Perceptions of Social Roles Across Cultures

MeiXing Dong(✉), David Jurgens, Carmen Banea, and Rada Mihalcea

University of Michigan, Ann Arbor, MI, USA
{meixingd,jurgens,carmennb,mihalcea}@umich.edu

Abstract. In this paper we introduce a data set of social roles and their aspects (descriptors or actions) as emerging from surveys conducted across a sample of over 400 respondents from two different cultures: US and India. The responses show that there are indeed differences of role perceptions across the cultures, with actions showcasing less variability, and descriptors exhibiting stronger differences. In addition, we notice strong shifts in sentiment and emotions across the cultures. We further present a pilot study in predicting social roles based on attributes by leveraging dependency-based corpus statistics and embedding models. Our evaluations show that models trained on the same culture as the test set are better predictors of social role ranking.

Keywords: Social role perceptions · Cultural differences · Word associations · Cultural biases · Natural language processing · Word representations

1 Introduction

Beliefs we hold about the world often manifest themselves in the way we use language. Understanding what people say or write can help us gain insight into their worldview, beliefs, and the way they are primed to interact with the surrounding world. Such analyses of language can also lead to new insights into cultural differences. Groups of people sharing certain characteristics – e.g., nationality, region, state, gender, or religion – would often have a shared understanding of the world, which in turn is reflected in their use of language.

While the connections between language and culture have traditionally been the purview of cultural psychology [42], more recent work in computational linguistics has also started to address these connections, resulting in models that can uncover the different use of words across cultures [22,24], the various distribution of topics in different cultures [34], or the word associations that people with different demographics tend to make [21,25].

The hypothesis driving our work is that we can use language to identify and understand the implicit perceptions and expectations that people hold with regards to social roles in our society. For instance, the frequent use of the descriptor *kind* or the action *help* in connection to the role *friend* can be an indication

I. Weber et al. (Eds.): SocInfo 2019, LNCS 11864, pp. 157–172, 2019.
https://doi.org/10.1007/978-3-030-34971-4_11

that friends are usually regarded as people who are kind and provide help. Moreover, we also hypothesize that there may be cultural differences in these social role perceptions, and that different groups of people may correspondingly use different descriptors or actions when they refer to the same social role.

This paper makes four main contributions. First, we examine what constitutes a social role, and we propose the use of descriptors (adjectives) and actions (verbs) as a way to understand the implicit perception of social roles as reflected in language. Second, we introduce a new data set, consisting of 49 frequent social roles (e.g., *mother*, *friend*, *lawyer*) and the associated descriptors and actions, as contributed by over 400 human judges from two different cultures (United States and India). Third, we perform several analyses to uncover cross-cultural variations in social role perception, and we identify roles with high, medium, and low variations. Finally, we propose two computational models that can predict the most likely social role based on a descriptor or an action. One model is based on statistics collected over a large syntactically annotated collection of texts authored by people from two cultures, while the second one relies on neural models that are aware of the syntactic relations between words.

Our main findings show that there are indeed differences in the perceptions associated with the roles between the two cultures, and that the degree of cultural similarity varies across the roles. The computational models show that it is possible to predict roles from the attributes that people associate with them. Furthermore, our models exhibit higher performance when the train and test set cultures match, indicating that our models encode cultural differences.

2 Related Work

The concept of "roles" is frequently considered by those in the social sciences as a way to analyze social structures and behaviors [2,7,18,26,38]. Roles can be characterized by the norms and expectations that society places on people of particular social or functional positions [26]. Such norms greatly influence how people act and interact with others [12], especially when one is acting as a member of a role [43]. The perceptions of others are important; depending on whether one acts according to role expectations, there exist rewards or punishments doled out by society [18]. By asking members of a group about the behaviors that a role is likely to participate in, one can analyze the differences in perceptions of roles between cultural groups, such as Hispanics versus the general US population [44].

We take inspiration from previous work that models latent character types, or personas (such as the "love interest" or "best friend") and their typical characteristics in films [4]. To extract character aspects, the authors look at a subset of the syntactic dependencies that involve the personas. We extract aspects in a similar way and focus on predicting a role based on its expected characteristics, in contrast to Bamman et al. that focus on partitioning types of roles. Additionally, films tend to create stereotypical personas with strong associations to their characteristics. Social roles, however, are constructed from societal expectations in aggregate and can be much more nuanced.

