gibbs sampling. Gibbs sampling is one member of the Markov Chain Monte Carlo class of algorithms (Gilks et al., 1995). This algorithm aims to build a Markov chain that arrives at the distribution of the latent variables conditional on the corpus as its steady state. In our case, we are interested of the posterior distributions of the topic weight vectors θ_m, the topic-word distribution $\phi_{i,j}$ and the topic word assignment θ_m^k. Note that the conditional distributions of θ_m as well as $\phi_{i,j}$ can be calculated just from θ_m^k. The *collapsed* Gibbs sampler uses this insight to facilitate the algorithm and just sample over θ_m^k. To this end the sampler needs to calculate the probability that a topic k_j is assigned to token t_n in document m given the mapping between all other tokens with topics in the entire corpus. Or, formally:

$$p(\theta_{n,m}^{k_j} | \boldsymbol{\theta}_{-\boldsymbol{n},\boldsymbol{m}}^{\boldsymbol{k_j}},\ \boldsymbol{\mathcal{M}}, \alpha, \beta) \tag{A 2}$$

where $\boldsymbol{\theta}_{-\boldsymbol{n},\boldsymbol{m}}^{\boldsymbol{k_j}}$ means all other current token – topic allocations except $\theta_{n,m}^{k_j}$. The complete derivation of the collapsed Gibbs sampling equation for LDA can be found in Carpenter (2010) or Heinrich (2008). These sources contain the solution to (A 2). Following Darling (2011) the sampling algorithm can be conceptualized as follows:

Input: $\boldsymbol{\mathcal{M}}$

Output: $\boldsymbol{\theta}_{\boldsymbol{m}}^{\boldsymbol{k}}$ and therefore implicitly counts how many tokens within a document can be attributed to which topic $(n_{m,k})$, counts how often a token was assigned to a topic $(n_{t,k})$ and counts how often topics appear in the corpus $(n_{k,\boldsymbol{\mathcal{M}}})$.

Start

Randomly assign a $\boldsymbol{\theta}_{\boldsymbol{m}}^{\boldsymbol{k}}$.

For i in # iterations do:

  For each t in $\boldsymbol{\mathcal{M}}$ do:

    Assign t to a k; due to this assignment in this iteration, change counters $n_{m,k}, n_{t,k}, n_{k,\boldsymbol{\mathcal{M}}}$

      For each topic in $k\ \in [0, \dots, 1-K]$:

      Compute new conditional probabilities that a token gets assigned a topic $p(\theta_m^k | \cdot) =$

      $(n_{m,k}\ \times \alpha)(\frac{n_{t,k}+\beta}{n_{k,\boldsymbol{\mathcal{M}}}+\beta\times\boldsymbol{\mathcal{M}}})$

      End

    Assign t a new topic by sampling from $p\left(\boldsymbol{\theta}_{\boldsymbol{m}}^{\boldsymbol{k}} \middle| \cdot\right)$; due to this assignment in this iteration, change counters $n_{m,k}, n_{t,k}, n_{k,\boldsymbol{\mathcal{M}}}$

  End

End

Return $\boldsymbol{\theta}_{\boldsymbol{m}}^{\boldsymbol{k}}$, $n_{m,k}, n_{t,k}, n_{k,\boldsymbol{\mathcal{M}}}$

End

B Topics

The lists of words are ordered in descending order: the first word in each cell in the right column has the highest probability in the topic.

Topic 1: “Frustration at school / job”	klausur, studium, aufgab, fuehlt, entschied, kolleg, frag, pruefung, bekam, geld, not, praktikum, stellt, sem, arbeitet, frustriert, unternehm, studi, hausarbeit, lern, punkt, gelernt, trotz, mitarbeit, frustrier (~19%)
Topic 2: „Insult and anger”	mensch, klass, schul, leut, red, beleid, lehr, grupp, liess, erinn, beleidigt, schlimm, kopf, gehoert, versucht, nam, fussball, sass, wuetend, fuehlt, beschimpft, schwer, fing, verhalt, unterricht (~19%)
Topic 3: „Yesterday“	Freund, freundin, gegang, wohnung, gemeinsam, gegess, traf, gefahr, entspannt, aufgestand, fing, mued, aufgewacht, beschloss, anschluss, broetch, fuhr, wochen, ass, lust, haus, naeh, bier, getrunk, sass (~26%)
Topic 4: „Help (elderly in public)”	sah, stand, bus, fuhr, aelt, strass, fiel, hilf, dacht, schwer, schoen, auto, helf, bedankt, platz, hand, lief, bemerkt, fahr, dam, blick, bot, kiel, freut, wartet (~24%)
Topic 5: „Everyday routine“	uni, ess, haus, fruehstueck, vorles, fahr, bett, dusch, anschliess, steh, kaffe, weck, treff, schlaf, koch, mensa, freund, aufsteh, schau, bus, klingelt, nachricht, mittag, nachmittag, fahrrad (~33%)

Topic 6: „Compassion (towards friends and family)”	freundin, freund, gefuehl, vat, kind, wusst, elt, famili, leb, mutt, erzaehlt, schwest, versucht, gespraech, helf, haus, traurig, aug, merkt, wohnt, lieb, oma, schoen, angst, gluecklich (~25%)

C Dictionary Method Results

C.1 Overall Word Category Counts

This table shows the LIWC and motivation word categories that were evaluated as well as the average count in these categories of narrative texts at the treatment level after text pre-processing.

