Personality Prediction of Narrative Characters from Movie Scripts

Anonymous ACL submission

Abstract

An NLP model that understands stories should be able to understand the characters in them. To support the development of neural models for this purpose, we construct a benchmark, `Story2Personality`. The task is to predict a movie character's personality based on the narratives of the character in the movie script. Experiments show that our task is challenging for the existing text classification models, as none is able to largely outperform random guesses. We further proposed a multi-view model to use both verbal and non-verbal descriptions for personality prediction, which gives improvement compared to using only verbal descriptions. The uniqueness and challenges in our dataset call for the development of narrative comprehension techniques from the perspective of understanding characters.[1]

1 Introduction

Plots and characters jointly play the central role in narratives (Riedl and Young, 2010). However, while plot comprehension in machine narrative understanding (Sims et al., 2019; Lal et al., 2021; Mou et al., 2021) has become an active topic, the issue of character comprehension is largely ignored.

In this work, we propose a new narrative NLP benchmark of personality prediction, the `Story2Personality`, to encourage the study of character understanding. The task collects movie characters' MBTI personalities judged by human raters in an online community and their movie scripts. The goal is to predict personality according to the character's narrative texts in the script.

Compared to existing datasets that focus on predicting real people's self-reported personality test results from their verbal expressions on social networks, e.g., Twitter (Qiu et al., 2012; Golbeck et al., 2011) or Reddit posts (Flekova and Gurevych, 2015; Gjurković and Šnajder, 2018), we identify

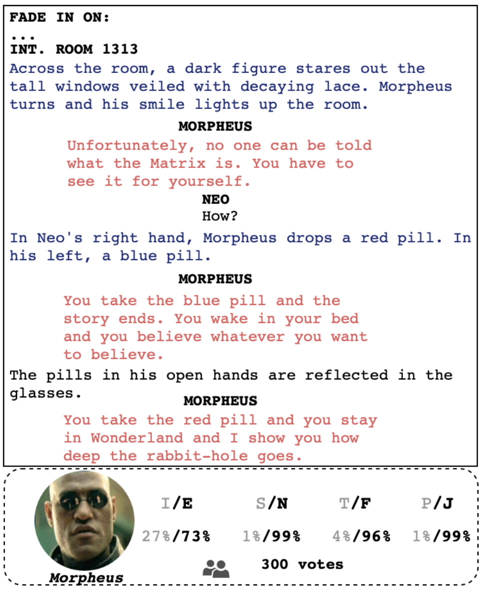

Figure 1: An example excerpt from "The Matrix" movie script. Blue utterances are mapped to the character *Morpheus's* **scene descriptions**, red are his **dialogues**. *Morpheus's* personality rated as **ENFJ** by 300 user votes

new challenges in personality prediction from narratives: (1) The writers usually use complex writing skills and styles[2] to engage the readers into the stories, making the understanding of narrative texts more difficult; (2) the inputs of the task are very long (>10K tokens on average), challenging the applications of Transformer-based models (Vaswani et al., 2017); (3) both the scene descriptions and dialogues are informative for the prediction, requiring models to jointly consider multi-view of inputs.

We make the following contributions:

- We establish the first large-scale dataset for personality prediction of narrative characters that can support the development of neural models. Our dataset consists of 3,543 characters with labels across all four MBTI dimensions. In comparison, the only existing related dataset (Flekova and Gurevych, 2015) contains only 298 book

[1] Our code and data will be released.

[2] Examples include variable narrative sequence (e.g., narrative, flashback, interpolation); and a variety of expressions (e.g., argument, lyricism, illustration).

characters and focuses on a single dimension. Our dataset is proved challenging — on this binary classification task, none of the baselines achieve higher than 60% macro-F1.

