information

Article

Detecting Moral Features in TV Series with a Transformer Architecture through Dictionary-Based Word Embedding

Paolo Fantozzi [1,†], Valentina Rotondi [2,3,†], Matteo Rizzolli [1,†], Paola Dalla Torre [4,†] and Maurizio Naldi [1,5,*,†]

[1] Department of Law, Economics, Politics, and Modern Languages, LUMSA University, 00192 Rome, Italy; p.fantozzi1@lumsa.it (P.F.); m.rizzolli@lumsa.it (M.R.)
[2] Department of Economics, Health and Social Care, SUPSI, 6928 Lugano, Switzerland; valentina.rotondi@supsi.ch
[3] Department of Sociology, University of Oxford, Oxford OX1 1JD, UK
[4] Department of Social Sciences—Communication, Education and Psychology, LUMSA University, 00193 Rome, Italy; p.dallatorre1@lumsa.it
[5] Department of Civil, Computer Science and Aeronautical Technologies Engineering, Roma Tre University, 00146 Rome, Italy
* Correspondence: m.naldi@lumsa.it
† These authors have contributed equally to this work.

Abstract: Moral features are essential components of TV series, helping the audience to engage with the story, exploring themes beyond sheer entertainment, reflecting current social issues, and leaving a long-lasting impact on the viewers. Their presence shows through the language employed in the plot description. Their detection helps regarding understanding the series writers' underlying message. In this paper, we propose an approach to detect moral features in TV series. We rely on the Moral Foundations Theory (MFT) framework to classify moral features and use the associated MFT dictionary to identify the words expressing those features. Our approach combines that dictionary with word embedding and similarity analysis through a deep learning SBERT (Sentence-Bidirectional Encoder Representations from Transformers) architecture to quantify the comparative prominence of moral features. We validate the approach by applying it to the definition of the MFT moral feature labels as appearing in general authoritative dictionaries. We apply our technique to the summaries of a selection of TV series representative of several genres and relate the results to the actual content of each series, showing the consistency of results.

Keywords: moral features; moral foundations theory; TV series; storytelling; language analysis; genre analysis; deep learning; transformers; SBERT

check for updates

Citation: Fantozzi, P.; Rotondi, V.; Rizzolli, M.; Dalla Torre, P.; Naldi, M. Detecting Moral Features in TV Series with a Transformer Architecture through Dictionary-Based Word Embedding. *Information* **2024**, *15*, 128. https://doi.org/10.3390/info15030128

Academic Editor: Katsuhide Fujita

Received: 18 January 2024
Revised: 15 February 2024
Accepted: 22 February 2024
Published: 24 February 2024

1. Introduction

The description and transmission of moral features through language have a long-standing history [1]. Moral content has been intrinsic to narration since ancient times when religious texts and myths aimed at leading the audience to draw some moral conclusions [2]. Actually, morality has been viewed as an unavoidable companion to any narration [3]. Also, moral prescriptions on behavior and expression were common in many manuals from the 16th to the 18th century [4]. The conveyance of morality through language has occurred in more recent forms of narration, such as movies [5,6] and TV series [7,8]. The relevance of TV series in the transmission of moral features is especially pertinent due to their role in shaping popular culture and influencing viewers' perceptions and attitudes [9]. Even when blockbusters are devised for audience success, moral patterns are somewhat included, either in the form of adherence to the audience's moral beliefs or as moral norm violations introduced in the conflicts among characters [10].

An established classification of moral features has been set through Moral Foundations Theory [11] (MFT), which posits five universal moral foundations— Harm/Care,

Fairness/Reciprocity, Ingroup/Loyalty, Authority/Respect, and Purity/Sanctity. This theoretical framework has been largely used to analyze the presence of moral content in traditional media (books and songs) as well as relatively recent ones (movies and video games) [12–14]. However, no attempts have been made in the literature to employ it in the context of TV series. The growing relevance of this entertainment outlet for the formation of popular culture poses a spurring need for its analysis.

Our paper aims to fill this research gap by adopting the MFT framework to analyze the presence of moral content in TV series. We propose a deep learning-based methodology using the Sentence-Bidirectional Encoder Representations from Transformers (SBERT) architecture, based on the combined use of a lexicon established as a companion to MFT (the *Moral Foundations Dictionary*, MFD) and a word embedding approach. We strive to provide a wide panorama of TV series by analyzing series encompassing several genres.

Our major contributions are the following:

- we propose a deep learning architecture to identify the presence of moral foundations in TV series;
- we validate the method through an authoritative general dictionary-based reference;
- we apply the novel architecture to a selection of seven TV series belonging to different genres, showing how the quantitative analysis of moral foundations relates to the content of those series.

The structure of the paper is as follows. We review the literature related to the quantitative analysis of TV series and the MFT framework in Section 2. We describe our datasets and methodology in Section 3 and finally provide the method validation and the results in Section 4.

2. Related Literature

Our study consists of applying quantitative techniques to TV series by employing the Moral Foundations Theory framework. In this section, we provide a brief literature review of quantitative studies concerning TV series and a wider description of the MFT framework and its applications.

2.1. The Quantitative Analysis of TV Series

TV series have started being analyzed quantitatively in the last decade using tools from graph theory and text mining.

The analysis first focused on the relationships between characters, embodied by the character network. Beveridge and Shan have shown how graph theory could be employed to examine the interactions among characters and the presence of clusters of characters in *Game of Thrones* [15]. Bost et al. extracted dynamic conversational networks from *Breaking Bad*, *Game of Thrones*, and *House of Cards* [16]. Tan et al. measured the graph characteristics of co-occurrence character networks in *Stargate* and *Star Trek* [17]. Complex networks were introduced to describe roles in *Mr. Love*, *Westworld*, and *The Dream of Red Mansions* [18]. The identification of leading roles among characters to distinguish between choral series, on the one hand, and series focusing on the main character, on the other hand, was tackled in [19] through the use of suitable metrics. Also, the influence of the network structure on the series' performance was examined by [20].

A more recent theme of investigation has been the analysis of the narrative structure of series. A classification of plots of medical series into medical cases plots, professional plots, and sentimental plots was proposed in [21]. Instead, the evolution of series over time was analyzed in [22] by examining the changing salience of characters and locations.

2.2. Moral Foundations Theory

Moral Foundations Theory, employed in this paper, has gained recognition as a psychological framework that uncovers the fundamental moral principles influencing human judgments and behavior. Although relatively recent, MFT has quickly gained acceptance in the field. The following section delves into this theory's historical development and formu-

lation. Next, we explore its application to identifying moral foundations within traditional media formats like books, songs, television, cinema, and emerging media platforms like video games. We analyze the media under scrutiny and the methodologies employed to ascertain the prominence of moral foundations, offering a comprehensive examination of their presence and impact.

Graham et al. introduced a comprehensive framework for classifying moral values, which they termed *Moral Foundations Theory* [11,23]. MFT builds on earlier theories of moral judgment and offers a unified framework to account for variations in human moral reasoning as the result of the combination of a few innate and modular foundations.

While alternative approaches exist (see, for instance, [24]), and although MFT is not impervious to criticism [25], its rapid ascent in popularity is unmistakable, evident in the proliferation of citations and the broad adoption of its questionnaire and associated instruments across various disciplines, including moral psychology, political science [26], economics [27], law [28], behavioral science [29], and beyond.

Within MFT, Graham et al. identified five fundamental dimensions, which they referred to as *foundations*. In their original formulation, the five dimensions were identified by the following pairs:

- *Harm/Care*;
- *Fairness/Reciprocity*;
- *Ingroup/Loyalty*;
- *Authority/Respect*;
- *Purity/Sanctity*.

According to Moral Foundations Theory, these five foundations are universal across cultures and reflect evolved psychological mechanisms that enable individuals and societies to solve collective action problems and navigate social complexity. However, different cultures and individuals may emphasize or prioritize these foundations differently, leading to different moral beliefs and values.

The *Harm/Care* foundation revolves around the notion that individuals and societies should protect and care for those who are vulnerable and in need. It encompasses values such as compassion, empathy, and kindness.

The *Fairness/Reciprocity* foundation centers on the idea that individuals and societies should strive for fairness, reciprocity, and justice. It encompasses values such as equality, proportionality, and impartiality.

The *Ingroup/Loyalty* foundation emphasizes the importance of loyalty and commitment to one's social groups, including family, tribe, or nation. It encompasses values such as patriotism, group solidarity, and sacrifice for the common good.

The *Authority/Respect* foundation underscores the significance of respecting and deferring to legitimate authority and hierarchical structures. It encompasses values such as obedience, respect for traditions, and deference to leaders.

The *Purity/Sanctity* foundation highlights the value of upholding purity, cleanliness, and sacredness within various domains, such as sexuality, food, and religion. It encompasses values such as reverence for the sacred, disgust for the impure, and adherence to traditional moral codes.

As you can see, [11] used a couple of terms to define each foundation. The terms in each couple were akin in most cases, but, in the case of *Harm/Care*, they seemed to identify opposite extremes of the same moral value. Similar couples of opposites have been identified in [23] for the other foundations as well, such as *Fairness/Cheating*, *Loyalty/Betrayal*, *Authority/Subversion*, and *Sanctity/Degradation*. In the following, we opt for the more recent formulation based on a couple of extremes for each moral foundation, as proposed at https://moralfoundations.org (accessed on 18 January 2024).

In order to identify the presence of those moral foundations in a corpus of texts, [30] built a lexicon for each foundation. The lexicon used in this study was developed through an iterative process comprising both an expansive and contractive phase. During the expansive phase, the researchers began with the core concepts of each moral foundation and

systematically generated a comprehensive list of associations, synonyms, and antonyms. This process involved consulting thesauruses, engaging in discussions with colleagues, and exploring various linguistic resources.

Subsequently, in the contractive phase, the lexicon underwent refinement. Words that appeared distant from the fundamental concepts of the moral foundations were purged from the lexicon. This selective process ensured that the final lexicon captured the essence of each foundation, encompassing both positive and negative polarities associated with them. By considering the moral foundations as cohesive entities, inclusive of their positive and negative dimensions, the lexicon provides a comprehensive and nuanced representation of the moral values under examination. For example, the lexicon for *Purity/Sanctity* includes *profane* and *pervert* as well as *pious* and *purity*, among others. The resulting lexicon was dubbed the *Moral Foundations Dictionary*.

In order to measure the presence of each foundation, [31] developed a questionnaire (the Moral Foundations Questionnaire, or MFQ for short) that included 30 items.

In addition to the comprehensive lexicon encompassing words associated with the five moral foundations, the researchers identified specific lexicons for the two opposite facets of each foundation. This recognition stems from the understanding that each foundation can be expressed in terms of both positive and negative dimensions, representing different sides of the moral spectrum. The researchers defined those two facets of the same foundation as *virtue* vs. *vice*. For a more detailed exploration of the lexicons associated with the virtue and vice sides of each foundation, interested readers can refer to the webpage available at http://moralfoundations.org/wp-content/uploads/files/downloads/moral%20foundations%20dictionary.dic (accessed on 17 January 2024). On this webpage, the lexicons associated with each facet of the moral foundations can be accessed, providing a comprehensive resource for understanding the specific words and concepts associated with the positive and negative dimensions of each foundation.

2.3. Moral Foundations Theory in Media Analysis

Understanding the moral dimensions portrayed in various forms of media is vital for comprehending their impact on individuals' attitudes, behaviors, and societal norms. The application of Moral Foundations Theory provides a systematic framework for delving into these moral underpinnings and examining their influence within media content. This subsection reviews the existing literature on the exploration of MFT in media analysis, highlighting its significance and contributions to the field. We consider first books, moving then to songs and movies, to end with video games.

Wheeler et al. employed the Google Books Ngram viewer to analyze books published in the 1900–2007 period [32]. They adopted the Moral Foundations Theory framework to examine how moral values shifted during the century. The Google Books Ngram viewer is a search engine that outputs the frequencies of search strings using n-grams in printed sources published between 1500 and 2019 in Google's text corpora in several languages. The total collection is far from complete and is claimed to contain more than 6% of all the books ever published [33]. The authors employed the terms comprising the *Moral Foundations Dictionary* (see Section 2.2 for a description of that lexicon) as search strings and found the resulting term frequencies for all the terms associated with the six moral foundations and derived the frequency of each moral foundation, restricting their analysis to books in English.