Table 1: Average counts of words from word categories by induced treatments

	Motivational States Part			Public goods game Part		
	Control	Anger	Care	Control	Anger	Care
“Pronoun”	1.13 (1.1)	1.76 (1.39)	2.98 (2.17)	0.93 (1.19)	2.34 (1.62)	2.43 (1.6)
“I”	0.53 (0.76)	0.57 (0.77)	1.07 (1.14)	0.33 (0.66)	0.73 (0.96)	0.91 (0.8)
„We“	0.11 (0.38)	0.24 (0.53)	0.17 (0.53)	0.14 (0.35)	0.16 (0.45)	0.2 (0.47)
“Self”	0.64 (0.77)	0.81 (0.99)	1.24 (1.16)	0.47 (0.78)	0.89 (0.98)	1.11 (0.94)
“You”	0.04 (0.21)	0.1 (0.37)	0.28 (0.62)	0.0 (0.0)	0.21 (0.52)	0.22 (0.48)
“Other”	0.27 (0.58)	0.21 (0.42)	0.85 (1.07)	0.04 (0.19)	0.47 (0.69)	0.69 (0.97)
“Mixed text components”	Control	Anger	Care	Control	Anger	Care
“Negate”	0.2 (0.5)	0.79 (0.72)	0.5 (0.66)	0.19 (0.52)	0.76 (0.86)	0.49 (0.83)
“Assent”	0.51 (0.73)	0.95 (1.13)	0.57 (0.72)	0.12 (0.33)	1.6 (1.46)	0.52 (0.81)
“Article”	0.38 (0.58)	0.31 (0.56)	0.2 (0.45)	0.37 (0.67)	0.42 (0.59)	0.37 (0.65)
“Preps”	4.33 (2.57)	5.29 (3.26)	4.91 (2.80)	4.84 (2.26)	5.71 (2.99)	4.69 (2.49)
“Numbers”	3.82 (2.35)	2.67 (1.96)	2.15 (1.73)	4.39 (3.25)	1.98 (1.49)	1.92 (1.74)

"Affect"	14.0 (6.0)	18.81 (7.44)	18.17 (7.19)	12.86 (5.13)	18.18 (6.09)	16.68 (5.29)
"Positive emotion"	10.58 (4.76)	10.4 (4.63)	11.52 (5.24)	9.86 (4.54)	10.45 (4.56)	10.92 (4.24)
"Positive feeling"	0.56 (0.76)	0.71 (1.02)	1.76 (1.48)	0.65 (0.92)	0.71 (0.84)	1.66 (1.63)
"Optimism"	1.62 (1.61)	2.02 (1.57)	1.87 (1.60)	1.46 (1.35)	1.82 (1.56)	2.0 (1.9)
"Negative emotion"	3.42 (2.46)	8.43 (4.58)	6.72 (3.64)	3.11 (2.6)	7.77 (3.83)	5.82 (2.56)
"Sad"	1.47 (1.25)	2.02 (1.98)	2.8 (2.25)	1.33 (1.26)	1.85 (1.66)	2.74 (1.91)
"Cognitive mechanism"	10.51 (4.48)	17.86 (6.9)	15.61 (6.01)	11.02 (4.27)	17.06 (5.98)	14.75 (6.21)
"Cause"	1.47 (1.39)	2.52 (1.63)	1.89 (1.8)	1.49 (1.15)	3.02 (1.93)	1.82 (1.54)
"Insight"	4.8 (2.18)	6.81 (3.42)	6.85 (3.59)	4.89 (2.46)	5.92 (3.03)	6.05 (3.46)
"Discrepancy"	0.49 (0.76)	1.48 (1.29)	1.07 (1.22)	0.58 (0.94)	1.08 (1.15)	0.97 (1.02)
"Inhibition"	0.51 (0.87)	0.93 (1.2)	0.33 (0.56)	0.53 (0.87)	0.97 (1.14)	0.45 (0.73)
"Tentative"	1.53 (1.56)	2.52 (2.16)	2.8 (2.07)	1.46 (1.58)	2.27 (1.74)	2.45 (1.78)
"Certain"	1.82 (1.7)	4.24 (2.51)	3.17 (2.29)	2.25 (1.79)	4.35 (2.33)	3.77 (2.82)
"Social"	10.44 (4.62)	15.69 (6.57)	17.93 (6.48)	9.14 (4.51)	16.98 (6.37)	16.09 (6.43)
"Communication"	4.24 (2.39)	8.38 (4.02)	6.96 (3.97)	4.16 (2.33)	8.48 (3.87)	5.68 (3.17)
"Other reference"	0.49 (0.76)	1.0 (1.01)	1.61 (1.58)	0.47 (0.68)	1.24 (1.22)	1.37 (1.29)
"Friends"	4.11 (2.53)	2.36 (2.21)	3.07 (2.74)	3.16 (2.38)	2.08 (2.26)	2.65 (2.37)
"Family"	0.89 (1.25)	1.52 (1.76)	4.17 (2.95)	0.82 (1.82)	2.31 (2.83)	3.98 (3.76)
"Humans"	0.18 (0.49)	1.88 (2.35)	3.3 (2.13)	0.39 (1.08)	2.74 (2.53)	3.49 (2.74)
"Time"	21.71 (6.25)	13.93 (4.71)	15.09 (4.66)	23.65 (6.41)	14.85 (5.11)	14.52 (4.85)
"Past"	15.24 (4.43)	20.79 (7.51)	22.33 (7.23)	15.04 (4.16)	20.69 (6.6)	19.63 (5.6)
"Present"	30.18 (8.06)	19.31 (9.08)	18.11 (6.65)	27.61 (6.58)	16.68 (5.74)	16.45 (7.03)
"Future"	0.84 (0.95)	0.45 (0.55)	0.3 (0.63)	0.93 (0.96)	0.34 (0.57)	0.37 (0.57)