- We develop a movie script parser to automatically process a script to a structured form with the verbal character dialogues and the non-verbal scene descriptions illustrating backgrounds. Human study shows that our parser is more accurate compared to previous rule-based tools.
- We propose an extension of the BERT (Devlin et al., 2018) classifier to handle the long and multi-view (verbal and non-verbal) inputs. It gives a 2-3% improvement over the baselines. This shows the potential of exploiting both verbal and non-verbal narratives of characters, which is consistent with psychology research (McCroskey and Richmond, 1996; Richmond et al., 2008); and suggests the future direction for new model development.

2 Background of MBTI

Personality is a psychological construct aimed at "distinguish(ing) internal properties of the person from overt behaviors" (Matthews et al., 2003) and is a "stable and measurable" individual characteristics (Vinciarelli and Mohammadi, 2014). Understanding the personalities of the characters is important for understanding the deeper message of the story. The most popular personality scales are **Big-5** (Digman, 1997) and **MBTI** (Myers and McCaulley, 1988). We chose the MBTI as the fictional character's personality, because the MBTI can be used as a **third-person evaluation technique** (Cohen et al., 1981). Human raters satisfy this criterion because they are fans of the movie, thus know the character well. Together with the voting from multiple raters, this ensures the accuracy of the labels.

MBTI has four dimensions: E/I: extravert (E) is seen as being generally active and objective while the intravert (I) is seen as generally passive and subjective (Sipps and Alexander, 1987). S/N: sensing (S) is seen as attending to sensory stimuli; intuition (N) describes a more detached, insightful analysis of events and stimuli (Boyle, 1995). T/F: thinking (T) involves logical reasoning and decision making; feeling (F) involves a more subjective and interpersonal approach (Thomas, 1983). J/P: judging (J) attitude is associated with prompt decision making; perception (P) involves greater patience and waiting for more information before making a decision.

An individual's MBTI type has a label based on her dominant preference for each dimension. In our example, Morpheus is an extraversion person, understanding the world with intuition, dealing with things with feeling, and organize the world around him by judging. Together gives an ENFJ type.

3 `Story2Personality` Dataset

We constructed our dataset in three stages: extracting movie scripts from the Internet Movie Script Database (*IMSDB*[3]), parsing the collected movie scripts into dialogue and scene sections, matching characters' personality types from *The Personality Database*[4] with their dialogues and scene sections.

3.1 Movie Scripts Collection

Movie scripts describe all of the elements that are required to tell a story (Jhala, 2008). We collected the movie scripts in the form of HTML files from IMSDB combined with movie scripts in NarrativeQA (Kočiskỳ et al., 2018). After removing corrupted or empty files, we got 1,464 usable scripts.

3.2 Our Statistical Movie Script Parser

As shown in Figure 1, a movie script usually has four basic format elements (Riley, 2009): **Scene Headings**, one line description of each scene's type, location, and time (i.e., `INT. ROOM 1313`); **Scene description**, the description of the actions of the characters (i.e., text in blue); **Dialogues**, names of characters and actual words they speak (i.e., text in red); **Transitions**, instructions for linking scenes together (i.e., `FADE IN ON`).

In order to extract information related to our task (i.e., dialogues and scene descriptions) in a structured form, we first split the scripts to sections, i.e., text chunks between two adjacent bolded chunks (i.e., scene headings or character names, which were stored as section titles). Then we designed a statistical method to classify the section types:

Rule-Based Pre-Processing We start with a rule to roughly classify the sections into dialogues and scenes. As shown in Figure 1, a common format of movie scripts is to align the shot headings, transitions and scene descriptions vertically, and uses a larger indentation for dialogues. Therefore the indentation size can be used to identify dialogues. Since the dialogues may have different indentations in the same script (same to scenes), the indentation

[3] https://imsdb.com/
[4] https://www.personality-database.com/

sizes of the same section type may vary. Our rule assumes the sections with larger indentations compared to the `FADE IN` marker in the same script as dialogues and the others as scenes.