Similarly, Long and Eveland conducted a lexicon-based analysis (again employing the MFT framework described in Section 2.2) on song lyrics, encompassing 13 genres, to explore the variations in moral content across these genres [34]. The songs were selected by referencing the primary US chart categories curated by Billboard. While these charts may feature songs in languages other than English, we speculate that the authors exclusively incorporated English-language song lyrics. Additionally, the authors administered a survey employing the Moral Foundations Questionnaire to assess the moral attitude of the respondents and record their genre preferences. By comparing the data derived from the analysis

of song lyrics with the questionnaire responses, the study aimed to examine whether the participants' moral progressivism aligned with the moral progressivism reflected in the lyrics of their preferred music genres.

While the available literature on the application of Moral Foundations Theory to the aforementioned media forms is limited, there exists a more substantial body of research focused on applying MFT in the analysis of TV and cinema.

Bowman et al. examined the mediating role of moral foundations in individuals' preferences for media genres, specifically focusing on movies (drama, action, and horror) and TV shows (comedy, news, and sports) [35]. They found that moral foundations play a mediating role in the relationship between nationality and genre preferences. In order to gauge the salience of moral foundations, the study employed the MFQ, administered to individuals from two different countries, namely the United States and Germany. The questionnaire for US respondents was the MFQ itself, while an adapted 29-item scale in German developed using back-translation was submitted to German respondents. Similarly, ref. [12] considered the influence of moral foundations on the consumption of TV crime dramas. In their study, the researchers also employed the MFQ.

In a study by Tamborini et al. [14], rather than using actual movies, ten summaries of fictional movies were created to examine whether viewers' moral beliefs influenced characters' appeal. Participants' moral beliefs were assessed using the standard MFQ, while viewers were asked to express their enjoyment of the movie outcome for the specific character (for example, if they enjoyed a negative outcome for a character violating a moral domain).

A similar investigation was carried out in [36] for the viewers of the series *Downton Abbey*. In the paper, the researchers examined Twitter posts to investigate how the behaviors of characters influenced viewers' perceptions in relation to moral norms. The salience of moral foundations was assessed by searching for the terms comprising the *Moral Foundations Dictionary* in the tweets.

Gehman et al. adopted a focus on children's movies, specifically examining three movies, with the aim of identifying the differences in moral foundations between heroes and villains [37]. The presence of moral foundations was investigated by having a group of human coders fill out the Moral Foundations Questionnaire for each character under analysis.

In another study conducted in [38], the focus was once again on children. The researchers explored the presence of moral foundations in the narratives contained within thematic textbooks used in Indonesian primary schools. The assessment of moral foundations involved tracing the occurrence of lexical tags associated with each moral foundation within the textbooks. In contrast to the aforementioned literature, which primarily relied on the direct search of *Moral Foundations Dictionary* (MFD) terms in texts, Araque et al. introduced the use of word embeddings to search for the presence of moral values by adopting a Wordnet-based extension of the MFD and using the SIMilarity-based sentiment projectiON (SIMON) [39] to compute the cosine similarity as a measure of semantic similarity between the words in the text and the words in the extended lexicon [40]. They applied their method to a corpus of tweets using a classification approach, where each tweet was assigned a single moral trait. Overcoming the binary classification task adopted in [39], Gonzalez et al. extended the number of moral foundations that can be assigned to a text using a 20% threshold, i.e., assigning 20% of the potential number of foundations to each movie [41]. They employed two embedding techniques, namely Word2Vec and BERT, to analyze tags extracted from synopses of a selection of movies (the authors downloaded the synopses from three popular online sources: IMDb, Rotten Tomatoes, and FilmAffinity, although it is not clear if they used all three sources). Tag extraction was performed through the YAKE! tool designed by Campos et al. [42]. Gonzalez et al. employed both the original *Moral Foundations Dictionary* (composed of ten items, including the opposite extremes of the five moral foundations) and an extended version, comprising twenty-four foundations. The analysis was conducted by computing the similarity between the embedding vectors of

the MFD terms and those of the words in the synopses. A threshold was applied to extract the most relevant foundations for each movie.

In addition to exploring moral dimensions in traditional media, it is worth mentioning a smaller but emerging area of research that examines the presence of moral dimensions within video games. However, unlike traditional media, the focus of this research is not solely on the moral content embedded in video games but rather on the moral behavior and decisionmaking processes evoked during gameplay.

The relationship between moral decisions and facial thermal variations was investigated by Guglielmo et al. [43], where specific moral foundations were assigned to decisions within the narrative of the video game *The Walking Dead*. These assignments were likely performed manually. In another study, Hodge et al. utilized a purpose-made video game to examine the time taken for moral decisions and the alignment between players and the game in terms of moral values [44].

Turning to the Chinese version of the popular video game *World of Warcraft*, Hornbeck et al. identified the presence of moral foundations elicited during gameplay through a survey [45]. Participants were asked to rate the frequency of experiencing feelings associated with moral foundations while playing their main *World of Warcraft* character.

Although recognizing that players may use their moral intuition when playing games, Krcmar et al. considered both moral decisions and strategic decisions (driven just by the desire to play and win the game) and traced the correspondence between the moral beliefs of the player and his/her behavior in the game [46]. The salience of the moral foundations was measured through a questionnaire. Similarly, Joeckel et al. analyzed the influence of moral intuition when making decisions in video games [13]. Again, the MFQ questionnaire was employed to measure moral foundations.

To summarize, previous studies have measured the salience of moral foundations through questionnaires or by quantifying the frequency of *Moral Foundations Dictionary* (MFD) terms in the analyzed text, except for the work of González-Santos et al. [41], which employed a deep learning-based approach. Additionally, the literature has primarily focused on media such as books, songs, movies, and video games, with only one exception in the TV series domain, as observed in the work of Ji and Raney [36], which analyzed viewers' perceptions rather than the actual moral content of the series.

In contrast to the existing literature, our approach differs in two key aspects: the dataset of interest and the methodology employed. We specifically focus on TV series, which have received limited attention in the current literature. Although Gonzalez et al. analyzed movies, they considered a selection of 20 movies in contrast to our massive scale of over 600 TV series episodes [41]. Moreover, we employ a deep learning-based approach to quantify the salience of moral foundations. While most previous studies relied on the use of the Moral Foundations Questionnaire and involved human coders, our approach is entirely automated, utilizing the SBERT architecture. This distinguishes our work from previous research, including the work by González et al. [41], who employed different embedding techniques (Word2Vec/BERT instead of SBERT). Further differences with the method employed by [41] are that we use the full text of the synopses rather than keywords extracted from them, which allows us not to ignore text that has not been distilled into keywords. They also sum embedding vectors rather than averaging them as we do. Averaging (our choice) rather than summing allows us to compare different movies (or different TV series) since summing outputs values that are influenced by the number of tags of each movie (more tags means more summing terms), making a comparison unfeasible. Finally, we provide the full range of salience values for moral foundations rather than just outputting one moral foundation [40] or two [41]. This allows us to obtain a more complete view of the moral foundations represented in the series.

3. Data and Methods

3.1. Data

We aim to employ our method to trace moral foundations in TV series, focusing on a selection of popular series. In this section, we provide a detailed description of our dataset, outlining the process of series selection, as well as the content and sources used.

To ensure a comprehensive representation of TV series, we adopted a top-down approach for our selection task. We began by considering a classification of TV series into various genres. However, it is worth noting that the classification of genres in TV series is subjective and lacks consensus. As a reference, we consulted Wikipedia, which lists a total of 95 genres on its webpage dedicated to TV genres (https://en.wikipedia.org/wiki/Category:Television_series_by_genre (accessed on 17 January 2024). Instead, IMDb lists the following 20 genres (https://www.imdb.com/list/ls023983860/ (accessed on 17 January 2024)):

- Drama;
- Comedy;
- Thriller;
- Crime;
- Mystery;
- Romance;
- Action;
- Sci-Fi;
- Fantasy;
- History;
- Adventure;
- Animation;
- Horror;
- Biography;
- War;
- Documentary;
- Music;
- Family;
- Western;
- Sport.

Finally, Creeber adopts the taxonomy of TV genres shown in Figure 1 [47].

Acknowledging the difficulty of establishing a universally accepted classification system, we have decided to rely on Creeber's classification and adopt a selection of the genres appearing in it. However, some (Crime, Fantasy, and Sci-Fi) are explicitly considered in IMDb too, and the others are subgenres of those employed by IMDb (Teen drama, Legal drama, and Hospital drama are subgenres of Drama, while Sitcom is a subgenre of Comedy). Our selection process involved choosing a popular series that exemplifies each genre. The following genres and their corresponding series have been selected:

- Crime (*Gomorrah*);
- Teen drama (*13 Reasons Why*);
- Fantasy (*Game of Thrones*);
- Hospital drama (*House, MD*);
- Sitcom (*Sex and the City*);
- Legal drama (*Suits*);
- Sci-Fi (*Westworld*).

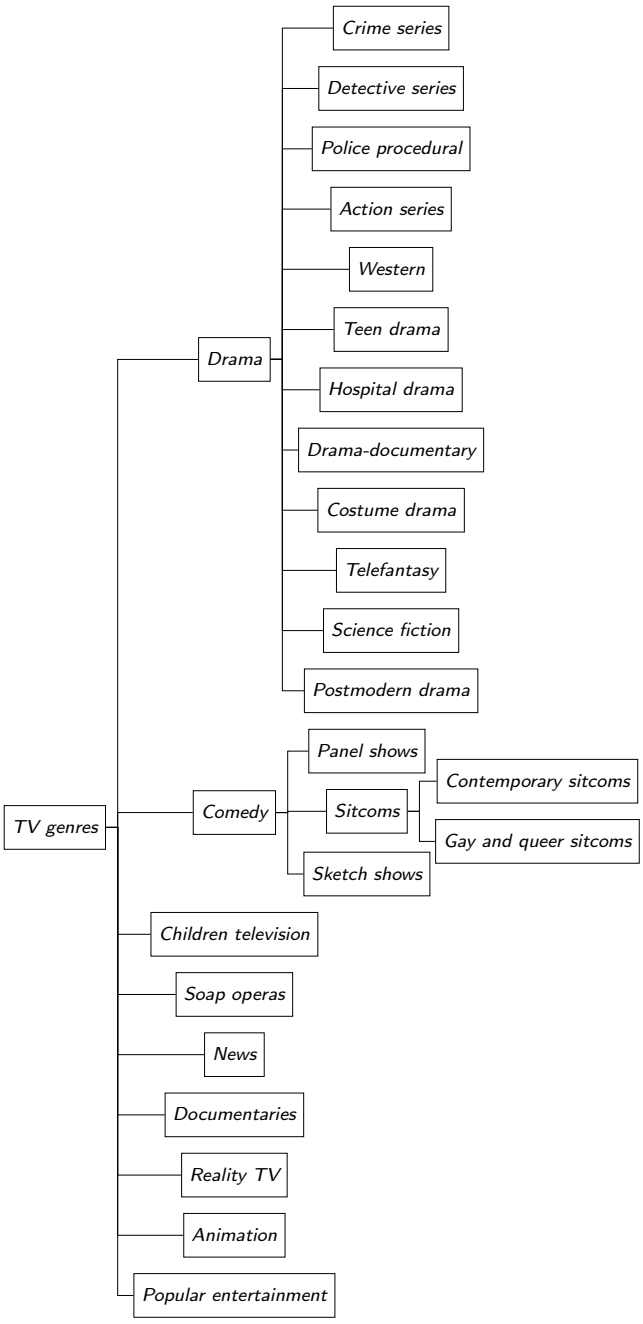

Figure 1. Taxonomy of TV genres by Creeber.

For each series in our dataset, we conducted our analysis on the summaries of the plots of each episode as reported on Wikipedia. We have chosen to analyze summaries because they exhibit several advantages over dialogues in describing the actual content of a series. They represent actions, which are not reflected in dialogues and would be lost if we considered dialogues instead of summaries. They may also represent states of mind that are not conveyed by dialogue but rather the appearance or gestures of a character. Also, they can effectively capture the structure and flow of the movie by highlighting the key plot points, major twists, and character developments in a coherent manner. Tags derived from summaries were employed in [41]. Our approach employing the full text of summaries avoids the potential misses introduced by the mediation layer represented by summary tagging.