“Space”	5.44 (3.16)	5.62 (3.2)	6.54 (3.67)	5.6 (2.8)	6.27 (3.27)	6.23 (3.53)
“Up”	0.84 (1.07)	0.55 (0.71)	0.7 (0.76)	0.74 (0.72)	0.56 (0.88)	0.52 (0.75)
“Down”	0.13 (0.34)	0.21 (0.47)	0.28 (0.58)	0.16 (0.41)	0.29 (0.55)	0.22 (0.45)
“Incl”	1.16 (1.17)	0.86 (1.03)	1.52 (1.31)	1.19 (1.47)	1.24 (1.14)	0.89 (0.87)
“Excl”	0.67 (0.74)	1.38 (1.55)	1.52 (1.5)	0.77 (0.95)	2.26 (1.81)	1.25 (1.23)
“Motion”	10.09 (4.67)	4.74 (3.21)	6.02 (3.26)	9.21 (3.84)	3.98 (2.22)	5.58 (2.6)
“Occupation”	21.04 (5.35)	17.81 (7.33)	12.43 (5.41)	22.65 (7.55)	17.79 (6.19)	11.08 (5.87)
“School”	8.51 (3.78)	7.29 (4.95)	3.65 (3.0)	10.07 (4.55)	7.24 (4.53)	3.71 (3.97)
“Job”	7.64 (4.16)	6.33 (4.07)	4.28 (3.1)	7.75 (3.96)	6.79 (3.78)	3.78 (2.62)
“Leisure activity”	12.38 (4.46)	4.74 (3.58)	4.96 (2.95)	11.11 (4.71)	4.5 (3.89)	4.26 (2.87)
“Home”	7.87 (3.18)	1.64 (1.86)	3.24 (2.5)	7.18 (3.38)	1.53 (2.04)	2.82 (2.3)
“Sports”	3.38 (2.01)	2.31 (2.7)	0.78 (1.01)	2.86 (2.06)	2.03 (3.2)	0.66 (1.09)
“TV”	1.2 (1.42)	0.21 (0.42)	0.24 (0.71)	1.04 (1.39)	0.27 (0.83)	0.18 (0.43)
“Music”	0.69 (1.1)	0.76 (0.82)	0.98 (1.27)	0.81 (0.99)	1.03 (1.45)	0.63 (0.88)
“Money”	2.29 (2.19)	2.02 (2.16)	2.04 (2.23)	2.12 (1.92)	1.98 (2.34)	2.0 (2.38)
“Metaphysical issues”	0.89 (0.93)	1.64 (1.56)	2.0 (1.45)	0.84 (0.86)	1.55 (1.4)	2.14 (1.82)
“Relig”	0.62 (0.78)	1.33 (1.52)	1.02 (1.16)	0.53 (0.76)	0.94 (1.1)	1.05 (1.35)
“Death”	0.27 (0.54)	0.31 (0.64)	1.02 (1.22)	0.32 (0.51)	0.65 (0.93)	1.2 (1.44)
“Physical”	11.11 (3.82)	2.83 (2.38)	4.43 (2.86)	11.84 (4.2)	3.18 (2.87)	3.85 (2.86)
“Body”	2.02 (1.69)	1.93 (1.85)	3.15 (2.77)	2.37 (1.77)	2.26 (2.19)	2.72 (2.29)
“Sex”	0.2 (0.5)	0.24 (0.58)	0.52 (0.81)	0.21 (0.45)	0.29 (0.69)	0.38 (0.84)
“Eat”	6.09 (2.33)	0.62 (1.36)	0.54 (0.91)	6.46 (3.33)	0.82 (1.66)	0.75 (1.08)
“Sleep”	3.0 (1.89)	0.24 (0.48)	0.48 (0.66)	2.21 (1.6)	0.18 (0.56)	0.57 (1.13)

"Grooming"	2.18 (2.0)	0.14 (0.42)	0.33 (0.73)	2.23 (2.11)	0.16 (0.52)	0.09 (0.34)
"Motives and others"	Control	Anger	Care	Control	Anger	Care
"Swear"	0.11 (0.32)	0.88 (1.25)	0.22 (0.55)	0.04 (0.19)	0.71 (1.15)	0.22 (0.62)
"Fillers"	0.04 (0.21)	0.0 (0.0)	0.02 (0.15)	0.02 (0.13)	0.03 (0.18)	0.02 (0.12)
"Achievement"	10.09 (4.13)	8.69 (3.26)	8.15 (3.74)	9.51 (3.97)	9.18 (3.56)	6.85 (3.11)
"Affiliation"	0.56 (0.69)	0.5 (0.59)	0.26 (0.57)	0.46 (0.8)	0.71 (0.91)	0.31 (0.79)
"Anger"	4.22 (2.45)	4.88 (2.87)	2.52 (1.86)	4.02 (2.39)	4.48 (2.48)	2.26 (1.57)
"Care"	2.4 (1.7)	3.14 (2.27)	4.33 (2.56)	2.61 (2.13)	3.98 (1.93)	4.45 (2.77)
"Consumption"	1.22 (1.62)	1.5 (1.49)	1.07 (1.02)	1.39 (1.16)	1.23 (1.3)	1.09 (1.23)
"Fear"	1.02 (1.1)	2.17 (1.68)	2.74 (2.13)	1.16 (1.1)	2.48 (2.03)	1.95 (1.44)
"Power"	3.73 (2.41)	6.43 (2.96)	5.57 (2.91)	4.49 (2.71)	6.87 (2.78)	6.06 (2.85)

Standard deviations in parentheses. In both parts the following categories were dichotomized: We, Negate, Assent, Article, Inhibition, Down, TV, Sex, Swear, Fillers, Affiliation. Note that this table shows mean category counts after pre-processing. This explains why the categories I, We, Self, You, Other, Negate, Assent and Article have so few counts: these categories consist of typical stopwords that were filtered out. In order to create more normally distributed variables, three new summary variables were produced. The variable PosEmo consists of the LIWC categories positive feeling, positive emotion and optimism. Likewise, the variable NegEmo sums up the categories negative emotion and sad. Finally, the variable Otherref contains the variables Other and Otherreference. Words from narratives and words from categories were tokenized and stemmed before the occurrences were counted.