Silver Parses Construction The above rough dialogues/scenes classification introduces a lot of noises. We use the following idea to automatically determine the threshold indentation of dialogues. First, we compute from the rough results the averaged ratio μ of dialogues in a script and its standard variation σ. Second, we keep adding sections with the largest indentation lengths to the set of dialogues, until the ratio of added sections becomes larger than $\mu+\sigma$. Finally, we keep the left sections as scenes, with the exception if none of the indentation length can threshold the ratio of dialogues in the range of $\mu\pm\sigma$, which are left as failure cases. We designated the successfully processed scripts with the dialogues/scene labels as the "silver" set. This consists of 29% of all the scripts.

Section Classifier For the failure scripts from the previous step and the scripts without `FADE IN` markers, we use a learning model to determine whether each section belongs to dialogue or scene, with both the section title and text as inputs. Using 137,042 labeled sections from the silver set, we trained a BERT-based section classifier to relabel all scenes and dialogues in all the scripts. The classifier achieved 99.31% accuracy on a heldout validation set. The outputs are our final parses.

3.3 Personality Collection and Mapping

Finally, we collect human rated MBTI types of movie characters from *The Personality Database*. In total, we collected 28,653 characters. Each character has an id, name, vote count, and voters' agreement on each MBTI dimension. For example, the MBTI profile in Figure 1 has 300 voters, with different agreement rate along each dimension, at the time of our data collection. To ensure the quality of personality voting, we removed character profiles with <3 voters and <60% agreement rate.

We then matched the characters' personality profiles to the scripts, if the name can be softly matched to the dialogue title or the recognized named entities in the scenes (details and example of the final processed data in Appendix A).

Table 1 shows the core statistics of our dataset. The types are mostly balanced, where N/S is the most uneven dimension. We filtered out the data with ≥60% agreement, so some characters do not have all the 4 dimensions. We include more details and examples in Table 4 in the appendix.

	Dimension	Train(%)	Dev(%)	Test(%)
(a)	E/I	45.9/51.8	49.6/49.0	52.6/44.2
	N/S	36.6/60.4	41.8/54.0	41.4/55.0
	T/F	54.7/43.2	45.8/50.8	46.0/52.8
	J/P	46.4/51.3	47.2/51.2	45.6/53.0

		Mean	Min	Max
(b)	# dialogues/character	76.90	0	776
	# words/dialogue	917.74	1	12, 536
	# scenes/character	41.08	0	495
	# words/scene	1,381.47	1	25,457

Table 1: Distribution of two personality types per dimension (a) and core statistics (b) in `Story2Personality`.

	Correct scene	Correct dialogue
Ramakrishna et al. (2017)	85%	93%
Our parser	**97%**	**100%**

Table 2: Comparison of correct parsing results.

4 Dataset Analysis

We conduct human study to verify the advantage of our script parser; then provide the human performance on our dataset.

Script Parsing results We compare with the state-of-the-art open-sourced script parser (Ramakrishna et al., 2017), which employs many human written rules, with a human study. We randomly selected five scene and five dialogue section in 10 common movies, giving 100 snippets for evaluation (40 from the silver set). Then we manually compared the processing results with the original movie scripts. Table 2 shows the result. Our parser outperforms Ramakrishna et al. (2017) with a large margin, while most mistakes from (Ramakrishna et al., 2017) is to recognize scenes as dialogues.

Human performance We take the majority vote of each character as the groundtruth. This gives an averaged 93.54% human accuracy across the four personality dimensions on our test data.

Computing humans' macro-F1 score lacks an analytical form from the agreement scores. Therefore we make an approximation by sampling 3 voters (minimum number of voters in our dataset) for each character and treating them like the predictions of 3 different models. This gives overall >95% scores, much higher than model performance (in Table 3). The raw statistic data of human agreement on each MBTI type can be found in Table 5 in the appendix.

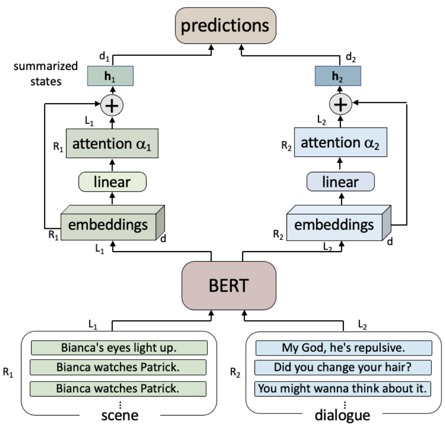

Figure 2: Multi-row multi-view BERT model architecture.