We show the major statistics for all the series in Table 1, including the minimum, average, and maximum length of summaries, computed both in words and characters.

Word tokenization has been carried out through the `en_core_web_lg` model in the spacy library. The language for all episodes is English.

Table 1. Length statistics of series expressed in words (chars).

Series	No. of Episodes	Min Length	Avg Length	Max Length
Gomorrah	58	92 (463)	329.9 (1669.6)	569 (2901)
13 Reasons Why	49	127 (628)	225.5 (1132.6)	487 (2366)
Game of Thrones	73	116 (571)	211.6 (1098.6)	284 (1424)
House, MD	176	34 (183)	92.5 (479.4)	345 (1739)
Sex and the City	94	37 (186)	70.3 (355.4)	221 (1047)
Suits	134	24 (120)	156.4 (794.3)	483 (2416)
Westworld	36	148 (774)	304.3 (1069.6)	288 (1455)

3.2. Methods

In order to identify the presence of moral concepts in the texts of interest (the summaries of TV series episodes), we adopt an approach based on identifying semantic similarities through a deep learning architecture. In this section, we describe the details of our approach.

Embedding Similarity through SBERT

The basic concept is to represent each text as a vector in a dense space of fixed dimension in a way such that semantically similar texts are represented by close vectors. Semantic similarity can then be measured through any vector distance metric. In this paper, we have opted to use cosine similarity as a distance metric. The representation of text is of the embedding type, which allows for achieving space density as opposed to intrinsically sparse representations like one-hot encoding, as described by [48]. We obtain vector embeddings both for the episode summaries and the words in the *Moral Foundations Dictionary* for each moral dimension reported in [30]. The average cosine similarity between the episode summary embedding and the words appearing in a specific moral dimension lexicon is considered a measure of the presence of that moral dimension in that episode. Such an approach has already been validated in [39], where cosine similarity between embeddings of words in a text and moral foundations words as appearing in a lexicon has been adopted as a measure of the presence of moral foundations and shown to agree with the results of human annotation. Aside from the specific application of moral foundations identification, the application of cosine similarity to word embeddings to compute semantic similarity was validated, e.g., in [49] for newspaper articles. In quite general terms, the property of word embedding to represent similar words by similar vectors is stated in Chapter 24.1 of the reference textbook by Russell and Norvig [50] such that a measure of distance like cosine similarity will allow us to evaluate the semantic distance between words expressed through embeddings. More precisely, if we examine n episodes in the TV series and make use of the *Moral Foundations Dictionary*, which adopts five moral dimensions (without loss of generality, we associate an integer number $j \in \{1,2,3,4,5\}$ to each moral dimension) and lists w_j words for the j-th moral dimension, the measure of the presence of the j-th moral dimension in the i-th episode is

$$p_{ij} = \frac{1}{w_j} \sum_{k=1}^{w_j} S_C(e_i, m_{jk}), \tag{1}$$

where e_i is the embedding vector for the i-th episode summary, and m_{jk} is the embedding vector for the k-th word ($k = 1,2,\ldots,w_j$) in the j-th moral dimension lexicon ($j \in \{1,2,3,4,5\}$).

We can perform a similar analysis by separately examining the lexicon pertaining to the positive and negative facets of each moral dimension, as described in Section 2.2. If we represent the pertaining presence as p_{ij}^{+} and p_{ij}^{-}, respectively, for the j-th moral dimension in

the i-th episode, we can obtain a measure of the prevalence of the positive over the negative polarity by computing the difference $q_{ij} = p_{ij}^{+} - p_{ij}^{-}$. If the difference is positive, the positive polarity (virtue) will be considered the dominant facet of that foundation; vice versa, the negative polarity (vice) will be considered the dominant one. Since we obtain a polarity value for each episode, we will plot the estimated probability density of polarity. For the empirical probability density function, we have employed the density function in R, which employs the Gaussian kernel method described by [51]. More precisely, the algorithm used in that function disperses the mass of the empirical distribution function over a regular grid of at least 512 points and then uses fast Fourier transform to convolve this approximation with a discretized version of the kernel and then uses a linear approximation to evaluate the density at the specified points. The bandwidth employed in the kernel method is determined as in [52].

The deep learning architecture we employ to obtain the embedding vectors is Sentence-BERT (SBERT), proposed by [53], which improves on the popular BERT transformer architecture proposed by [54] by drastically reducing the computational effort required for large-scale semantic similarity tasks. The improvement of SBERT over BERT is achieved by adding a pooling layer at the output of BERT that computes the mean of all the output vectors and by using a Siamese training approach over pairs of sentences, with a distance function over representations to be as close as possible. The differences with BERT can be seen in Figure 2, taken from [53], where the SBERT architecture is shown for training (left) and inference (right), and the use of two BERT networks in parallel and the addition of a pooling layer are visible.

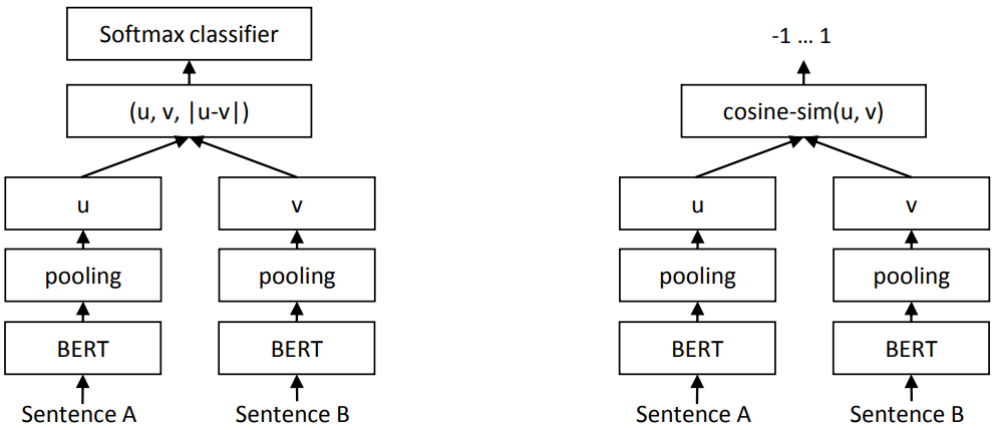

Figure 2. The SBERT architecture.

The model is a pre-trained, one named `paraphrase-multilingual-MiniLM-L12-v2` [53] (https://huggingface.co/sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2) (accessed on 17 January 2024), which is fine-tuned using a Siamese network structure: couples of text are input to the Siamese networks to minimize their difference.

4. Results

Our main objective is to examine the moral content of TV series along the five foundations proposed by [11]. Of course, the complex plots that characterize modern TV series are all likely to harness multiple foundations of morality. In addition to examining the moral content of a series on its own, our study also embraces a comparative perspective: different series/genres are compared to each other both regarding how salient each foundation is for a particular series/genre and regarding which side of the virtue–vice spectrum dominates for each foundation.

In this section, we show the results of our analysis. In order to visualize the results, we resort to a radar chart, where each radius represents one of the five moral dimensions described in Section 2.2, and the resulting pentagon cuts each radius at a value representing

the relative importance of the corresponding moral dimension. For the results to be easily comparable, we have normalized the presence values so that they sum to 1 for each episode; i.e., the following condition holds:

$$\sum_{j=1}^{5} p_{ij} = 1 \qquad \forall i \tag{2}$$

We will begin by presenting a comprehensive validation of our findings and then conducting an individual analysis of the selected series, followed by a comparative assessment.

4.1. Validation

In order to ensure the reliability and accuracy of our algorithm, it is essential to validate its performance. Recognizing the presence of moral foundations in TV series is a complex classification task as multiple foundations can coexist simultaneously with varying degrees of prominence. Moreover, we cannot access a definitive *ground truth* that would indicate which specific foundation is present in each episode.

Given these challenges, we have developed a validation approach that relies on the semantic associations of words related to moral foundations.

To ensure a robust validation process, we have obtained the definitions and sample sentences of the five moral foundation names and their polarity subdivisions from the widely respected *Oxford English Dictionary* [55]. By leveraging these authoritative linguistic sources, we aim to establish a reliable benchmark for evaluating our algorithm's performance.

In the validation phase, instead of employing the episode summaries, we fed our algorithm with those extracted definitions and sample sentences. This approach allows us to assess whether the algorithm can accurately identify the moral foundation corresponding to the provided definition. The algorithm generates a similarity value as its output, reflecting the degree of similarity between the input definition and each moral foundation.

To evaluate the algorithm's recognition accuracy, we have ranked the moral foundations based on their respective similarity scores. If the algorithm correctly identifies the moral foundation, it will rank first, indicating a higher similarity score compared to the other foundations. This ranking-based approach provides an explicit criterion for assessing the algorithm's performance.

Tables 2 and 3 present the results in a matrix format for foundations and foundation polarities, respectively. In these tables, the rows represent the actual foundations, while the columns indicate the output provided by the algorithm. The numbers inside the table indicate the rankings obtained. Ideally, if the algorithm were capable of perfectly recognizing the foundation, we would observe a diagonal consisting of 1s. We observe that this holds true for four out of five foundations, with the correct foundation ranking second in the missed case. Furthermore, for the polarity task, we achieve this perfect recognition in seven out of ten cases, with the correct polarity ranking second in the missed cases. It is important to note that this cannot be treated solely as a classification task since there is no perfect mutual exclusivity of foundations, as clarified previously. Nevertheless, if we were to consider it as such, the accuracy achieved would be 80% for foundations and 70% for foundation polarities.

Table 2. Validation ranking for moral foundations.

	Authority/Subversion	**Care/Harm**	**Fairness/Cheating**	**Loyalty/Betrayal**	**Sanctity/Degradation**
Authority/Subversion	**2**	5	1	3	4
Care/Harm	3	**1**	2	5	4
Fairness/Cheating	2	5	**1**	4	3
Loyalty/Betrayal	2	4	3	**1**	5
Sanctity/Degradation	4	3	2	5	**1**

Table 3. Validation ranking for moral foundations polarity.

	Authority	**Betrayal**	***Care***	**Cheating**	**Degradation**	**Fairness**	**Harm**	**Loyalty**	**Sanctity**	**Subversion**
Authority	**1**	5	7	3	8	2	10	6	9	4
Betrayal	8	**2**	7	5	4	9	3	6	10	1
Care	2	4	**1**	5	6	7	10	3	9	8
Cheat	7	2	10	**1**	5	4	6	9	8	3
Degradation	10	5	6	3	**1**	7	2	9	8	4
Fairness	3	8	5	2	7	**1**	10	6	4	9
Harm	9	7	5	1	6	3	**2**	10	8	4
Loyalty	5	7	6	2	10	3	9	**1**	8	4
Sanctity	4	9	2	6	7	3	10	5	**1**	8
Subversion	6	4	10	1	5	3	7	9	8	**2**

4.2. Measuring Moral Dimensions in TV Series

After assessing the capability of our approach to identify the presence of moral features, in this section, we report the results we have obtained by applying our method to the dataset described in Section 3.1. We devote a subsection to each series and the last subsection to an overall comparison of our findings.

4.2.1. *Gomorrah*

Gomorrah is an Italian TV series based on Roberto Saviano's book [56]. It focuses on the Camorra, a violent organized crime syndicate in Naples. The show is set in gritty neighborhoods and depicts the Camorra's operations and power struggles, affecting both criminals and ordinary people. The series does not shy away from portraying the harsh realities of the criminal underworld, showcasing violence, intimidation, and betrayal.

We consider first the radar chart where all the episodes are shown in Figure 3.