C.2 Word Category Manipulation Results

Hypothesis C.2.1 –The *Care* treatment leads to the highest counts of care words per subject relative to narratives under *Anger* and *Control*.

Hypothesis C.2.2 –The *Anger* treatment leads to the highest counts of used anger words per subject compared to the *Care* and *Control* treatments.

Result C.2.1: The use of care words differs significantly between treatments and is significantly the highest in the narratives of the Care treatment. Hypothesis C.2.1 finds support.

Result C.2.2: The use of anger words differs significantly between treatments and is the highest in the narratives of the Anger treatment while it is significantly the lowest in the narratives of the Care treatment. Hypothesis C.2.2 is supported.

Table C

	"Motivational states part"			"Public goods game part"		
Motives	*Control*	*Anger*	*Care*	*Control*	*Anger*	*Care*
"Achievement"	10.09 (4.13)	8.69 (3.26)	8.15 (3.74)	9.51 (3.97)	9.18 (3.56)	6.85 (3.11)
"Affiliation"	0.56 (0.69)	0.5 (0.59)	0.26 (0.57)	0.46 (0.8)	0.71 (0.91)	0.31 (0.79)
"Anger"	**4.22** **(2.45)**	**4.88** **(2.87)**	**2.52** **(1.86)**	**4.02** **(2.39)**	**4.48** **(2.48)**	**2.26** **(1.57)**
"Care"	**2.4** **(1.7)**	**3.14** **(2.27)**	**4.33** **(2.56)**	**2.61** **(2.13)**	**3.98** **(1.93)**	**4.45** **(2.77)**
"Consumption"	1.22 (1.62)	1.5 (1.49)	1.07 (1.02)	1.39 (1.16)	1.23 (1.3)	1.09 (1.23)
"Fear"	1.02 (1.1)	2.17 (1.68)	2.74 (2.13)	1.16 (1.1)	2.48 (2.03)	1.95 (1.44)
"Power"	**3.73** **(2.41)**	**6.43** **(2.96)**	**5.57** **(2.91)**	**4.49** **(2.71)**	**6.87** **(2.78)**	**6.06** **(2.85)**

Mean counts of motivation category words by treatment and for both parts of the experiment (ms-p and pgg-p). Bold means significant differences at below 1% significance level between treatments as measured by KW test for both parts. For the pgg-p on the right side, all categories are significantly different from another as indicated by KW test except consumption category. This is potentially driven by larger sample size. For the ms-p and the pgg-p, *Anger* leads to the highest usage of anger motivation words, but not significantly so. What is however significant at the 1 % level is the reduction in the usage of anger words in *Care* narratives compared to *Control*. Writing about compassionate and helping memories reduces the usage of anger motivation words. As a side result, we also find evidence for a finding that was presented by Schultheiss (2013) which states that power and anger are positively associated. In both samples, *Anger* leads to significantly more use of power motivation words than *Control* (1% significance level).

D Robustness Checks

Table D

Dependent variable: Self-reported compassion motivation rating in “motivational states part”		
Model	D1	D2
Care treatment dummy	-	-1.19
	(-)	(1.18)
Anger treatment dummy	-	-1.67
	(-)	(1.20)
Topic 1: Frustration at uni / job	-0.64	1.83
	(1.59)	(2.39)
Topic 2: Anger at school	-2.33	0.68
	(1.62)	(2.72)
Topic 4: Help others in public	4.86***	6.84***
	(1.78)	(2.49)
Topic 6: Empathy / compassion towards close ones	3.48*	5.03*
	(2.00)	(2.76)
“Other reference”	0.03	0.02
	(0.11)	(0.11)
“Positive emotion”	0.05	0.04
	(0.03)	(0.03)
“Care”	0.01	0.01
	(0.09)	(0.09)
“Affiliation”	0.56	0.51
	(0.38)	(0.39)
“Tentative”	-0.11	-0.11
	(0.10)	(0.10)
“Communication”	-0.06	-0.04
	(0.05)	(0.06)
“Friends”	-0.09	-0.06
	(0.08)	(0.09)
“Family”	-0.09	-0.08
	(0.10)	(0.10)
“Humans”	0.23	0.32
	(0.52)	(0.53)
Constant	3.94***	3.35***
	(0.69)	(0.82)
N	133	133
R^2	0.22	0.23

OLS; *** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

Table E

Dependent variable: Self-reported anger motivation rating in "motivational states part"		
Model	E1	E2
Care treatment dummy	-	1.99**
	(-)	(0.94)
Anger treatment dummy	-	2.05**
	(-)	(0.95)
Topic 1: Frustration at uni / job	2.70**	-0.24
	(1.33)	(1.93)
Topic 2: Anger at school	2.67**	-1.10
	(1.29)	(2.23)
Topic 4: Help others in public	0.52	-2.66
	(1.36)	(1.98)
Topic 6: Empathy / compassion towards close ones	-0.95	-4.32**
	(1.26)	(2.12)
"Power"	0.04	0.06
	(0.06)	(0.06)
"Anger"	0.12*	0.13*
	(0.07)	(0.07)
"Tentative"	-0.11	-0.13
	(0.09)	(0.08)
Constant	0.86*	1.53**
	(0.51)	(0.59)
N	133	133
R^2	0.15	0.18

OLS; *** = $p < 0.01$; ** = $p < 0.05$; * = $p < 0.1$;