5 Experiments

Baselines We build two baseline models: **SVM**, the LinearSVC from sklearn.svm. We extracted top 20K word unigram, bigram, and trigram features according to term frequency after removing stop words. We set C=0.1; and **BERT**, fine-tuning the out-of-box BERT, with a linear head on the '[CLS]' token's final layer embedding for classification.

Our Method We further propose the multi-view multi-row BERT (**MV-MR BERT**) classifier (Fig. 2), an extension of BERT to deal with the long inputs and handle the verbal and non-verbal information differently, so as to better fit our task.

First, to handle the long input per character, we borrow the idea from fusion-in-decoder (Izacard and Grave, 2020). The key idea is, since the complexity of Transformers is $O(RL^2)$ (with R the number of rows and L the length per row), when L is very large, we can split it to multiple segments to reduce the quadratic term. Then we rely on the attention over all the segments to fuse the information. Specifically, we split the input content $\mathcal{D}$ of a character into multiple segments $\mathcal{D} = \{\mathcal{S}_i\}_{i=1}^R$, and encode all the segments in a minibatch as $\mathbf{H} = \text{BERT}(\mathcal{S}_i) \in \mathbb{R}^{R\times L\times d}$, where d is the hidden state size. Then a linear head is applied to get attention score across tokens in all the rows as $\alpha = \text{softmax}(\mathbf{H}W + b) \in [0,1]^{R\times L}$. The final summarized representation of the input $\mathcal{D}$ is thus the weighted summation $\mathbf{h}_\mathcal{D} = \sum_{i=1}^R \sum_{j=1}^L \alpha_{ij}\mathbf{H}_{ij}$.

Second, to handle both dialogues and behavioral descriptions in the scenes of a character,[5] our multi-view model receives an input pair $(\mathcal{D}^{\text{dial}}, \mathcal{D}^{\text{scene}})$, then uses a shared BERT and separated linear heads to compute the summarized states $\mathbf{h}_\mathcal{D}^{\text{dial}}$ and $\mathbf{h}_\mathcal{D}^{\text{scene}}$. The two vectors are fed into a fully-connected layer for prediction. For the scene description, we prepend a special token "[ent]" to the target character's name to denote its position. The attention α^{scene} is only computed on these special tokens.

[5] Both have been proved useful by psychology studies (McCroskey and Richmond, 1996; Richmond et al., 2008), where sometimes non-verbal behaviors (e.g., stance, gesture, and body movements) even dominate.

Model	E/I	N/S	T/F	J/P
SVM	54.65	55.41	52.83	56.18
BERT	$56.06_{\pm 0.73}$	$55.59_{\pm 3.36}$	$57.13_{\pm 0.97}$	$57.59_{\pm 1.40}$
MV-MR BERT	$\mathbf{57.50_{\pm 2.04}}$	$\mathbf{57.42_{\pm 4.27}}$	$\mathbf{60.33_{\pm 0.93}}$	$\mathbf{59.83_{\pm 1.42}}$
- multiview	$57.30_{\pm 1.91}$	$57.05_{\pm 1.80}$	$57.04_{\pm 2.05}$	$57.39_{\pm 2.21}$
Human Perf.	$98.19_{\pm 0.60}$	$97.82_{\pm 0.10}$	$98.51_{\pm 0.67}$	$98.03_{\pm 0.19}$

Table 3: Macro F1 scores on the four dimensions

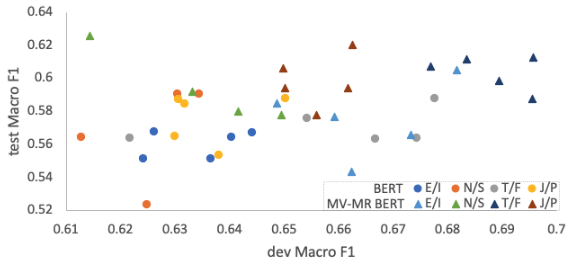

Figure 3: Dev vs. test F1 scores of BERT-based models.