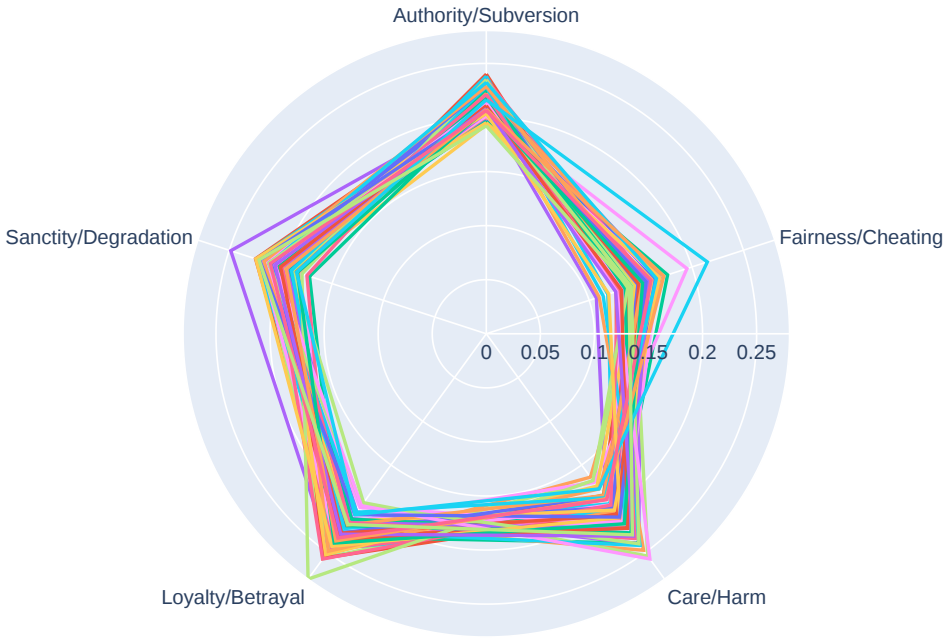

Figure 3. Moral foundation scores for all episodes in *Gomorrah*.

As observed, the pentagons representing the moral dimensions for all the episodes exhibit a significant degree of overlap, indicating remarkable consistency in the presence of moral dimensions across the episodes.

To evaluate the prominence of each moral dimension within the TV series, we employ a ranking based on the presence score, denoted as p_{ij}. By calculating the value of p_{ij} for each moral dimension in each episode, we can generate rankings that consistently highlight the most prominent moral dimension in the drama. Of course, the picture is

not extremely sharp: there is no moral dimension ranking first in all the episodes. We can obtain an overall picture of the presence of each moral dimension by computing the number of times it ranks first, second, etc. Table 4 presents the results of this analysis, showcasing the presence rankings of each moral dimension across the episodes. We see that the *Ingroup* moral dimension is indisputably the most present one, ranking first in 31 out of 58 episodes. At the same time, *Harm* is the sure runner-up, ranking first in sixteen out of the fifty-eight episodes, whereas *Authority* ranks first in six episodes and second in twenty-six. On the other end of the spectrum, we have *Fairness*, which holds the wooden spoon of this competition, being the least present moral dimension in fifty-six episodes, i.e., in all but two.

Table 4. Count of moral foundation rankings over all episodes in *Gomorrah*.

Ranking	**1**	**2**	**3**	**4**	**5**
Authority	6	26	18	8	0
Fairness	0	1	0	1	56
Harm	16	10	12	18	2
Ingroup	31	13	10	4	0
Purity	5	8	18	27	0

We can now examine the actual polarity of each foundation. In Figure 4, we see the resulting empirical density of q_{ij} for each moral foundation. We see that the negative polarity prevails for all the foundations, in particular for the *Ingroup* foundation, whose mode is largely in the negative semi-axis.

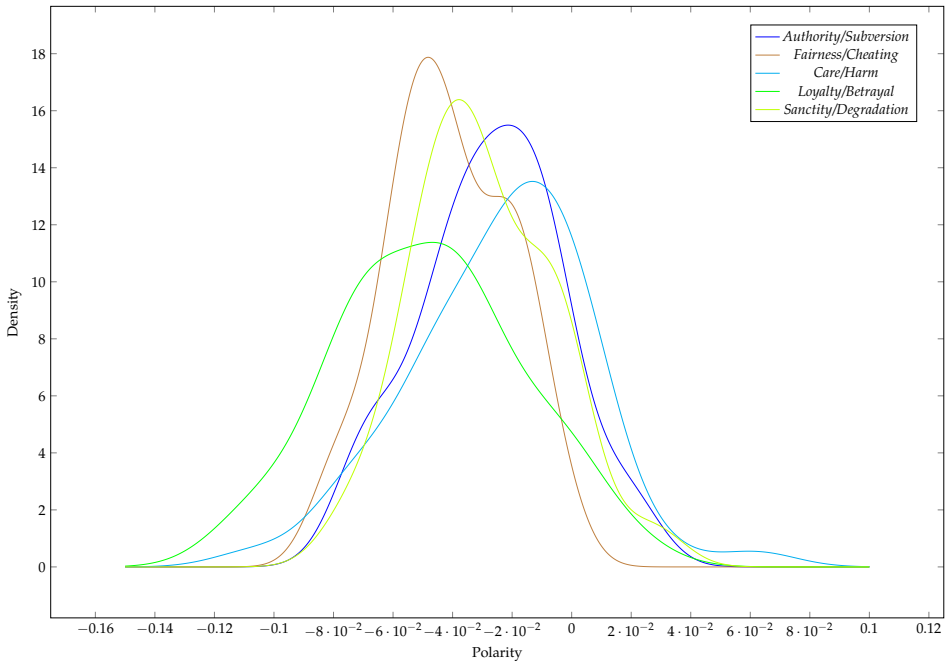

Figure 4. Empirical probability density of foundation polarity in *Gomorrah*.

The findings regarding *Gomorrah*'s dominant moral foundations are consistent with our expectations. The presence of the *Care/Harm* foundation aligns with the portrayal of the characters in the series. While some characters in the series may show signs of caring for their families or loved ones, a significant portion of the show revolves around violence and harm. The instances of characters inflicting pain, hurting, or even killing others without regard for their well-being demonstrate the prevalence of harm, indicating a negative polarity within this foundation. Regarding the *Ingroup/Loyalty* foundation, the series accurately reflects the theme of loyalty within the context of family or criminal organizations. *Gomorrah*'s characters are expected to show unwavering loyalty to their bosses

and associates. However, betrayal and treachery indicate that loyalty is not always absolute as characters may switch allegiances or betray others for personal gain. This observation aligns with our analysis, which shows a mixed picture of loyalty and lack thereof.

As to fairness (see Table 4), the absence of this moral dimension in the world of organized crime portrayed in *Gomorrah* is consistent with our findings. The series often depicts characters resorting to violence, coercion, or bribery to achieve their objectives rather than engaging in fair exchanges or negotiations. This absence of fairness further confirms its limited role within the moral dynamics of the series.

Finally, the analysis of the *Authority/Subversion* foundation reflects the negative aspect, where characters tend to operate outside the law or exploit their positions of power to mistreat others. The focus on bosses and their ability to establish new rules and subvert existing hierarchies reinforces the presence of subversion within the series. This observation aligns with the dynamics of power and control depicted in *Gomorrah*.

4.2.2. *13 Reasons Why*

Let us delve into the analysis of the teen drama series, *13 Reasons Why*.

13 Reasons Why is a Netflix original series that is based on the 2007 novel of the same name by Jay Asher [57]. The storyline revolves around the life of Hannah Baker, a high school student who tragically takes her own life. Before her death, Hannah records a series of cassette tapes in which she explains the *13 Reasons Why* she felt compelled to end her life. Each episode focuses on one of the reasons and explores the individuals who played a role in Hannah's life, including her classmates and friends. The series deals with themes that resonate with many teenagers (e.g., bullying, sexual assault, mental health, and suicide) and examines the role of social media and technology in the lives of teenagers, exploring how these platforms can connect and isolate individuals. The series unfolded over four seasons, comprising a total of 49 episodes, sparking significant discussions and raising awareness about mental health issues and the importance of suicide prevention.

We show the moral foundation scores for all the episodes in Figure 5. Here, we observe significant differences among the episodes. However, the *Care/Harm* dimension is largely dominant, with *Ingroup* being a close second (the distance may be considerable in a few episodes, but the overall rankings shown in Table 5 are rather similar).

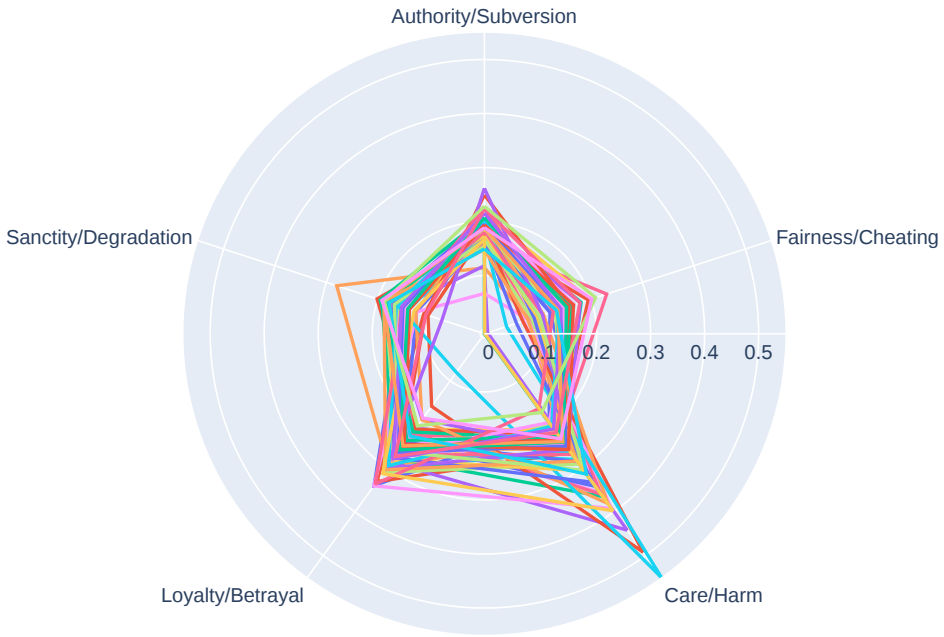

Figure 5. Moral foundation scores for all episodes in *13 Reasons Why*.

Table 5. Count of moral foundation rankings over all episodes in *13 Reasons Why*.

Ranking	**1**	**2**	**3**	**4**	**5**
Authority	2	10	21	14	2
Fairness	1	2	0	12	34
Harm	24	17	7	0	1
Ingroup	22	19	6	1	1
Purity	0	1	15	22	11

If we look at the polarity in Figure 6, again, all the modes are in the negative half. The *Subversion* component is dominant, with *Harm* being quite negative as well. On the other end of the spectrum, *Fairness*, although mostly negative, exhibits the smallest imbalance towards vice.

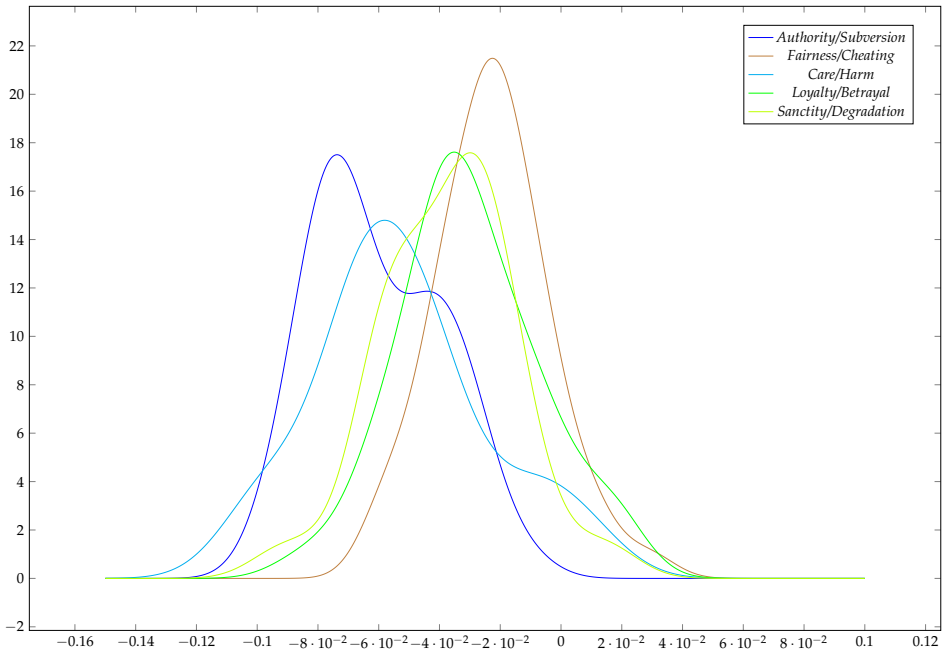

Figure 6. Empirical probability density of foundation polarity in *13 Reasons Why*.