Table E.A

Dependent variable:	Public good contribution	Belief about other's contribution	Perception of the game (1 = competitive)	Attention to own payoff	Attention to other's payoff	Probability to be conditional contributor
Topic 1: Frustration at uni / job	1.99 (2.39)	3.16 (2.16)	1.81 (1.86)	3.64 (2.25)	1.46 (1.92)	-0.64 (1.89)
Topic 2: Anger at school	0.34 (2.16)	3.90** (1.96)	2.63** (1.65)	-3.22 (2.03)	1.32 (1.74)	-3.93** (1.67)
Topic 4: Help others in public	3.73* (2.11)	3.73* (1.92)	-1.30 (1.78)	2.20 (1.99)	2.02 (1.70)	0.16 (1.80)
Topic 6: Empathy / compassion towards close ones	-1.59 (2.07)	-0.86 (1.88)	1.36 (1.68)	-1.10 (1.95)	-2.35 (1.67)	-0.56 (1.73)
"Power"	0.10 (0.09)	0.09 (0.08)	-0.07 (0.08)	0.01 (0.09)	0.02 (0.07)	0.04 (0.07)
"Anger"	0.13 (0.11)	0.06 (0.10)	0.00 (0.09)	-0.04 (0.10)	-0.12* (0.08)	-0.06 (0.08)
"Other reference"	0.03 (0.16)	0.05 (0.14)	0.21* (0.12)	0.04 (0.15)	0.09 (0.13)	0.31** (0.15)
"Positive emotion"	-0.04 (0.04)	-0.02 (0.04)	0.05 (0.04)	-0.01 (0.04)	-0.01 (0.04)	-0.01 (0.04)
"Care"	-0.10 (0.11)	-0.13 (0.10)	-0.11 (0.09)	0.01 (0.10)	-0.01 (0.09)	0.00 (0.09)
"Affiliation"	0.30 (0.28)	-0.14 (0.25)	-0.07 (0.22)	0.26 (0.26)	0.37* (0.22)	0.59* (0.31)
"Tentative"	0.21 (0.15)	0.05 (0.13)	0.05 (0.12)	0.10 (0.14)	-0.06 (0.12)	-0.11 (0.12)
"Communication"	-0.13 (0.08)	-0.20*** (0.07)	-0.03 (0.06)	0.09 (0.07)	0.04 (0.06)	0.04 (0.07)
"Friends"	0.27** (0.11)	0.29*** (0.10)	0.05 (0.09)	0.00 (0.11)	0.21** (0.09)	-0.03 (0.09)
"Family"	0.04 (0.10)	0.01 (0.09)	-0.09 (0.08)	-0.12 (0.09)	-0.01 (0.08)	-0.06 (0.08)
"Humans"	0.03 (0.12)	0.04 (0.11)	0.15* (0.09)	-0.17 (0.11)	-0.09 (0.09)	0.09 (0.10)
Constant	4.34*** (0.96)	3.79*** (0.87)	-2.25*** (0.82)	5.35*** (0.90)	3.73*** (0.77)	1.46* (0.76)
N	184	184	184	184	184	184
R^2/ *AIC*	0.09	0.11	AIC: 217.34	0.11	0.11	218.75

Table E .A2

Dependent variable:	Public good contribution	Belief about other's contribution	Perception of the game (1 = competitive)	Attention to own payoff	Attention to other's payoff	Probability to be conditional contributor
Human Care encoding	-0.09 (0.28)	-0.22 (0.25)	0.33 (0.24)	0.06 (0.26)	0.03 (0.22)	-0.13 (0.23)
Human Anger encoding	-0.02 (0.27)	-0.23 (0.24)	0.10 (0.23)	0.33 (0.25)	0.30 (0.22)	-0.39* (0.22)
Human motivational intensity encoding	-0.24 (0.31)	0.00 (0.28)	-0.36 (0.26)	0.05 (0.29)	0.01 (0.25)	0.25 (0.25)
Topic 1: Frustration at uni / job	2.72 (2.53)	3.89* (2.29)	2.45 (2.02)	2.47 (2.38)	0.46 (2.03)	-0.05 (2.05)
Topic 2: Anger at school	0.70 (2.35)	4.80** (2.13)	2.74 (1.82)	-4.60** (2.21)	0.07 (1.89)	-2.81 (1.80)
Topic 4: Help others in public	4.56* (2.19)	4.16** (1.99)	-1.31 (1.89)	2.20 (2.06)	2.16 (1.76)	-0.39 (1.91)
Topic 6: Empathy / compassion towards close ones	-0.29 (2.27)	-0.31 (2.06)	1.75 (1.84)	-1.10 (2.14)	-2.14 (1.83)	-1.49 (1.90)
"Power"	0.11 (0.09)	0.10 (0.08)	-0.08 (0.08)	0.00 (0.09)	0.01 (0.07)	0.05 (0.08)
"Anger"	0.12 (0.11)	0.06 (0.10)	0.02 (0.09)	-0.07 (0.10)	-0.16* (0.09)	-0.03 (0.08)
"Other reference"	0.05 (0.16)	0.08 (0.14)	0.20* (0.12)	0.02 (0.15)	0.07 (0.13)	0.33** (0.15)
"Positive emotion"	-0.04 (0.05)	-0.02 (0.04)	0.04 (0.04)	0.00 (0.04)	0.00 (0.04)	-0.02 (0.04)
"Care"	-0.08 (0.11)	-0.11 (0.10)	-0.12 (0.09)	0.00 (0.10)	-0.02 (0.09)	0.01 (0.09)
"Affiliation"	0.25 (0.28)	-0.15 (0.26)	-0.08 (0.22)	0.26 (0.26)	0.36* (0.23)	0.60** (0.30)
"Tentative"	0.24 (0.15)	0.09 (0.14)	0.05 (0.12)	0.08 (0.14)	-0.07 (0.12)	-0.11 (0.12)
"Communication"	-0.12 (0.08)	-0.19** (0.07)	-0.04 (0.06)	0.09 (0.07)	0.03 (0.06)	0.04 (0.07)
"Friends"	0.28** (0.11)	0.32*** (0.10)	0.05 (0.09)	-0.02 (0.11)	0.20** (0.09)	-0.01 (0.09)
"Family"	0.03 (0.10)	0.02 (0.09)	-0.09 (0.09)	-0.15 (0.10)	-0.04 (0.08)	-0.03 (0.09)
"Humans"	0.07 (0.12)	0.06 (0.11)	0.16* (0.09)	-0.18 (0.11)	-0.09 (0.10)	0.07 (0.10)
Constant	4.59*** (0.99)	4.07*** (0.90)	-2.47*** (0.87)	5.19*** (0.93)	3.63*** (0.80)	1.60** (0.79)
N	184	184	184	184	184	184
R^2/ *AIC*	0.11	0.12	AIC: 220.55	0.12	0.12	220.66