Results and Analysis Following Flekova and Gurevych (2015), we use macro-averaged F1 as evaluation metric. Table 3 shows the main results on the four MBTI dimensions. Peak performance was achieved by our MV-MR BERT. The result suggests using both dialog and action scene descriptions consistently improved model performance.

The results are generally low compared to human performance, showing the task is challenging to existing models. Figure 3 gives further evidence for the challenge of our task, which plots the dev versus test scores during our model selection. It shows the dev and test results are not highly-correlated, meaning that by achieving near perfect accuracy on the training data, the models largely overfit the noises instead of capturing real clues.

6 Conclusion

We present a new narrative understanding task, `Story2Personality`, a large scale personality prediction dataset of movie characters. We evaluate several classifiers on our task – while our multi-view BERT model achieves a substantial improvement over the SVM and BERT baselines, there is a huge gap compared to human performance. This indicates our dataset a valuable and challenging task for future research.

References

Gregory J Boyle. 1995. Myers-briggs type indicator (mbti): some psychometric limitations. *Australian Psychologist*, 30(1):71–74.

David Cohen, Marilye Cohen, and Herbert Cross. 1981. A construct validity study of the myers-briggs type indicator. *Educational and Psychological Measurement*, 41(3):883–891.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

John M Digman. 1997. Higher-order factors of the big five. *Journal of personality and social psychology*, 73(6):1246.

Lucie Flekova and Iryna Gurevych. 2015. Personality profiling of fictional characters using sense-level links between lexical resources. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1805–1816.

Matej Gjurković and Jan Šnajder. 2018. Reddit: A gold mine for personality prediction. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, pages 87–97.

Jennifer Golbeck, Cristina Robles, Michon Edmondson, and Karen Turner. 2011. Predicting personality from twitter. In *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*, pages 149–156. IEEE.

Gautier Izacard and Edouard Grave. 2020. Leveraging passage retrieval with generative models for open domain question answering. *arXiv preprint arXiv:2007.01282*.

Arnav Jhala. 2008. Exploiting structure and conventions of movie scripts for information retrieval and text mining. In *Joint International Conference on Interactive Digital Storytelling*, pages 210–213. Springer.

Carl Jung. 2016. *Psychological types*. Routledge.

Tomáš Kočiskỳ, Jonathan Schwarz, Phil Blunsom, Chris Dyer, Karl Moritz Hermann, Gábor Melis, and Edward Grefenstette. 2018. The narrativeqa reading comprehension challenge. *Transactions of the Association for Computational Linguistics*, 6:317–328.

Yash Kumar Lal, Nathanael Chambers, Raymond Mooney, and Niranjan Balasubramanian. 2021. Tellmewhy: A dataset for answering why-questions in narratives. *arXiv preprint arXiv:2106.06132*.

Gerald Matthews, Ian J Deary, and Martha C Whiteman. 2003. *Personality traits*. Cambridge University Press.

James C McCroskey and Virginia P Richmond. 1996. *Fundamentals of human communication: An interpersonal perspective*. Waveland PressInc.

Xiangyang Mou, Chenghao Yang, Mo Yu, Bingsheng Yao, Xiaoxiao Guo, Saloni Potdar, and Hui Su. 2021. Narrative question answering with cutting-edge open-domain qa techniques: A comprehensive study. *arXiv preprint arXiv:2106.03826*.

Isabel Briggs Myers and Mary H McCaulley. 1988. *Myers-Briggs type indicator: MBTI*. Consulting Psychologists Press Palo Alto.

Lin Qiu, Han Lin, Jonathan Ramsay, and Fang Yang. 2012. You are what you tweet: Personality expression and perception on twitter. *Journal of research in personality*, 46(6):710–718.