Another related line of research has considered the prediction of words that are most likely to be associated with a stimulus word [21]. Our task differs in that we go beyond free-form associations and instead hone in on specific aspect types, namely actions and descriptors as they relate to a given social role.

To use natural language, we must build word representations. A straightforward approach is to treat words as discrete symbols, leading to many bag-of-words methods for representing text [41,46]. While useful for many tasks, this representation does not encode relations between words or semantics. Many recent word representation methods model words as continuous, dense vectors derived from neural networks [5,30] or word co-occurrence information [35], also known as word embeddings. These have been shown to perform well across numerous tasks [39]. Additionally, [1,3,10,20,21] have sought to encode additional sources of information to be captured in word embedding vectors.

One of our models is derived from dependency-based embedding models [28], where dependency links are used to form the contexts in a skip-gram model. The resulting embeddings encode functional similarity rather than topical similarity. For instance, *rapping*, *busking*, and *breakdancing* are among the most similar words for "dancing" when using dependency-based embeddings, as opposed to topically related words surfaced by regular linear context embeddings, such as "dancer", "dance", and "dances." We adapt the former model to focus on specific types of dependencies that encode aspects, distinguishing between the different functional uses of a word. For example, we can find roles that are most relevant to a given aspect, rather than the words that are generally related either by domain or by function.

3 Collecting a Cross-Cultural Data Set of Social Roles

The perception of a social role can be characterized by the descriptors or actions that people associate with it. We created a data set by surveying a large and demographically diverse audience on Amazon Mechanical Turk (AMT) about the aspects they associate with different roles. Our survey task is similar to that of gathering word associations, where survey participants are provided with a list of *stimulus words* and are asked to provide the first word that comes to their mind [21,27,33]. However, rather than asking for free-form associations, as done before, we added structure to our prompts to induce responses that correspond to descriptive aspects. Specifically, we asked survey participants to provide actions and descriptors for each stimulus role, given prompts such as *What is a friend like?* and *What does a friend do?*

Selecting Social Roles. Language abounds with the names of the many social roles that people partake in, from common names (like mother or teacher) to less common ones (like debtor or occultist). Here, we aim to curate a set of social roles for annotation that meet three criteria: (1) occur with high frequency in text, (2) appear in daily life, and (3) have relatively unambiguous words associated with them. We detail the selection process next.

A large set of candidate social roles were selected using WordNet [31], a large lexical database for English. WordNet provides an ontological organization of a word's meanings and contains a semantic network of how these meanings (i.e., *senses*) relate to one another. In particular, WordNet specifies the hyponymy relationship between senses that allows us to identify more specific meanings of people; for example, *mother* and *father* are both hyponyms of *parent*. To get all potential social roles, we collected the 8,654 words that are children of *person* in the hyponymy tree.

As WordNet contains many infrequent words, we extracted frequency counts for each role from a large collection of blog data from India and the US, described in detail in Sect. 5.1. We tagged each blog sentence with part-of-speech information, and then counted the frequency of each candidate role occurring as a noun.

Finally, we analyzed the most frequently occurring candidate roles and identified roles that occurred in blogs from both countries, that are generally unambiguous, and are likely to be encountered in day-to-day life. For instance, we did not include *queen* because most people are unlikely to interact with queens, and therefore descriptors and actions are unlikely to reflect personal experiences. We also excluded ambiguous roles such as *official* or *director*, since their attributes can change depending on the context. Ultimately, the selection process resulted in a set of 49 social roles.

Crowdsourcing Setup. The descriptors and actions for each social role were collected through AMT English surveys[1], targeted to individuals in India and the US. We chose countries that were likely to differ in terms of cultural and societal norms, but still have many English speakers to bypass translation issues. Each participant was presented with five social roles and asked to provide three actions and three descriptors for each role. Participants were also asked to indicate how often they interact with the role and how positively they view those interactions. A demographic questionnaire was included at the end of the each survey containing questions about the respondent's gender, age, level of education, ethnicity, and nationality. Responses were collected from 200 participants from each country for each role. This resulted in 600 actions and 600 descriptors collected for each social role, for each country.