E Translated Instructions for Human Coding

Compassion motivation

Persons motivated by compassion share with others without having an ulterior motive of expecting a favor in return. They care for the wellbeing of others and make decisions that benefit others; in particular when these others are close ones. Compassion motivation leads one to help others.

Compassion motivation is about intimacy, to look after someone and to shepherd someone, having empathy for a person and expressing this empathy through supporting behavior as well as being present for others. Compassion-motivated persons notice how needy of care others really appear to them. Compassion motivation is associated with positive emotions like warmth, love, affection, satisfaction, fulfillment, empathy, quiescence, choppiness and tender-heartedness.

Persons that are motivated by compassion are susceptible to signals of affection like for example a soft, soothing way of communicating, compassionate physical contact, laughing and smiling, or eye-contact. Moreover, they pay attention to the expression of the eyes. Also, there exists increased attention vis-à-vis needs of family members.

Compassion-motivated persons appear to others as warm, inartificial and caring. Furthermore, these persons tend to reveal themselves and to be trustful.

Anger motivation

Anger-motivated persons lean towards aggressive and risk-seeking behavior. At the same time, these persons are less cooperative and make decisions that do not take into account the wellbeing of others.

Persons with an active anger motivation feel negatively valenced emotions like anger, rage, or feelings of revenge and retaliation for the most part. In particular, these states manifest themselves if a person motivated by anger perceives her physical and psychological integrity as threatened. Causes for such a threat perception are transgressions of behavioral norms by others, own frustration and extreme physical conditions (hunger, coldness, injury). Under certain conditions, anger motivation can be associated with positive emotions like pride, power or self-esteem. Oftentimes under anger, heart rate and blood pressure increase.

Persons motivated by anger increasingly react to angry faces and outbursts of fury. Also, they show less refusal vis-à-vis immoral activities. Furthermore, persons under anger motivation realize less options for action within a situation. This results from the fact that anger-motivated persons can not consider new information differentiated, but process these through heuristical thinking in a fast and reactionist way.

Task

For all narratives in the folder, please indicate three assessments in numbers from 0 to 11:

1.) How compassion-motivated were behavior, perceptions and feelings of the narrating person in the described situation? (0-11)
2.) How anger-motivated were behavior, perceptions and feelings of the narrating person in the described situation? (0-11)

3.) How high was the intensity of the described situation for you / how much did the described situation touch you or change something in you? (0-11)

Z References

Abatayo, A. L., Lynham, J., & Sherstyuk, K. (2018). Facebook-to-Facebook: online communication and economic cooperation. *Applied Economics Letters*, *25*(11), 762-767.

Abele, A. E., Uchronski, M., Suitner, C., & Wojciszke, B. (2008). Towards an operationalization of the fundamental dimensions of agency and communion: Trait content ratings in five countries considering valence and frequency of word occurrence. *European Journal of Social Psychology*, *38*(7), 1202-1217.

Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in Economics* (pp. 235-251).

Akerlof, G. A., & Shiller, R. J. (2015). *Phishing for phools: The economics of manipulation and deception*. Princeton University Press.

Akerlof, G. A., & Snower, D. J. (2016). Bread and bullets. *Journal of Economic Behavior & Organization*, *126*, 58-71.

Alexander, B. (2017). *The New Digital Storytelling: Creating Narratives with New Media--Revised and Updated Edition*. ABC-CLIO.

Andreoni, J., & Rao, J. M. (2011). The power of asking: How communication affects selfishness, empathy, and altruism. *Journal of public economics*, *95*(7-8), 513-520.

Angie, A. D., Connelly, S., Waples, E. P., & Kligyte, V. (2011). The influence of discrete emotions on judgement and decision-making: A meta-analytic review. *Cognition & Emotion*, 25(8), 1393-1422.

Arango, L. E., Pantoja, J., & Velásquez, C. (2017). *Effects of the central bank’s communications in Colombia. A content analysis*(No. 1024). Banco de la Republica de Colombia.

Babin, J. J. (2016). A Picture is Worth a Thousand Words: Emojis, Computer-Mediated Communication, and Trust.

Bandiera, O., Hansen, S., Prat, A., and Sadun, R. (2017). CEO Behavior and Firm Performance. *NBER Working Papers 23248*, National Bureau of Economic Research, Inc.

Batson, C. D., O'Quin, K., Fultz, J., Vanderplas, M., & Isen, A. M. (1983). Influence of self-reported distress and empathy on egoistic versus altruistic motivation to help. *Journal of personality and social psychology*, *45*(3), 706.

Berger, A. A. (1997). *Narratives in popular culture, media, and everyday life*. Sage.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan), 993-1022.

Bluck, S., Alea, N., Habermas, T., & Rubin, D. C. (2005). A tale of three functions: The self–reported uses of autobiographical memory. *Social Cognition*, *23*(1), 91-117.