Anil Ramakrishna, Victor R Martínez, Nikolaos Malandrakis, Karan Singla, and Shrikanth Narayanan. 2017. Linguistic analysis of differences in portrayal of movie characters. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1669–1678.

Virginia P Richmond, James C McCroskey, and Mark Hickson. 2008. *Nonverbal behavior in interpersonal relations*. Allyn & Bacon.

Mark O Riedl and Robert Michael Young. 2010. Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, 39:217–268.

Christopher Riley. 2009. *The Hollywood standard: the complete and authoritative guide to script format and style*. Michael Wiese Productions.

Matthew Sims, Jong Ho Park, and David Bamman. 2019. Literary event detection. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 3623–3634.

Gary J Sipps and Ralph A Alexander. 1987. The multifactorial nature of extraversion-introversion in the myers-briggs type indicator and eysenck personality inventory. *Educational and Psychological Measurement*, 47(3):543–552.

Charles R Thomas. 1983. Field independence and myers-briggs thinking individuals. *Perceptual and motor skills*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. In *Advances in neural information processing systems*, pages 5998–6008.

Alessandro Vinciarelli and Gelareh Mohammadi. 2014. A survey of personality computing. *IEEE Transactions on Affective Computing*, 5(3):273–291.

A Details of Dataset Construction

Soft Name Matching Algorithm We created two movie-character dictionaries to associate the characters with the movies using the characters' full names and their subcategories (i.e., movie names) in personality profile data, as well as section titles (i.e., character names or scene headings) and movie names in the movie script data. Then, we tokenized and lowercased the character names. We matched both the exact same full names and the intersections of tokens such as the first or last name of the full name when the movie names are matched. To identify a character's scene descriptions, we extracted named entities from scene descriptions and then matched the characters and scenes based on their names using the same method. After matching the character name with the movie name, we store the MBTI personality, vote count, dialogue and scene descriptions into a dictionary for each character.

Example Data Item for One Character Figure 4 shows the example of information for one character *Gary* from the movie "Joker" in our `Story2Personality`, stored in json format. The data item contains the ID ('id'), character name ('mbti_profile') and movie name ('subcategory') in the PDB website; together with the human voted MBTI types and the number of votes. Finally we save the dialogues of the character and the scenes he appears in two separated entries. For scenes, we save both the scene texts and the soft matched name mentions in the texts for the target character. The name mention is used to prepend the special tokens in our MV-MR BERT model.

B Additional Information of the Dataset

Table 4 lists the distribution of all the 16 MBTI types in our dataset, together with a representative movie character for each type.

Personality	%	Example
ISTJ	8.41%	*Darth Vader* ("Star Wars")
ISTP	8.07%	*Shrek* ("Shrek")
ESTP	8.21%	*Han Solo* ("Star Wars")
ESTJ	6.52%	*Boromir* ("The Lord of the Rings")
ISFJ	6.41%	*Forrest Gump* ("Forrest Gump")
ISFP	6.49%	*Harry Potter* ("Harry Potter")
ESFJ	4.88%	*Cher Horowitz* ("Clueless")
ESFP	7.06%	*Jack Dawson* ("Titanic")
INFJ	4.80%	*Edward Cullen* ("Twilight")
INFP	5.42%	*Amélie Poulain* ("Amélie")
ENFP	3.90%	*Anna* ("Frozen")
ENFJ	3.75%	*Judy Hopps* ("Zootopia")
INTJ	4.26%	*Michael Corleone* ("The God Father")
INTP	3.75%	*Neo* ("The Matrix")
ENTP	4.94%	*Tyler Durden* ("Fight Club")
ENTJ	4.88%	*Patrick Bateman* ("American Psycho")

Table 4: Distribution of the 16 MBTI personality types in `Story2Personality`

C Statistics of Human Agreement

Table 5 lists the human agreement score on each MBTI dimension, on which we compute the human accuracy and approximate the human macro-F1 scores.