To ensure answer quality, we included a spam-check question that asked for the answer to an earlier question. This filtered out participants that responded without reading the prompts. Built-in form restrictions prevented the submission of answers that were given as examples, or empty answers. As a final check, we manually spot-checked responses before accepting them, to make sure participants did not fill in random words. We lemmatized all of the responses and for each given social role we kept those responses that occurred five times or more as culturally-salient aspects of the role.

Previous studies [6,13] have shown that while Turkers tend to be younger and more educated, it is possible for the data they supply to reflect aspects of the population at large, such as ideology. The data we gathered serves as an

[1] English is one of the official languages of India and the second most-spoken language behind Hindi.

Table 1. Top survey responses for societal role words.

Role word	Actions		Descriptors	
	US	India	US	India
Mother	Care, love, cook	care, love, cook	Loving, caring, nurturing	Caring, lovable, loving
Baby	Cry, sleep, eat	Cry, play, smile	Loving, sweet, kind	Cute, innocent, chubby
Doctor	Diagnose, prescribe, examine	Treat, care, cure	Smart, intelligent, helpful	Caring, god, helpful
Policeman	Protect, arrest, serve	Arrest, protect, help	Strong, brave, helpful	Strict, brave, strong
Student	Study, learn, read, write, work	Study, play, learn, read, write	Studious, smart, young	Obedient, intelligent, studious
Politician	Lie, campaign, speak, talk, cheat	Speak, vote, lead, promise, rule	Dishonest, greedy, corrupt	Powerful, honest, influential

additional resource to complement existing cross-cultural resources, providing insight into cultural differences pertaining to how social roles are perceived. Despite the potential skew in demographics, we still find differences between the two countries, as detailed in later sections.

Table 1 shows the top responses for a sample set of social roles.

4 Demographic Variations in Social Roles

The characteristics associated with social roles in different countries can reveal cultural similarities and differences. Many aspects are associated with a role regardless of the underlying culture, such as a *mother* being *caring* and a *policeman* being *brave*. On the other hand, *doctors* are more associated with preliminary actions in the treatment process such as *examine*, *diagnose* and *prescribe* in the US, while in India they are more associated with treatment results, such as *treat* and *cure*. Also, Indian descriptors show a stronger perception of doctors as being *caring*, versus *smart* and *intelligent* in the US. Additionally, US participants associate many negative aspects with *politician*, reflecting the current political climate, while in India, the actions are mostly associated with positive aspects.

Intra-group and Inter-group Similarities. We measure the agreement between respondents within and across cultural groups. Given the set of response words for a social role from a single held-out respondent, we determine whether any of these responses match the most frequent response or any of the top 25 responses of the remaining respondents in the group. If so, then we consider this respondent in agreement with the group. We define the agreement score as the ratio of participants whose responses are in agreement with the group. Similarly, we measure the agreement between each survey respondent in one group with the most frequent or top 25 most frequent responses from the other group.

The intra-group and inter-group analyses are shown in Table 2. From the intra-group similarities, we can see that there is high agreement among both the top and top 25 responses given by participants from the same country, with the US having higher agreement in general than India. Overall, action responses are more cohesive across the two countries compared to descriptor answers; we noted earlier that there is more variation and subjectivity in regards to descriptors.

Table 2. Left: intra-group similarities (higher similarity indicates a more cohesive group). Right: inter-group similarities (higher similarity indicates a less distinct group).

Intra-group similarity			Inter-group similarity		
Demographic	Primary	Top 25	Demographic	Primary	Top 25
US-US (Descriptors)	0.33	0.89	US-India (Descriptors)	0.19	0.78
India-India (Descriptors)	0.24	0.76	India-US (Descriptors)	0.15	0.61
US-US (Actions)	0.40	0.93	US-India (Actions)	0.35	0.90
India-India (Actions)	0.40	0.89	India-US (Actions)	0.32	0.85

When we look at how much participants from one country agree with participants from *the other country*, we find a much lower agreement for descriptors, both in terms of primary response and the top 25 responses. For example, the similarity drops by 0.08 (from 0.40 to 0.32) between India-India and India-US for the most frequent response for actions. We see the agreement drop in all cases when comparing intra- versus inter-group similarity. We conclude that the agreement for actions between the countries is comparable to the agreement within countries, implying that the actions attributed to roles are more objective and universal.