Once again, we observe a strong correlation between the moral foundation scores and the thematic content of the series. The presence of the *Care/Harm* foundation is evident as the show addresses topics such as mental health, suicide, bullying, and sexual assault, all of which are closely related to the concept of harm. For instance, the character Hannah Baker, like many others, experiences harm through bullying and sexual assault, ultimately leading to her feeling neglected and unimportant. The *Loyalty/Betrayal* foundation is also prominent, particularly in the relationships between characters. Hannah feels betrayed by her former friend Jessica when she starts dating Hannah's crush, and Clay Jensen struggles with conflicting loyalties between Hannah and his other friends. The series predominantly adopts the viewpoint of teenage characters, and their relationships with parents are either absent from the narrative or depicted as negative, highlighting a lack of understanding and support. The relationships depicted in the series primarily revolve around peer interactions, often characterized by betrayal and broken trust. This emphasis on negative polarities further explains the dominant presence of negative moral dimensions in the moral foundation scores.

4.2.3. *Game of Thrones*

For the Fantasy genre, we have examined the series *Game of Thrones*.

It is an epic fantasy TV series based on George R.R. Martin's *A Song of Ice and Fire* novel. Set in the fictional world of Westeros, it follows noble houses competing for power

and the Iron Throne. Known for complex plotlines, world-building, and a large ensemble cast, it explores the themes of power, loyalty, and morality. *Game of Thrones* was aired from 2011 to 2019, spanning eight seasons and 73 episodes. It gained immense popularity and critical acclaim, winning numerous awards. The series captivated audiences with its mix of violence, sex, and political intrigue.

In Figure 7, the *Care/Harm* foundation appears dominant, ranking first in 47 episodes out of 73 (see Table 6). *Ingroup* is another strong foundation, with 18 first-place rankings. *Fairness*, as in *Gomorrah*, plays a negligible role, ranking last in all the episodes.

Table 6. Count of moral foundation rankings over all episodes in *Game of Thrones*.

Ranking	**1**	**2**	**3**	**4**	**5**
Authority	7	12	45	9	0
Fairness	0	0	0	0	73
Harm	47	16	6	4	0
Ingroup	18	39	9	7	0
Purity	1	6	13	53	0

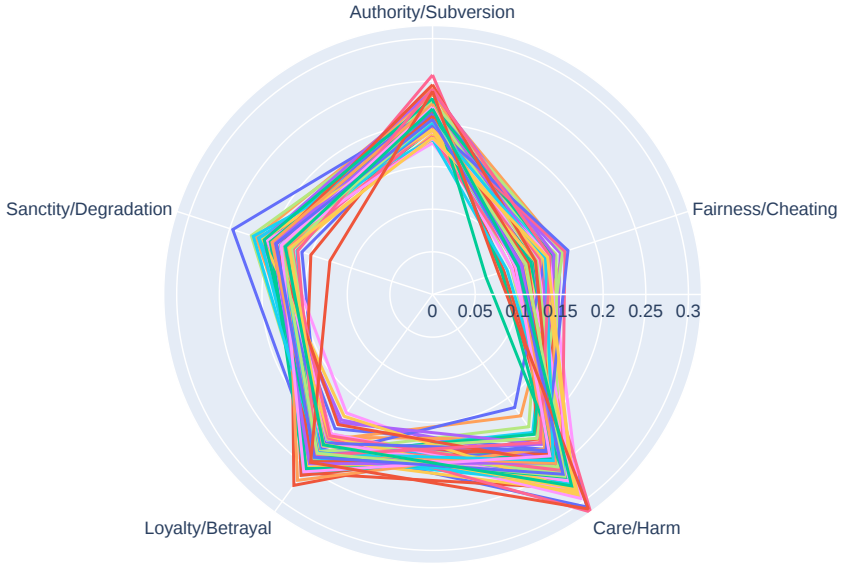

Figure 7. Moral foundation scores for all episodes in *Game of Thrones*.

All the polarity values are heavily shifted towards the vice side (see Figure 8), in particular for *Authority/Subversion* (hence, the *Subversion* facet predominates) and *Loyalty/Betrayal* (hence, *Betrayal* prevails). The least negative foundation is *Fairness/Cheating*.

We can relate the presence of the *Care/Harm* foundation through the characters' attitudes toward protecting the innocent and vulnerable. Characters like Jon Snow and Daenerys Targaryen prioritize the safety and well-being of those under their care. We also observe that the reference for the fantasy genre is the classical tragedy, where the dominant characters are mostly anti-heroes who end up being destroyed by their own vices. Furthermore, the show addresses mental health and trauma, which are connected to the *Care/Harm* foundation.

The themes of loyalty and betrayal, representing the ingroup foundation, play a central role in the series. Characters form alliances and allegiances with one another, emphasizing loyalty to family, friends, or allies. However, these relationships are often tested, leading to conflicts and betrayals, resulting in a negative polarity score. The consequences of breaking oaths and betraying trust are explored throughout the show.

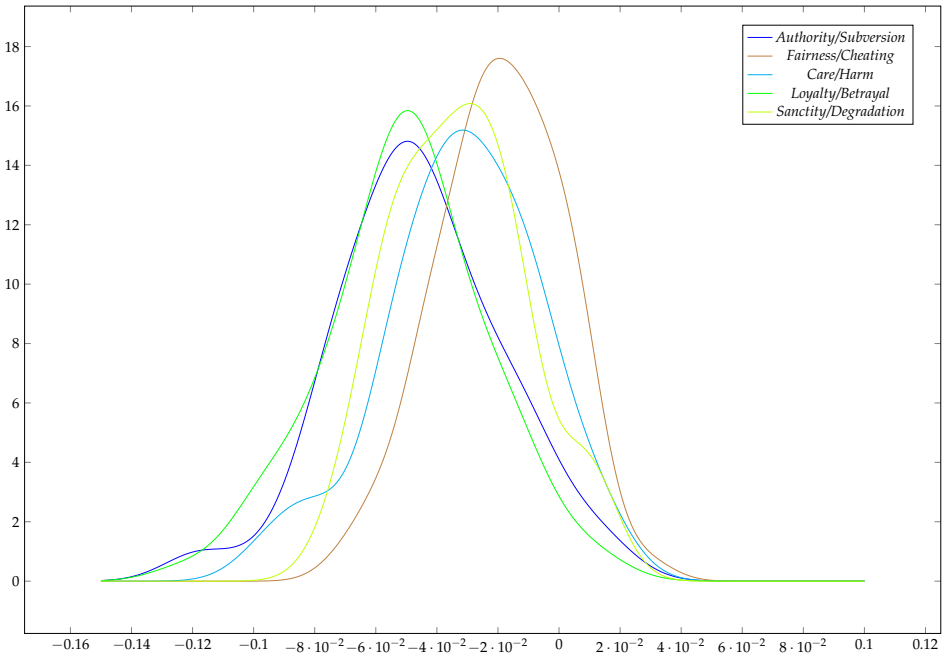

Figure 8. Empirical probability density of foundation polarity in *Game of Thrones*.

4.2.4. *House, MD*

House, MD is a medical drama series that was aired from 2004 to 2012, spanning eight seasons and 177 episodes. Starring Hugh Laurie as Dr. Gregory House, the show follows a team of doctors at Princeton-Plainsboro Teaching Hospital. The series' themes include medical ethics, the challenges of practicing medicine, and the personal struggles of the characters. It tackles complex medical cases, with House and his team using unconventional methods to uncover diagnoses. The show also features a recurring cast of supporting characters who contribute to the story.

This series stands out against those examined so far in that its mix of moral foundations is rather balanced, as can be seen in Figure 9, except a single episode where the *Sanctity/Degradation* pair takes the entire spotlight.

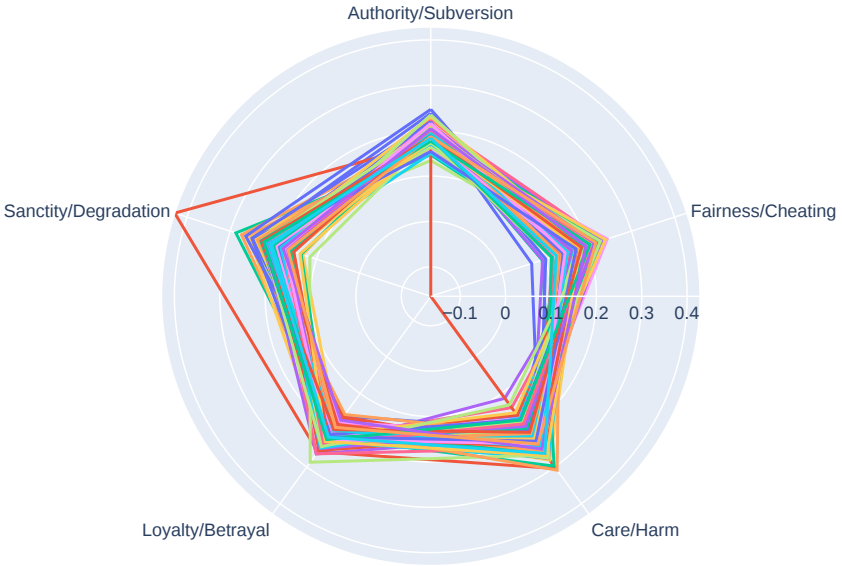

Figure 9. Moral foundation scores for all episodes in *House, MD*.

However, while all the moral foundations are represented in the series, there are a few that stand out by ranking higher than the others, although the differences in their rankings are not stark. If we look at Table 7, we see that *Ingroup* and *Harm* remain the most present foundations. *Fairness* is, again, the least present, although not lagging behind as much as in the other series.

Table 7. Count of moral foundation rankings over all episodes in *House, MD*.

Ranking	1	2	3	4	5
Authority	6	27	52	70	21
Fairness	10	13	19	35	99
Harm	64	39	32	18	23
Ingroup	71	49	36	14	6
Purity	25	48	37	39	27

The dominant polarity is negative for all the foundations, except *Care/Harm*, where the two components show quite a similar weight (see Figure 10).

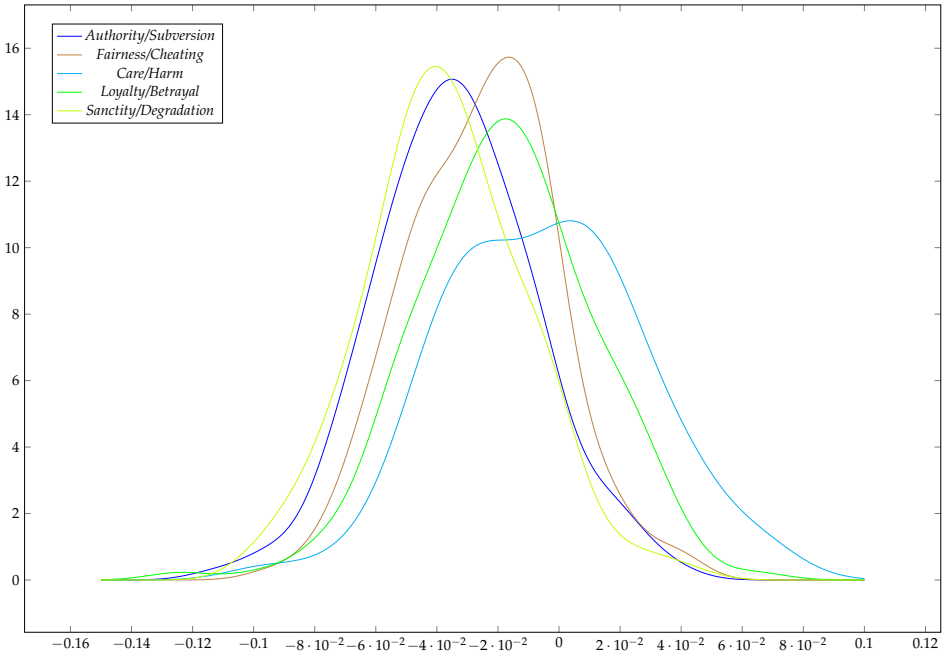

Figure 10. Empirical probability density of foundation polarity in *House, MD*.