Bock, O., Baetge, I., & Nicklisch, A. (2014). hroot: Hamburg registration and organization online tool. *European Economic Review, 71*, 117-120.

Boudoukh, J., Feldman, R., Kogan, S., & Richardson, M. (2013). *Which news moves stock prices? a textual analysis* (No. w18725). National Bureau of Economic Research.

Brandts, J., & Cooper, D. J. (2007). It's what you say, not what you pay: An experimental study of manager-employee relationships in overcoming coordination failure. *Journal of the European Economic Association*, *5*(6), 1223-1268.

Budak, C., Goel, S., Rao, J. M., and Zervas, G. (2014). Do-not-track and the economics of third-party advertising. Boston University, School of Management Research Paper 2505643.

Carpenter, B. (2010). Integrating out multinomial parameters in latent Dirichlet allocation and naive Bayes for collapsed Gibbs sampling. *Rapport Technique*, *4*, 464.

Charness, G., & Dufwenberg, M. (2006). Promises and partnership. Econometrica, 74 (6),

1579–1601.

Charness, G., & Dufwenberg, M. (2009). Hidden information, participation, and communication. Unpublished Paper (University of California, Santa Barbara).

Chen, R., & Chen, Y. (2011). The potential of social identity for equilibrium selection. The

American Economic Review, 2562–2589.

Chierchia, G., Przyrembel, M., Lesemann, F. P., Bosworth, S., Snower, D., & Singer, T. (2018, April 4). Navigating Motivation: A Semantic and Subjective Atlas of 7 Motives. http://doi.org/10.17605/OSF.IO/WJHMQ

Cohen, G. (1998). The effects of aging on autobiographical memory. Autobiographical memory: Theoretical and applied perspectives, 105-123.

Cooper, D. J., & Kagel, J. H. (2005). Are two heads better than one? Team versus individual play in signaling games. American Economic Review, 95(3), 477-509.

Crocker, J., & Canevello, A. (2012). Consequences of self-image and compassionate goals. In Advances in experimental social psychology (Vol. 45, pp. 229-277). Academic Press.

Darling, W. M. (2011, December). A theoretical and practical implementation tutorial on topic modeling and gibbs sampling. In *Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies* (pp. 642-647).

Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, *346*(6210), 1243089.

Falk, A., & Tirole, J. (2016). *Narratives, imperatives and moral reasoning* (pp. 707-711). mimeo.

Finucane, A. M. (2011). The effect of fear and anger on selective attention. *Emotion*, *11*(4), 970-974.

Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics letters*, *71*(3), 397-404.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental economics, 10(2)*, 171-178.

Fraiberger, S. P. (2016): “News Sentiment and Cross-Country Fluctuations,” Available at SSRN.

Ganz, M. (2011). Public narrative, collective action, and power. *Accountability through public opinion: From inertia to public action*, 273-289.

Gentzkow, M., & Shapiro, J. M. (2010). What drives media slant? Evidence from US daily newspapers. *Econometrica*, *78*(1), 35-71.

Gentzkow, M., Shapiro, J. M., & Taddy, M. (2016). *Measuring polarization in high-dimensional data: Method and application to congressional speech* (No. w22423). National Bureau of Economic Research.

Gentzkow, M., Kelly, B. T., & Taddy, M. (2017). *Text as data* (No. w23276). National Bureau of Economic Research.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. (Eds.). (1995). *Markov chain Monte Carlo in practice*. CRC press.

Glaeser, E. L. (2005). The political economy of hatred. *The Quarterly Journal of Economics*, *120*(1), 45-86.

Goetz, J. L., Keltner, D., & Simon-Thomas, E. (2010). Compassion: an evolutionary analysis and empirical review. *Psychological bulletin*, *136*(3), 351.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, *101*(suppl 1), 5228-5235.

Gu, R., Huang, Y. X., & Luo, Y. J. (2010). Anxiety and feedback negativity. *Psychophysiology*, *47*(5), 961-967.

Habermas, T., & Bluck, S. (2000). Getting a life: the emergence of the life story in adolescence. *Psychological bulletin*, *126*(5), 748.

Hansen, S., McMahon, M., & Prat, A. (2017). Transparency and deliberation within the FOMC: a computational linguistics approach. *The Quarterly Journal of Economics*, *133*(2), 801-870.

Heinrich, G. (2008). *Parameter estimation for text analysis,"Univ. Leipzig, Leipzig*. Germany, Tech. Rep., http://faculty. cs. byu. edu/~ ringger/CS601R/papers/Heinrich-GibbsLDA. pdf.

Heston, S. L., and N. R. Sinha (2015): "News versus sentiment: Predicting stock returns from news stories," Robert H. Smith School Research Paper.

Hogenraad, R. (2003). The words that predict the outbreak of wars. *Empirical studies of the Arts*, *21*(1), 5-20.

Hogenraad, R. (2005). What the Words of War Can Tell Us About the Risk of War. *Peace and Conflict: Journal of Peace Psychology*, *11*(2), 137.

Holland, A. C., & Kensinger, E. A. (2010). Emotion and autobiographical memory. *Physics of life reviews*, *7*(1), 88-131.

Houser, D., & Xiao, E. (2011). Classification of natural language messages using a coordination game. *Experimental Economics*, *14*(1), 1-14.

Klimecki, O., & Singer, T. (2012). Empathic distress fatigue rather than compassion fatigue? Integrating findings from empathy research in psychology and social neuroscience. *Pathological altruism*, 368-383.

Kuziemko, I., & Washington, E. (2015). *Why did the democrats lose the south? Bringing new data to an old debate* (No. w21703). National Bureau of Economic Research.

Larsen, S. F. (1992). Personal context in autobiographical and narrative memories. In *Theoretical perspectives on autobiographical memory* (pp. 53-71). Springer Netherlands.