Levels of Social Role Similarity Across Cultures. We closely examine how various roles are perceived differently across countries. To measure how similar a role is between India and the US, we compute the cosine similarity between the frequencies pertaining to the set of aspects resulting from the union of the responses for that role for each country. Table 3 shows a sample of roles that display various levels of similarity ranging from high to low in regards to their associated actions or descriptors across the countries in question. We notice that *soldier* exhibits the highest similarity level both for actions (*fight, protect*) and descriptors (*brave, strong*). *Actor*, on the other hand, showcases a medium action-based similarity, as actors in India regularly engage in dancing, unlike their US counterparts. Interestingly, *friend*, despite its ubiquitousness as a social role, displays among the lowest scoring action-based similarity. We note that in the US, *friend* is more associated with communication-focused actions such as *listen*, *talk*, *laugh*, while in India, given the more collectivist culture, people primarily think of friends in the context of being *helpful* and *caring*.

Levels of Aspect Similarity Across Cultures. We further analyze the frequency of aspect usage across roles to identify how predictive a given aspect is of a social role. Table 4 aggregates the responses at the aspect level. We see that some actions are highly predictive of a role. For instance, *arrest* occurs with *police* and *vote* appears with *citizen, politician* roles in both countries. However, *sacrifice* occurs in the action-focused answers for *soldier* in the US, while in India, *mother* and *father* also trigger this response. *Counsel* also displays a divergent usage, in the US being associated with *lawyer*, while in India, with *priest*. Similarly, descriptors also show variations in their associations with roles. These range from a high similarity of 1 for *religious* (which always appears in the context of *priest*), to mid-range (0.46) for *obedient* (which in the US caries a stronger meaning of loyal, and applies to a hierarchical organization, e.g. army for *soldier* or country for *citizen*, while in India it is more indicative of filial piety and the need to listen to one's elders, whether as a *student*, *son*, or *daughter*), to low (0) for *committed* (which in the US occurs in prompts for *wife* and *husband*, while in India it appears in prompts for *farmer*).

Sentiment and Emotion in Social Role Perceptions. Social roles can evoke a variety of emotional responses, such as feelings of authority, love, or even fear. Viewed in aggregate, the aspects used to describe roles can potentially reveal which emotional aspects of social roles are most important to a culture. Therefore, we perform two analyses where we convert the actions and descriptions for each role into their sentiment and emotion associations. For sentiment, we map each word to its mean score in SentiWordNet [19] and then average across all the words for each aspect of a role for its estimated sentiment score. For emotion, we repeat a similar process with the NRC Emotion Lexicon [32]. This lexicon maps individual words to a binary indicator of whether they have an association with each of the eight Plutchik emotions [36]. Here, we compute the probability that an aspect word for a role has an association with each emotion. We then average the sentiment and emotion-association probabilities across all roles.

For sentiment, Fig. 1 shows clear differences between India and the US responses, with AMT workers from India using significantly more positive descriptors about roles. No significant difference is seen for actions, though this is expected, as adjectives (descriptors) typically carry more sentiment than verbs (actions); for example, common sentiment lexicons like SentiWordNet [19] and OpinionFinder [47] contain more adjectives than verbs, and adjectives have been shown to outperform verbs as features in sentiment classification [48]. Examining AMT workers' explicit sentiment ratings for roles, we see that their ratings have high correlation with the inferred sentiment, with Pearson's r ranging from 0.51 to 0.61. This result suggests that the inferred ratings are capturing representative attitudes but, crucially, that roles' aspect words convey more than the workers' sentiment about the role.

Table 3. Social roles exhibiting different levels of similarity (H(igh), M(edium), L(ow)) between the US and India based on the differences between the top 20 responses.