The presence of the *Care/Harm* foundation is reasonably expected. Its presence shows through the medical cases that the main character, Dr. House, and his team handle. The team's goal is to provide care to patients and prevent harm to them. Dr. House, while often being cynical and irascible, is ultimately motivated by a desire to help his patients and cure their illnesses. His attitude makes a strong case for the dominant presence of *Care/Harm*, as appearing in Table 7, and also for its balance of polarities. The medical perspective leading to acting for the good of people tilts the picture towards the positive side, while Dr. House's behavior pushes the foundation towards the negative polarity. Dr. House's character is the opposite of the caring doctor in previous medical dramas. His rudeness and impoliteness make his role unique in that genre. The theme of loyalty and betrayal (*Ingroup*) is also present in the relationships between the characters. For example, the team members are loyal to Dr. House, even when they disagree with him or find his behavior difficult. Similarly, Dr. House is loyal to his patients, even when it means going against hospital policy or risking his own reputation. Overall, the series actually shows a conflict of loyalty-related behaviors: the team members may shift between being loyal to the institution or Dr. House, and a similar conflict is present in Dr. House himself, who may

decide to be loyal to his patients against the institution to which he belongs. The overall attitude is shown by the mostly negative polarity of this foundation in Figure 10.

4.2.5. *Sex and the City*

Sex and the City is a beloved romantic comedy–drama series that aired on HBO from 1998 to 2004, spanning six seasons and a total of 94 episodes. It revolves around the lives of four women, Carrie, Charlotte, Miranda, and Samantha, as they navigate love, relationships, and careers in the vibrant city of New York. The themes touched upon include love, sexuality, female empowerment, and the complexities of modern dating. It gained immense popularity and became a cultural phenomenon, influencing fashion and popular culture. The show's candid discussions about sex and its portrayal of independent women resonated with audiences worldwide.

The radar chart of foundations in Figure 11 shows a relatively balanced image, with all the foundations being present. However, some foundations rank consistently higher than others.

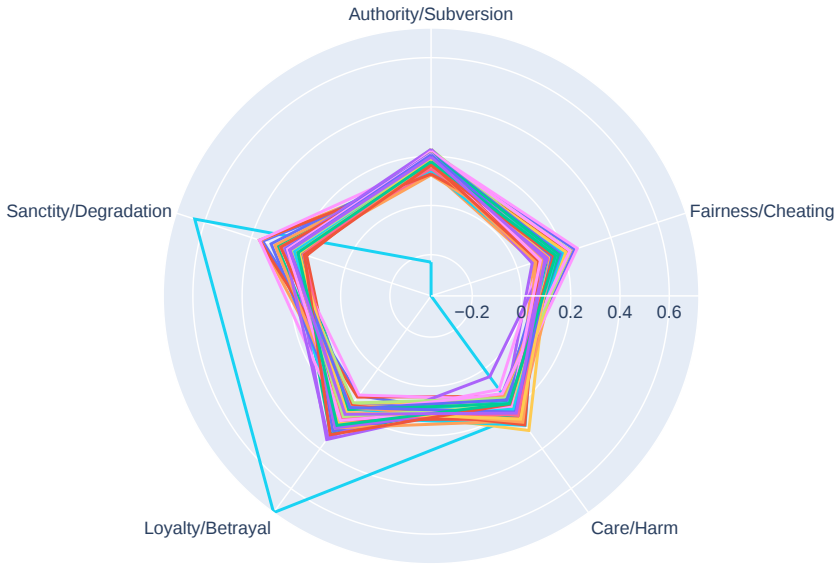

Figure 11. Moral foundation scores for all episodes in *Sex and the City*.

In Table 8, we observe that *Ingroup* and *Purity* are the dominant ones, while *Fairness* and *Authority* typically occupy the lowest positions.

Table 8. Count of moral foundation rankings over all episodes in *Sex and the City*.

Ranking	**1**	**2**	**3**	**4**	**5**
Authority	2	7	40	40	5
Fairness	3	6	5	14	66
Harm	10	21	16	26	21
Ingroup	42	24	16	11	1
Purity	37	36	17	3	1

The picture of polarities shifts away from the nearly entirely negative side we have observed so far. We can see in Figure 12 that the *Purity* foundation exhibits a dominant positive polarity, and *Loyalty* has a significant positive tail as well.

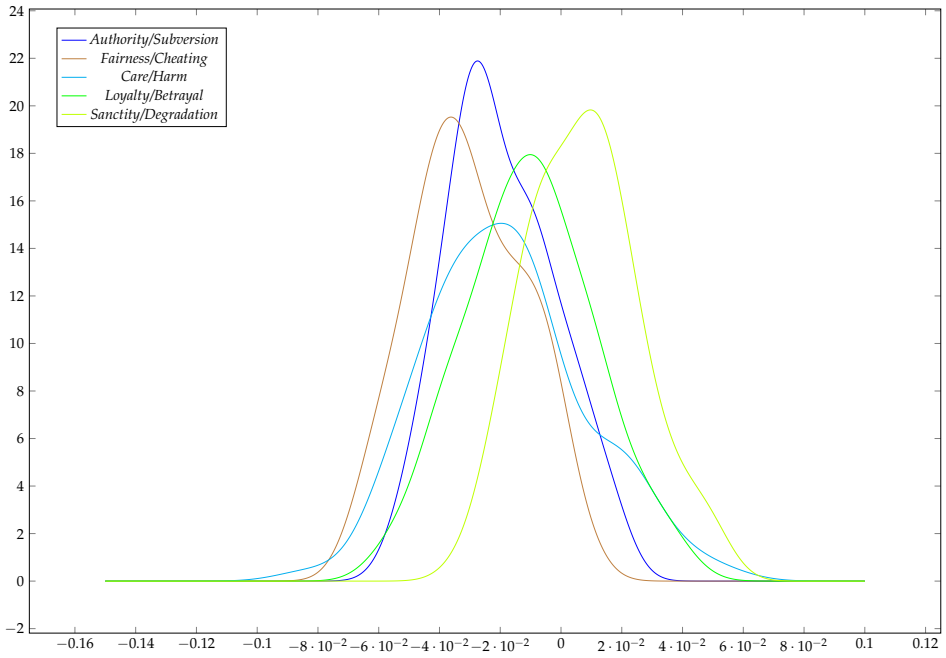

Figure 12. Empirical probability density of foundation polarity in *Sex and the City*.

We can recognize that *Ingroup* is the dominant foundation shown in Table 8. For example, the main characters are loyal to each other and often support each other through difficult times. However, the show also explores the ways in which loyalty can be tested and betrayed as the characters occasionally experience conflicts or betrayals in their relationships. We experience some difficulties in reconciling the results of our analysis with the series' content for the *Sanctity/Degradation* pair. The *Sanctity/Degradation* foundation is present in the show through the characters' attitudes toward sex and sexuality. For example, the show portrays a range of sexual behaviors and attitudes, some of which challenge the traditional notions of sanctity and morality. The characters often grapple with issues of shame and guilt around their sexual behavior, and the show explores the ways in which sexual expression can be stigmatized or celebrated. We would expect a strong presence of the *Purity* foundation (which we actually observe) but tilted towards the negative polarity, while we observe a different dominant polarity in Figure 12. The low rank of the *Authority* foundation is well-explained by the characters' identity. They are a group of friends who share their joys, fears, and loves. They consider themselves independent and liberated under a feminist framework and do not recognize the authority and the rules of a chauvinist society. Actually, they ignore those rules rather than openly subverting them, recognizing just their friendship and loyalty to one another.

4.2.6. *Suits*

Suits is a captivating legal drama series that was aired on USA Network from 2011 to 2019, spanning nine seasons and a total of 134 episodes. The show revolves around the story of Mike Ross, a brilliant college dropout, who finds himself working as a law associate for the charismatic and successful lawyer Harvey Specter despite lacking a law degree. The series explores the intricacies of the law, showcasing legal strategies, client relationships, and the cutthroat world of corporate politics. It also delves into the personal lives of the characters, exploring romantic relationships, familial conflicts, and personal growth. The diverse ensemble cast includes lawyers, paralegals, secretaries, and corporate executives, working together to overcome challenges and achieve their goals. The success of the series led to a spinoff titled *Pearson* in 2019, focusing on the character Jessica Pearson.

The picture of the moral foundation scores in all the episodes in Figure 13 appears quite blurry. The two dimensions of *Care/Harm* and *Fairness/Cheating* exhibit significant peaks but also a substantial dispersion.

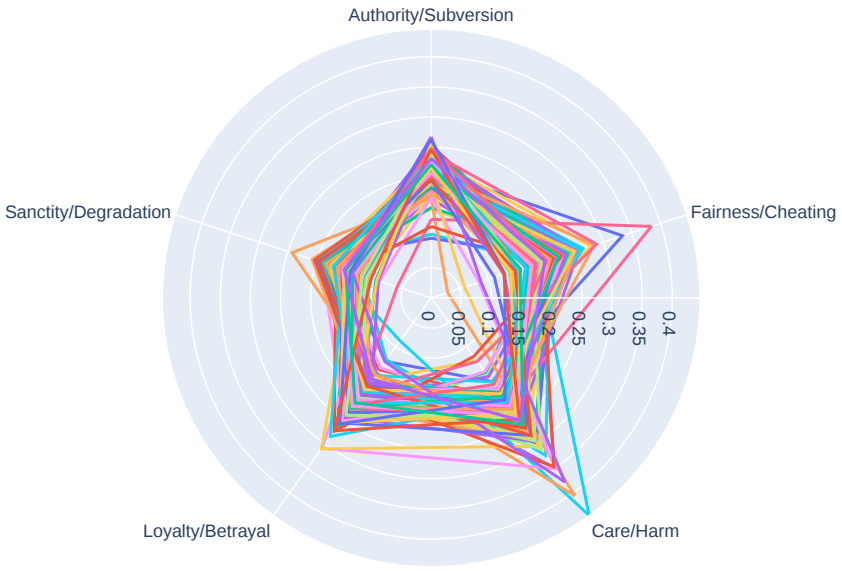

Figure 13. Moral foundation scores for all episodes in *Suits*.

The picture becomes clearer if we look at the relative positioning in Table 9. The *Care/Harm* dimension emerges as the dominant one in most episodes, with *Fairness/Cheating* and *Ingroup* being distant runner-ups. Instead, we find *Purity* lying nearly always at the bottom.

Table 9. Count of moral foundation rankings over all episodes in *Suits*.

Ranking	**1**	**2**	**3**	**4**	**5**
Authority	10	39	58	21	6
Fairness	33	18	20	41	22
Harm	65	37	16	11	5
Ingroup	26	37	29	32	10
Purity	0	3	11	29	91

The actual polarity of all the foundations is again mostly negative, as shown in Figure 14. The mode is well in the negative field for *Authority/Subversion* and *Care/Harm*.

We can confirm the dominant presence of the *Care/Harm* foundation through the legal cases that the main characters handle. The characters often work to protect their clients and prevent them from being harmed by legal and business issues. Additionally, the show explores issues of mental health and addiction, both of which relate to the *Care/Harm* foundation. The polarity analysis agrees more with the latter observation than with the former. In addition to the previously mentioned foundations, *Suits* also touches upon the *Fairness/Cheating* foundation through its portrayal of the characters' attitudes towards justice and the legal system. The main characters consistently strive for fairness in their legal cases, working diligently to expose corruption and uncover the truth. However, the show also delves into the darker side, exploring how cheating and deception can be employed to gain an advantage within the legal and business realms. This aspect is prominently reflected in our polarity analysis. Furthermore, the themes of loyalty and betrayal (*Ingroup*) comprise a significant element within the show, influencing the dynamics between the characters. Loyalty is a crucial virtue, both personally and professionally, as the characters display unwavering support for one another. However, *Suits* also delves into the complexities of loyalty, showcasing instances where trust is tested and betrayal occurs, resulting in conflicts within relationships. Once again, our polarity analysis emphasizes this latter aspect, highlighting the presence of negative polarities.