The raters are most divided in annotation of N, with an average agreement is 91.06% and the standard deviation 0.11. One reason is that the perceptual style dimension N/S measures how the individual obtain information. Comparing with dimensions related to attitudes (E/I) or decision making (T/F, J/P) (Jung, 2016) perceptual style is more implicit. Specifically, S is seen as attending to sensory stimuli, while N describes a more detached, insightful analysis of events and stimuli (Boyle, 1995). They are more difficult to determine from the explicit story narratives.

	Mean	Min	Max	STD	#Character
I	94.43%	60%	100%	0.10	1,783
E	94.22%	60%	100%	0.10	1,679
N	91.06%	60%	100%	0.11	1,347
S	93.32%	60%	100%	0.11	2,082
T	94.22%	60%	100%	0.10	1,851
F	93.68%	60%	100%	0.10	1,617
P	93.68%	60%	100%	0.10	1,825
J	93.72%	60%	100%	0.10	1,644

Table 5: Descriptive statistics of voters' agreement

```
{'id': 62784,
'mbti_profile': 'Gary',
'subcategory': 'joker',
'vote_count_mbti': 35,
'I': 100.0,
'N': nan,
'F': 97.0,
'P': nan,
'E': nan,
'S': 100.0,
'T': nan,
'J': 86.0,
'dialog_text': ['Of course. No problem. Another time.',
'(screaming) What the fuck what the fuck WHAT',
'Hey, Art?',
"Hey Art, I heard what happened-- I'm sorry man.",
'No clue.',
"(re: his look) Hey Arthur, how's it going? You get a new gig?",
"(interrupting) They didn't talk to me.",
"We don't wanna bother you. Randall just thought we should come and pay our respects.",
'Arthur,-- Hoyt wants to see you in his office.',
'(to Randall) Since when do you use a prop gun?'],
'scene_text': [('gary',
"Gary backs away toward the door. Joker sits there for a moment, breathing heavy, wipes Randall's blood off his face--"),
('gary', "Gary doesn't answer. Doesn't move--"),
('gary',
"AND JOKER STABS THE SCISSORS AS DEEP AS HE CAN into Randall's neck. Blood spurts. Randall screams. Gary stumbles back in shock--"),
('gary',
"Joker turns, sees Gary at the front door. He points up high to the chain-lock. He can't reach it. Joker just shakes his head to himself and gets up to unlock the door. He walks past Gary who's still trembling almost too afraid to look up at him. Joker leans over him and undoes the chain, opens the door. Gary bolts, running down the hallway as fast as he can--"),
('gary',
"Joker walks into the locker room, sees Randall half-dressed for work, red nose, big pants, big shoes, no wig yet, sitting with Gary, TWO OTHER CLOWNS AND A MAGICIAN around the small table, shooting the shit, drinking coffee. They nod hello at Joker or give him a perfunctory wave, most of his co-workers think he's a freak."),
('gary',
"Joker pulls them out and jams them into Randall's eye before he can react. The sound is sickening. Gary's screaming in the background-- Randall blindly fights back, screaming in pain, flailing his arms, his own blood blinding him-- Joker grabs Randall by the head -- all of his pent up rage and frustration pouring out of him -- AND SLAMS HIS HEAD"),
('sees gary',
"AGAIN. And AGAIN. And AGAIN. Joker lets go of Randall's head, and Randall drops to the ground. Joker leans back against the wall, out of breath, kind of slides down the wall to the floor-- Sees Gary huddled in the corner, trembling with fear--"),
('gary',
'Gary turns to go. Randall pauses for a moment, has something else to say before he leaves--'),
('gary',
"Joker unlocks the locks, keeping the security chain latched, and cracks open the door,-- Sees Randall. Looks down, and sees Gary next to him. Undoes the chain and opens the door for them-- Randall and Gary get a look at Joker's face, his dyed green hair still wet, streaking white grease-paint smeared over part of his face--")],
}
```


Figure 4: An example except from `Story2Personality`. Each character has ID ('id'), character name ('mbti_profile'), movie name ('subcategory'), dialogue, and scene descriptions.