Similarity	Role	Score	Top US aspects	Top IN aspects
Actions				
H	Soldier	0.94	Fight, protect, defend	Fight, protect, shoot
	Professor	0.91	Teach, grade, lecture	Teach, guide, educate
	Mother	0.89	Care, love, cook	Care, love, cook
M	Girlfriend	0.77	Love, kiss, listen	Love, care, help
	Policeman	0.77	Protect, arrest, serve	Arrest, protect, help
	Actor	0.71	Act, perform, pretend	Act, dance, perform
L	Doctor	0.66	Diagnose, prescribe, examine	Treat, care, cure
	Politician	0.59	Lie, campaign, speak	Speak, vote, lead
	Friend	0.58	Listen, talk, laugh	Help, care, play
Descriptors				
H	Soldier	0.89	Brave, strong, loyal	Brave, strong, patriotic
	Writer	0.88	Creative, imaginative, smart	Creative, imaginative, good
	Mother	0.72	Caring, loving, nurturing	Caring, lovable, kind
M	Researcher	0.62	Smart, intelligent, curious	Intelligent, knowledgeable, brilliant
	Prisoner	0.60	Angry, sad, guilty	Bad, criminal, guilty
	Friend	0.59	Fun, loyal, funny	Helpful, caring, honest
L	Farmer	0.44	Strong, hardworking, diligent	Hardworking, poor, helpless
	Judge	0.38	Fair, powerful, smart	Honest, intelligent, knowledgeable
	Politician	0.17	Dishonest, greedy, corrupt	Powerful, honest, influential

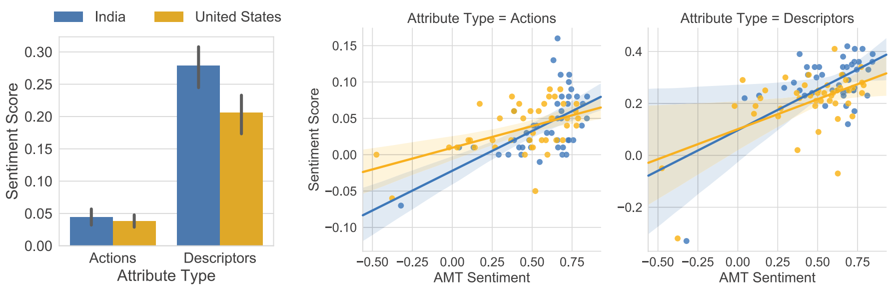

Fig. 1. India and the US differ significantly in the sentiments of roles attributes (left); indeed, AMT workers' explicit sentiment ratings for each role were highly correlated with inferred sentiments of their descriptors and actions (right). Bars and shaded regions show 95% confidence intervals.

The emotion trends, shown in Fig. 2, reveal a more complex picture with Indian respondents being more likely to use emotionally-associated language than their US counterparts. This heightened emotionality occurs both for positive emotions like trust, surprise, and anticipation, as well as negative emotions like disgust and sadness. However, US respondents are more likely to evoke anger or fear; yet, these emotions are the two least-frequently used in our data. While cross-cultural studies of emotion have shown differences between India and the

Table 4. Aspects exhibiting different levels of similarity (H(igh), M(edium), L(ow)) between the US and India based on the differences between the top 20 responses.

Similarity	Aspect	Score	Top US roles	Top IN roles
Actions				
H	Arrest	1.0	Police	Police
	Vote	0.99	Citizen, politician	Citizen, politician
	kiss	0.93	Girlfriend, boyfriend, mother	Girlfriend, boyfriend, husband
M	Medicate	0.71	Nurse	Doctor, nurse
	Sacrifice	0.61	Soldier	Father, mother, soldier
	Forgive	0.51	Priest	Friend, priest, mother
L	Invent	0.15	Engineer, chef, writer	Scientist, researcher, engineer
	Meditate	0.0	—	Priest
	Counsel	0.0	Lawyer	Priest
Descriptors				
H	Religious	1.0	Priest	Priest
	Loving	0.96	Mother, husband, wife	Mother, husband, sister
	Curious	0.91	Tourist, researcher, journalist	Journalist, tourist, researcher
M	wise	0.71	Father, priest, professor	Professor, teacher, judge
	Loyal	0.65	Friend, husband, wife	Friend, citizen, soldier
	Obedient	0.46	Son, soldier, citizen	Student, son, daughter
L	Faithful	0.15	Wife, husband	Priest, secretary, chef
	Jovial	0.0	—	Politician
	Committed	0.0	Wife, husband	Farmer