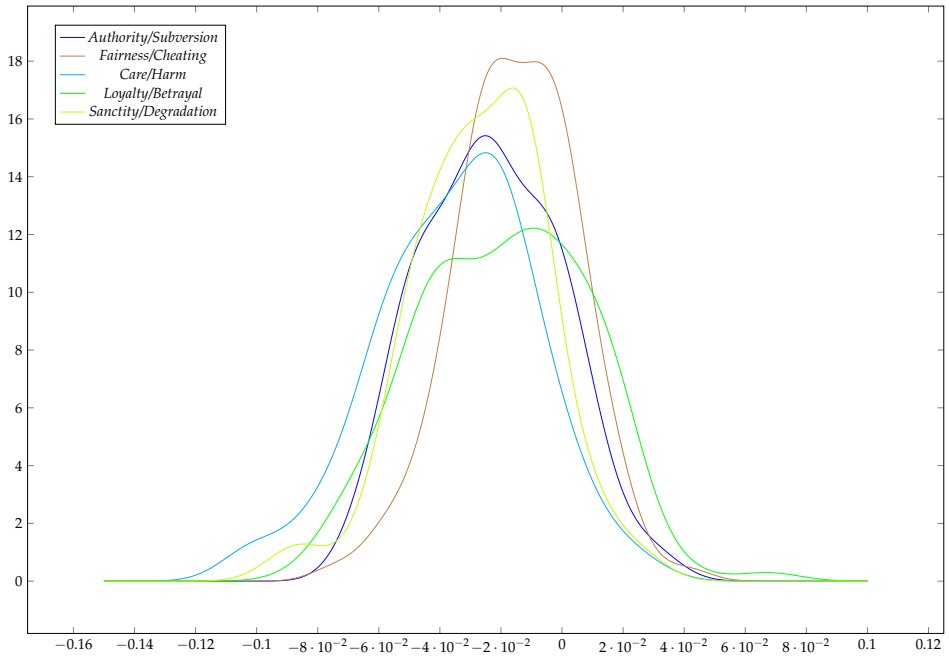

Figure 14. Empirical probability density of foundation polarity in *Suits*.

4.2.7. *Westworld*

Westworld is a science fiction television series that premiered on HBO in 2016 and is still ongoing. The show is inspired by the 1973 film of the same name. It is set in a futuristic amusement park called *Westworld*, where guests can live out their wildest fantasies in a hyper-realistic Wild West environment populated by android hosts. The series explores themes related to artificial intelligence, consciousness, morality, and the nature of reality as the characters confront the ethical and existential implications of creating and interacting with lifelike androids. The show is known for its intricate plotlines, complex characters, and thought-provoking philosophical themes. The show has also spawned a devoted fanbase that eagerly anticipates each new season and dissects each episode for hidden clues and plot twists.

In the radar chart shown in Figure 15, two foundations emerge: *Care/Harm* and *Sanctity/Degradation*. We observe again an outlier episode, where *Care/Harm* is highly dominant over the other foundations.

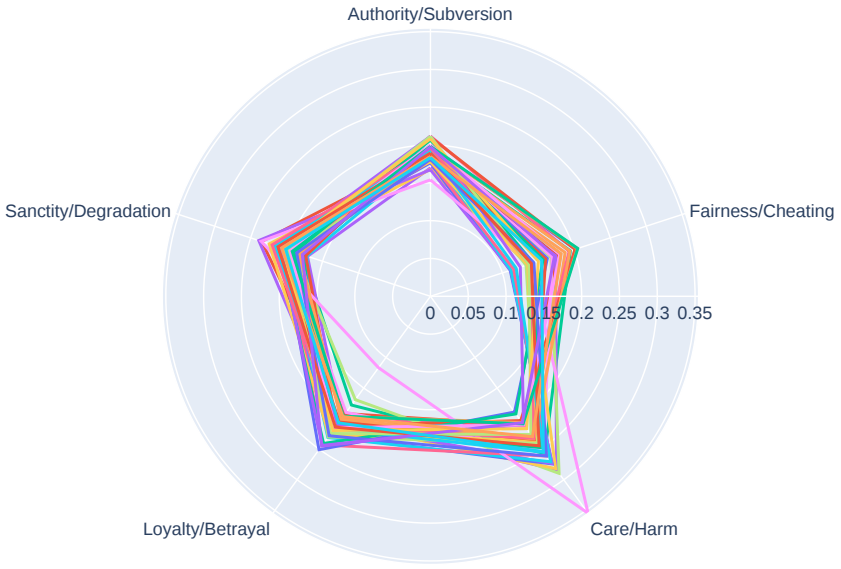

Figure 15. Moral foundation scores for all episodes in *Westworld*.

This picture is to be slightly reconsidered when we take a look at the table of rankings in Table 10, which confirms the dominant role of *Care/Harm* but ranks *Ingroup* as the runner-up rather than *Purity*. The least present foundation is *Fairness*.

Table 10. *Westworld*: count of moral foundation rankings over all episodes.

Ranking	1	2	3	4	5
Authority	0	3	14	18	1
Fairness	0	3	4	4	25
Harm	26	6	1	3	0
Ingroup	9	13	9	3	2
Purity	1	11	8	8	8

The polarity density in Figure 16 shows a sharp negative polarity for all the foundations except two, *Fairness* and *Purity*, where the positive polarity does not lag behind as much as in the other foundations.

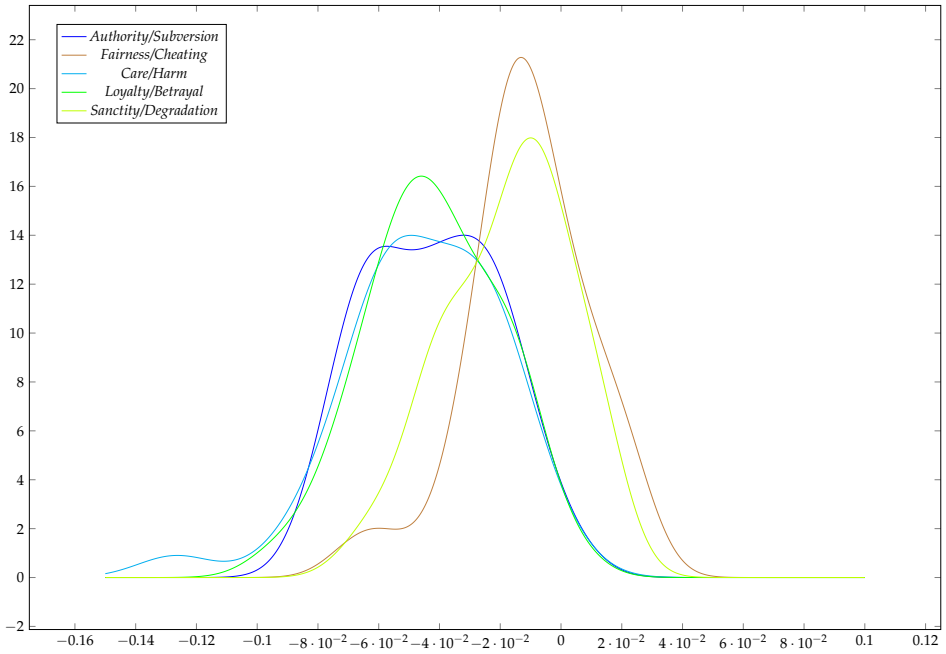

Figure 16. Empirical probability density of foundation polarity in *Westworld*.

In the TV show being discussed, the foundation of *Care/Harm* holds significant importance, evident through its exploration of consciousness and empathy. The AI hosts, who attain self-awareness, begin to question their own suffering and the suffering of their fellow hosts. This introspection prompts an examination of the ethical implications of humans exploiting the hosts for their own ends. Consequently, the themes of consciousness and empathy serve as a central pillar in the narrative.

Moreover, the foundation of *Sanctity/Degradation* emerges prominently, focusing on matters of morality and ethics. Both the hosts and the human characters grapple with profound moral and ethical dilemmas as they delve into questions surrounding existence and consciousness. The repercussions of violating moral and ethical boundaries in the pursuit of scientific progress and personal desires are explored throughout the show.

4.2.8. Overall Comparison

Figure 17 offers a comprehensive overview of the prevalence of the moral dimensions across the series we have examined, allowing readers to compare and contrast the moral foundations that underpin each show. This visual representation serves as a valuable tool for understanding the moral landscapes that shape the narratives of these series.

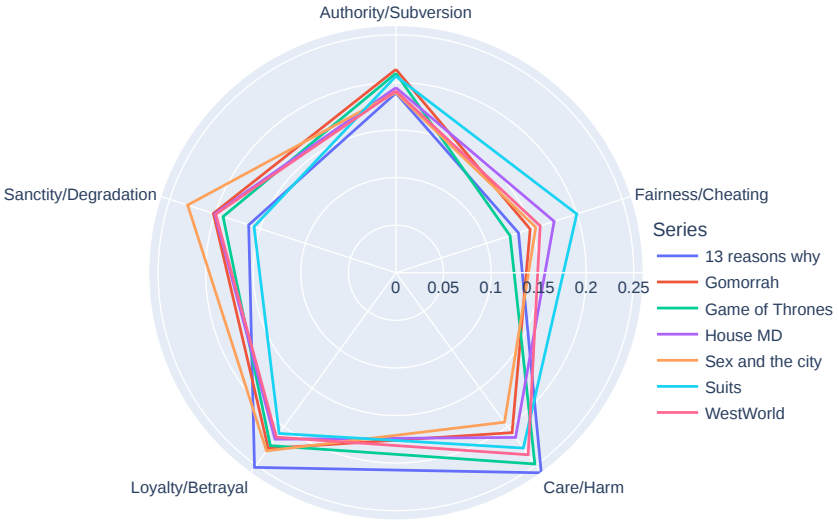

Figure 17. Radar chart of moral foundations distribution over the series.

By examining the radar chart, readers can gain insights into the relative prominence of different moral dimensions within each show. This information facilitates a deeper exploration of the ethical considerations and thematic emphasis present in the series, stimulating further discussion and analysis.

Figure 17 reveals the presence of multiple moral foundations across the selected series. The *Care/Harm* foundation emerged prominently in many shows, highlighting themes of empathy, protection of the vulnerable, and the consequences of harm inflicted upon others. This foundation was particularly evident in series such as *13 Reasons Why*, where issues of mental health, bullying, and sexual assault were explored, and *Westworld*, where the consciousness and suffering of AI hosts were central to the storyline.

The *Fairness/Cheating* foundation was also prevalent, as depicted through the pursuit of justice, exposure of corruption, and exploration of ethical dilemmas. It was prominent in series such as *Suits*, where the characters sought fairness in the legal system while navigating the complexities of a cutthroat industry. This foundation was also observed in *Westworld*, where the exploitation of hosts for personal gain raised questions about fairness and the consequences of cheating.

Additionally, the *Sanctity/Degradation* foundation emerged in the exploration of morality, ethics, and the violation of moral codes. Series like *Westworld* examined the consequences of violating ethical boundaries in the pursuit of scientific advancement and personal gain. In contrast, *Game of Thrones* explored the moral complexities and consequences of actions in a power-driven world.

Finally, the *Loyalty/Betrayal* foundation emerged as a significant theme in several series, vividly portraying the dynamics of trust, loyalty, and the potential for betrayal. One notable example is the series *Gomorrah*, which delves into the gritty and dangerous world of a criminal society.

5. Conclusions

This paper has provided an insightful exploration of the moral dimensions present in a selection of popular television series across different genres. By analyzing these series, valuable insights have been gained into the moral landscapes that shape their narratives, offering a foundation for deeper exploration and discussion.

Through the comparison of moral dimensions, the distinct moral landscapes of each series have been highlighted. It is evident that each show, within its genre and storytelling, prioritizes specific moral foundations, influencing the themes, character dynamics, and overall messages conveyed to the audience. However, while certain moral foundations may take precedence within a particular series, other foundations are still present to varying degrees, showing the complexity and richness of the storytelling.

TV series have the potential to shape viewers' moral attitudes and beliefs, making it crucial for viewers to engage critically with these series. By actively considering the moral foundations presented, questioning their portrayals, and reflecting on how they align with personal beliefs and experiences, viewers can deepen their understanding of the cultural, social, and ethical significance of these shows.

Further research is needed to explore the link between the moral dimensions embedded in TV shows and the moral behavior of their viewers. Understanding the impact of these shows on viewers' moral complexities and behavior will provide valuable insights into the societal implications of TV series as a medium for exploring moral questions.