Lerner, J. S., & Tiedens, L. Z. (2006). Portrait of the angry decision maker: How appraisal tendencies shape anger's influence on cognition. *Journal of Behavioral Decision Making*, *19*(2), 115-137.

Loughran, T., and B. McDonald (2011): "When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks," The Journal of Finance, 66(1), 35–65.

Lowe, T., Brown, K., Dessai, S., de França Doria, M., Haynes, K., & Vincent, K. (2006). Does tomorrow ever come? Disaster narrative and public perceptions of climate change. *Public understanding of science*, *15*(4), 435-457.

McAdams, D. P. (2001). The psychology of life stories. *Review of general psychology*, *5*(2), 100.

Mokyr, J. (2016). *A culture of growth: the origins of the modern economy*. Princeton University Press.

Nimark, K. P. and Pitschner, S. (2016). Delegated information choice. CEPR Discussion Papers 11323, C.E.P.R. Discussion Papers.

Nye Jr, J. (2008). *The powers to lead*. Oxford University Press.

Nye Jr, J. S. (2013). Transcript of witness testimony to the House of Lords Select Committee on Soft Power and UK Influence, 15 October.

Nyman, R., Kapadia, S., Tuckett, D., Gregory, D., Ormerod, P., & Smith, R. (2018). News and narratives in financial systems: exploiting big data for systemic risk assessment.

O'Connor, P. E. (2000). *Speaking of crime: Narratives of prisoners* (Vol. 17). U of Nebraska Press.

Pancost, R. D. (2017). Climate change narratives. *Nature Geoscience*, *10*(7), 466.

Pasupathi, M., & Hoyt, T. (2009). The development of narrative identity in late adolescence and emergent adulthood: The continued importance of listeners. *Developmental Psychology*, *45*(2), 558.

Penczynski, S. P. (2016). Using machine learning for communication classification.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of personality and social psychology*, *77*(6), 1296.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. Linguistic Inquiry and Word Count (LIWC): LIWC2001, 2001.

Pillemer, D. (2003). Directive functions of autobiographical memory: The guiding power of the specific episode. *Memory*, *11*(2), 193-202.

Pöder, K. (2010). The lighthouse in Estonia: the provision mechanism of "Public Goods". *Discussions on Estonia Economic Policy*, *18*, 323-347.

Rodrik, D. (Ed.). (2003). *In search of prosperity: Analytic narratives on economic growth*. Princeton University Press.

Roselle, L., Miskimmon, A., & O'Loughlin, B. (2014). Strategic narrative: A new means to understand soft power. *Media, War & Conflict*, *7*(1), 70-84.

Roseman, I. J. (2011). Emotional behaviors, emotivational goals, emotion strategies: Multiple levels of organization integrate variable and consistent responses. *Emotion Review*, *3*(4), 434-443.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological review*, *96*(2), 341.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological review*, *69*(5), 379.

Schultheiss, O. C. (2013). Are implicit motives revealed in mere words? Testing the marker-word hypothesis with computer-based text analysis. *Frontiers in psychology*, *4*, 748.

Seidenstücker, G., & Seidenstücker, E. (1974). Versuche zur Computerauswertung des Leistungs-TAT von Heckhausen. *Psychol. Beiträge*, *16*, 68-92.

Shapiro, A. H., Sudhof, M., & Wilson, D. (2018, March). Measuring news sentiment. Federal Reserve Bank of San Francisco.

Shiller, R. J. (2017). Narrative economics. *The American Economic Review*, *107*(4), 967-1004.

Singer, T., & Klimecki, O. M. (2014). Empathy and compassion. *Current Biology*, *24*(18), R875-R878.

Smith, M. S. (1968). The computer and the TAT. *Journal of School Psychology*, *6*(3), 206-214.

Smith, C. A., & Lazarus, R. S. (1990). Emotion and adaptation. In: L. A. Pervin (Ed.), *Handbook of Personality: Theory and Research*, 609-637. New York: Guilford Press.

Smorti, A., & Fioretti, C. (2016). Why narrating changes memory: a contribution to an integrative model of memory and narrative processes. *Integrative Psychological and Behavioral Science*, *50*(2), 296-319.

Stellar, J. E., Cohen, A., Oveis, C., & Keltner, D. (2015). Affective and physiological responses to the suffering of others: Compassion and vagal activity. *Journal of Personality and Social Psychology*, *108*(4), 572.

Sutter, M., & Strassmair, C. (2009). Communication, cooperation and collusion in team tournaments—an experimental study. *Games and Economic Behavior*, *66*(1), 506-525.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, *62*(3), 1139-1168.

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, *63*(3), 1437-1467.

Thorsrud, L. A. (2016). Nowcasting using news topics. Big Data versus big bank.

Turowski, J., & Mikfeld, B. (2013). Gesellschaftlicher Wandel und politische Diskurse. *Überlegungen für eine strategieorientierte Diskursanalyse. Werkbericht*, *3*.

Van Lange, P. A., Rockenbach, B., & Yamagishi, T. (Eds.). (2017). *Trust in social dilemmas*. Oxford University Press.

Wolf, M., Horn, A. B., Mehl, M. R., Haug, S., Pennebaker, J. W., & Kordy, H. (2008). Computergestützte quantitative textanalyse: äquivalenz und robustheit der deutschen version des linguistic inquiry and word count. *Diagnostica*, *54*(2), 85-98.

Wydick, B. (2015). How narratives influence human behavior. Retrieved from https://www.weforum.org/agenda/2015/01/how-narratives-influence-human-behaviour/

Xiao, E., & Houser, D. (2005). Emotion expression in human punishment behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(20), 7398-7401.

Young, L., and S. Soroka (2012): "Affective news: The automated coding of sentiment in political texts," Political Communication, 29(2), 205–231.