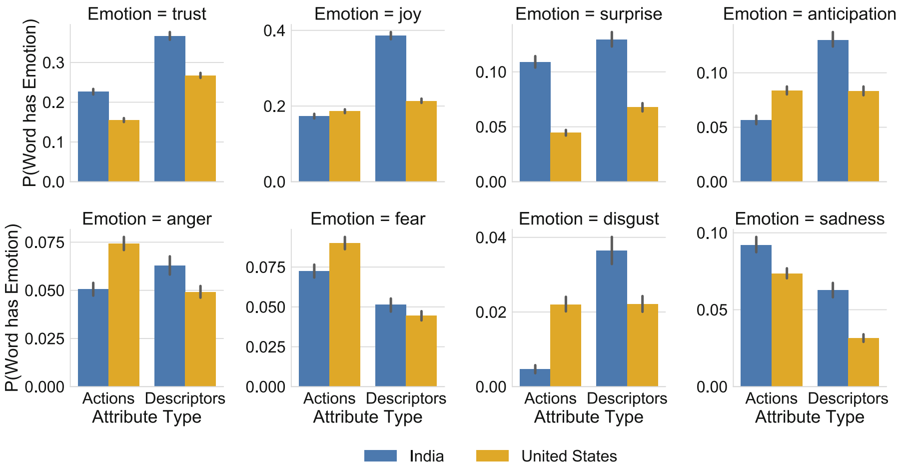

Fig. 2. AMT answers' emotions. The emotionality of actions and descriptors across social roles shows clear cultural differences in how each is conceived; plots show the probability of an action or descriptor using a word associated with each emotion in the NRC Emotion Lexicon, with bars showing 95% confidence intervals of mean probability.

US [15,37,40], these studies have typically looked at specific settings such as childhood development, rather than general attitudes; our data set provides a valuable new source of comparison.

5 Pilot Evaluation

We conduct two initial experiments to gauge how demographic-aware roles can be predicted from descriptors and actions using textual data. We evaluate our models on the task of predicting the most likely social roles for a given aspect. For example, if we think of *lovely,* we want to identify the roles in each culture that are most strongly associated with this descriptor. This enables us to underscore the particularities of each culture, where some roles are associated with softer traits, while others with stronger traits.

We take the top 20 aspects for each role from the survey responses of both countries and combine them into a set of descriptors and a set of actions. Table 5 shows statistics pertaining to these aspects. For each aspect, the set of expected roles (i.e., ground truth) are the ones for which the aspect appears in the top 20 responses. Evaluations are conducted on each aspect type separately. Our models rank the roles for each aspect, which we compare with the expected roles. We report the precision and recall at 5, averaged over the aspects. We also report Pearson correlation (which is typical in word similarity tasks [11,16]), as this gauges how accurate the model is in arranging the roles in order of association with the given aspect.

5.1 Blogger Data Set

To train our models, we need to employ text written by authors whose demographic location is known. For that, we use a set of blogs written between 1999 and 2016 collected from Google Blogger [22]. We select those blogs that also contain location information and only consider those with authors in India or the US. This allows us to analyze the cultural differences between India and the US that may appear as differences in the meaning and usage of roles. We filter out sentences with more than 150 words[2] or with more than 25% non-English words[3]. The remaining sentences are cleaned from HTML tags and truncated to the first 50 words. We then use the Stanford CoreNLP library [29] to obtain dependency parses for the sentences in our data set, as well as part-of-speech tags and lemmatized versions of the tokens. Only lemmas that occur 5 or more times are considered.

Because the US blog data is roughly twice the size of the Indian blog data, we balance the data by downsampling the US data to match the number of sentences in the Indian data. Table 6 provides statistics of the resulting data set.

[2] Normal sentences are rarely this long, and upon manual inspection we found that these tend to be malformed sentences.

[3] https://github.com/rfk/pyenchant.

Table 5. Statistics for unique aspect words given by survey responses.

Type	US-only	India-only	Both	**All**
Action	126	76	199	**401**
Descriptor	154	136	156	**446**

Table 6. Blog data statistics.