Author Contributions: The authors' individual contributions are shown in the following: Conceptualization, M.R., M.N. and V.R.; methodology, M.N. and P.F.; software, P.F.; validation, M.R., V.R., M.N. and P.F.; formal analysis, M.N. and P.F.; investigation, M.R., M.N., P.F., V.R. and P.D.T.; data curation, P.F.; writing—original draft preparation, M.N., M.R., V.R. and P.D.T.; writing—review and editing, M.N., M.R., V.R. and P.D.T.; visualization, P.F.; supervision, M.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: No new data were created.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BERT	Bidirectional Encoder Representations from Transformers
DOAJ	Directory of open access journals
MDPI	Multidisciplinary Digital Publishing Institute
MFD	*Moral Foundations Dictionary*
MFQ	Moral Foundations Questionnaire
MFT	Moral Foundations Theory
SBERT	Sentence-Bidirectional Encoder Representations from Transformers

References

1. Bergmann, J.R. Introduction: Morality in discourse. *Res. Lang. Soc. Interact.* **1998**, *31*, 279–294. [CrossRef]
2. Hemming, J. Morality after myth. *J. Moral Educ.* **1996**, *25*, 39–45. [CrossRef]
3. White, H. The value of narrativity in the representation of reality. *Crit. Inq.* **1980**, *7*, 5–27. [CrossRef]
4. Burke, P. *The Art of Conversation*; John Wiley & Sons: Hoboken, NJ, USA, 2013.
5. Shaw, D. *Morality and the Movies: Reading Ethics through Film*; Bloomsbury Publishing: London, UK, 2012.
6. Carroll, N. Movies, the moral emotions, and sympathy. *Midwest Stud. Philos.* **2010**, *34*, 1–19. [CrossRef]
7. Bilandzic, H.; Hastall, M.R.; Sukalla, F. The morality of television genres: Norm violations and their narrative context in four popular genres of serial fiction. *J. Media Ethics* **2017**, *32*, 99–117. [CrossRef]
8. Vaage, M.B. Blinded by familiarity: Partiality, morality, and engagement with TV series. In *Cognitive Media Theory*; Routledge: London, UK, 2014; pp. 268–284.
9. La Ferrara, E. Mass media and social change: Can we use television to fight poverty? *J. Eur. Econ. Assoc.* **2016**, *14*, 791–827. [CrossRef]
10. Bilandzic, H. The complicated relationship between media and morality. *J. Media Psychol.* **2011**, *23*, 46–51. [CrossRef]
11. Haidt, J. The new synthesis in moral psychology. *Science* **2007**, *316*, 998–1002. [CrossRef]
12. Jöckel, S.; Früh, H. 'The world ain't all sunshine': Investigating the relationship between mean world beliefs, conservatism and crime TV exposure. *Communications* **2016**, *41*, 195–217. [CrossRef]
13. Joeckel, S.; Bowman, N.D.; Dogruel, L. Gut or game? The influence of moral intuitions on decisions in video games. *Media Psychol.* **2012**, *15*, 460–485. [CrossRef]
14. Tamborini, R.; Eden, A.; Bowman, N.D.; Grizzard, M.; Weber, R.; Lewis, R.J. Predicting media appeal from instinctive moral values. *Mass Commun. Soc.* **2013**, *16*, 325–346. [CrossRef]
15. Beveridge, A.; Shan, J. Network of thrones. *Math Horiz.* **2016**, *23*, 18–22. [CrossRef]
16. Bost, X.; Labatut, V.; Gueye, S.; Linares, G. Narrative smoothing: Dynamic conversational network for the analysis of TV series plots. In Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), San Francisco, CA, USA, 18–21 August 2016; pp. 1111–1118.

17. Tan, M.S.; Ujum, E.A.; Ratnavelu, K. A character network study of two Sci-Fi TV series. In Proceedings of the AIP Frontiers in Physics Conference: 4th International Meeting. American Institute of Physics, Kuala Lumpur, Malaysia 27–30 August 2013; Volume 1588, pp. 246–251.
18. Zhang, L.; Li, C.; Fan, L.; Shi, M. Analysis of character relationship in tv series based on complex network. In Proceedings of the 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), Singapore, 6–8 June 2018; pp. 723–727.
19. Fronzetti Colladon, A.; Naldi, M. Concentration indices for dialogue dominance phenomena in TV series: The case of the Big Bang Theory. In *Proceedings of the Text Analytics: Advances and Challenges*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 55–64.
20. Fronzetti Colladon, A.; Naldi, M. Predicting the performance of TV series through textual and network analysis: The case of Big Bang Theory. *PLoS ONE* **2019**, *14*, e0225306. [CrossRef]
21. Rocchi, M.; Pescatore, G. Modeling narrative features in TV series: Coding and clustering analysis. *Humanit. Soc. Sci. Commun.* **2022**, *9*, 333. [CrossRef]
22. Naldi, M.; Dalla Torre, P. Uncovering the Narrative Structure of "Breaking Bad" Through a Multi-Dimensional Quantitative Analysis. *Ser. Int. J. Ser. Narrat.* **2022**, *8*, 17–32.
23. Graham, J.; Haidt, J.; Koleva, S.; Motyl, M.; Iyer, R.; Wojcik, S.P.; Ditto, P.H. Moral foundations theory: The pragmatic validity of moral pluralism. In *Advances in Experimental Social Psychology*; Elsevier: Amsterdam, The Netherlands, 2013; Volume 47, pp. 55–130.
24. Curry, O.S.; Chesters, M.J.; Van Lissa, C.J. Mapping morality with a compass: Testing the theory of 'morality-as-cooperation' with a new questionnaire. *J. Res. Personal.* **2019**, *78*, 106–124. [CrossRef]
25. Zakharin, M.; Bates, T.C. Remapping the foundations of morality: Well-fitting structural model of the Moral Foundations Questionnaire. *PLoS ONE* **2021**, *16*, e0258910. [CrossRef]
26. Haidt, J. *The Righteous Mind: Why Good People are Divided by Politics and Religion*; Knopf Doubleday Publishing Group: New York, NY, USA, 2012.
27. Welsch, H. Moral foundations and voluntary public good provision: the case of climate change. *Ecol. Econ.* **2020**, *175*, 106696. [CrossRef]
28. Silver, J.R. Moral foundations, intuitions of justice, and the intricacies of punitive sentiment. *Law Soc. Rev.* **2017**, *51*, 413–450. [CrossRef]
29. Hopp, F.R.; Amir, O.; Fisher, J.T.; Grafton, S.; Sinnott-Armstrong, W.; Weber, R. Moral foundations elicit shared and dissociable cortical activation modulated by political ideology. *Nat. Hum. Behav.* **2023**, *7*, 2182–2198. [CrossRef]
30. Graham, J.; Haidt, J.; Nosek, B.A. Liberals and conservatives rely on different sets of moral foundations. *J. Personal. Soc. Psychol.* **2009**, *96*, 1029. [CrossRef]
31. Graham, J.; Nosek, B.A.; Haidt, J.; Iyer, R.; Koleva, S.; Ditto, P.H. Mapping the moral domain. *J. Personal. Soc. Psychol.* **2011**, *101*, 366. [CrossRef]
32. Wheeler, M.A.; McGrath, M.J.; Haslam, N. Twentieth century morality: The rise and fall of moral concepts from 1900 to 2007. *PLoS ONE* **2019**, *14*, e0212267. [CrossRef]
33. Lin, Y.; Michel, J.B.; Lieberman, E.A.; Orwant, J.; Brockman, W.; Petrov, S. Syntactic annotations for the Google Books Ngram corpus. In Proceedings of the ACL 2012 System Demonstrations, Jeju, Republic of Korea, 8–14 July 2012; pp. 169–174.
34. Long, J.A.; Eveland, W.P., Jr. Entertainment use and political ideology: Linking worldviews to media content. *Commun. Res.* **2021**, *48*, 479–500. [CrossRef]
35. Bowman, N.D.; Jöckel, S.; Dogruel, L. A question of morality? The influence of moral salience and nationality on media preferences. *Commun.-Eur. J. Commun. Res.* **2012**, *37*, 345–369. [CrossRef]
36. Ji, Q.; Raney, A.A. Morally judging entertainment: A case study of live tweeting during Downton Abbey. *Media Psychol.* **2015**, *18*, 221–242. [CrossRef]
37. Gehman, R.; Guglielmo, S.; Schwebel, D.C. Moral foundations theory, political identity, and the depiction of morality in children's movies. *PLoS ONE* **2021**, *16*, e0248928. [CrossRef]
38. Imelwaty, S.; Gunawan, F.; Kuraedah, S.; Mardiana, W.; Wardhana, D.E.C.; Boulahnane, S. Moral values in Indonesian primary school thematic textbooks: The marrying of language appraisal and moral foundation theories. In *Education 3-13*; Taylor and Francis: Milton Park, UK, 2022; pp. 1–13.
39. Araque, O.; Zhu, G.; Iglesias, C.A. A semantic similarity-based perspective of affect lexicons for sentiment analysis. *Knowl.-Based Syst.* **2019**, *165*, 346–359. [CrossRef]
40. Araque, O.; Gatti, L.; Kalimeri, K. MoralStrength: Exploiting a moral lexicon and embedding similarity for moral foundations prediction. *Knowl.-Based Syst.* **2020**, *191*, 105184. [CrossRef]
41. González-Santos, C.; Vega-Rodríguez, M.A.; Pérez, C.J.; López-Muñoz, J.M.; Martínez-Sarriegui, I. Automatic assignment of moral foundations to movies by word embedding. *Knowl.-Based Syst.* **2023**, *270*, 110539. [CrossRef]
42. Campos, R.; Mangaravite, V.; Pasquali, A.; Jorge, A.M.; Nunes, C.; Jatowt, A. Yake! collection-independent automatic keyword extractor. In Proceedings of the Advances in Information Retrieval: 40th European Conference on IR Research, ECIR 2018, Grenoble, France, 26–29 March 2018; Proceedings 40; Springer: Berlin/Heidelberg, Germany, 2018; pp. 806–810.

43. Guglielmo, G.; Klincewicz, M. The temperature of morality: A behavioral study concerning the effect of moral decisions on facial thermal variations in video games. In Proceedings of the 16th International Conference on the Foundations of Digital Games, Online, 2–6 August 2021; pp. 1–4.
44. Hodge, S.E.; Taylor, J.; McAlaney, J. (A) morally demanding game? An exploration of moral decision-making in a purpose-made video game. *Media Commun.* **2020**, *7*, 213–225. [CrossRef]
45. Hornbeck, R.G. Moral Cognition Empowers Spiritual Experience in Chinese World of Warcraft. In *Religious Cognition in China: "Homo Religiosus" and the Dragon*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 179–194.
46. Krcmar, M.; Cingel, D.P. Moral foundations theory and moral reasoning in video game play: Using real-life morality in a game context. *J. Broadcast. Electron. Media* **2016**, *60*, 87–103. [CrossRef]
47. Creeber, G. *The Television Genre Book*; Bloomsbury Publishing: London, UK, 2015.
48. Li, Y.; Yang, T. Word embedding for understanding natural language: A survey. In *Guide to Big Data Applications*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 83–104.
49. Sitikhu, P.; Pahi, K.; Thapa, P.; Shakya, S. A comparison of semantic similarity methods for maximum human interpretability. In Proceedings of the 2019 Artificial Intelligence for Transforming Business and Society (AITB), IEEE, Kathmandu, Nepal, 5 November 2019; Volume 1, pp. 1–4.
50. Russell, S.J.; Norvig, P. *Artificial Intelligence: A Modern Approach*; Pearson Educaiton Limited: Harlow, UK, 2021.
51. Silverman, B.W. *Density Estimation for Statistics and Data Analysis*; CRC Press: Boca Raton, FL, USA, 1986; Volume 26.
52. Sheather, S.J.; Jones, M.C. A reliable data-based bandwidth selection method for kernel density estimation. *J. R. Stat. Soc. Ser. B Methodol.* **1991**, *53*, 683–690. [CrossRef]
53. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019; pp. 3982–3992.
54. Devlin, J.; Chang, M.W.; Lee, K.; Toutanova, K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv* **2018**, arXiv:1810.04805.
55. Stevenson, A. *Oxford Dictionary of English*; Oxford University Press: Oxford, UK, 2010.
56. Saviano, R. *Gomorrah*; Pan Macmillan: New York, NY, USA, 2019; Volume 82.
57. Asher, J. *13 Reasons Why*; Penguin: New York, NY, USA, 2017.