	# Sentences	# Tokens
US	17,476,527	348,479,631
US (balanced)	7,394,484	146,347,629
India	7,426,583	148,710,411

5.2 Computational Models to Predict Social Roles

We propose two computational models to predict social roles. The first model focuses on corpus statistics using dependency link counts, while the second uses neural-network dependency-based word embeddings.

Dependency Link Count (DLC). We first look for the actions and descriptors that engage in a syntactic relation with a role in a sentence as a way of modeling the way people associate roles and their aspects. In order to extract role-aspect relations, we leverage dependency parsing information.

Let us consider the following example: *"The attentive policeman arrested the perpetrators."* The dependency parse results in the following relations (the relations in which *policeman* appears in are italicized): *det* (policeman-3, The-1), *amod* (policeman-3, attentive-2), *nsubj* (arrested-4, policeman-3), root (ROOT-0, arrested-4), det (perpetrators-6, the-5), dobj (arrested-4, perpetrators-6).

The first three relations showcase scenarios where words appear with our target role *policeman*. Since we are interested in finding descriptors for the target role, we focus on the *AMOD* relationship (or adjectival modifier) where the role appears in the source position in the dependency relation; to identify associated actions for the target role, we utilize the *NSUBJ* relations (or nominal subject) where the role appears in the target position in the dependency relation. Consequently, *attentive* is marked as participating in an AMOD relation, while arrested participates in a NSUBJ relation, corresponding to descriptors and actions, respectively.

Word co-occurrence is used extensively to model relationships between words [8,9,17]. Therefore, we rank the roles for an aspect according to how frequently they co-occur in the link type corresponding to the aspect type.

Dependency Aspect Embedding (DAE). Neural word embeddings have proven useful for a large variety of tasks [14,45,49]. Here, we make use of their representation power, but aim to capture demographic-focused embeddings for social role aspects in particular. Previous work has shown that dependency-based word embeddings induce different word similarities [28], yielding more functional similarities, rather than topical similarities.

Rather than training our embedding models on all dependency links, we consider only the links that correspond to descriptors and actions. We train separate models for the two different link types for each of the two countries. This yields four models for: actions in the US, descriptors in the US, actions in India, and descriptors in India. We use the Python *Gensim* Word2Vec library and use 300 latent dimensions with negative sampling.

Table 7. Role prediction results for *actions* (left) and *descriptors* (right). Metrics: precision at 5, recall at 5 and Pearson correlation.

		Actions						Descriptors					
		US			India			US			India		
Model		P@5	R@5	Corr.	P@5	R@5	Corr.	P@5	R@5	Corr.	P@5	R@5	Corr.
DLC	US	**0.13**	**0.26**	**0.15**	**0.10**	**0.19**	**0.11**	**0.12**	**0.20**	**0.13**	0.09	0.14	0.07
	India	0.12	0.24	0.14	**0.10**	0.18	**0.11**	0.11	**0.20**	0.11	**0.10**	**0.16**	**0.09**
DAE	US	**0.12**	**0.28**	**0.14**	**0.11**	0.24	0.12	**0.09**	**0.19**	**0.09**	0.08	**0.18**	0.08
	India	0.11	0.25	0.12	**0.11**	**0.25**	**0.13**	**0.09**	0.17	**0.09**	**0.09**	**0.18**	**0.09**

For a given aspect (a descriptor or action), we compute the cosine similarity between the embedding pertaining to the aspect and the embedding pertaining to each of the social roles. The roles are then ranked according to their cosine similarity, where higher values imply a greater likelihood that the aspect is associated with the role.

5.3 Evaluations and Discussions

Our experiments analyze models that are widely assumed to capture social information [23] and we test the degree to which they are able to do so on a data set designed with this information in mind. The results for the role prediction task are provided in Table 7[4]. The columns represent the source countries of the survey responses, used as the gold standard data for evaluation, while the rows indicate the country of the blogs on which the models were trained. Bold values represent the best performance, when comparing between model countries, for a given combination of model type and gold standard evaluation.

The DAE model is able to achieve a higher recall for actions than the DLC model, but otherwise the two perform comparably. Notably, these two models achieve equal or better performance when the country of the gold standard responses matches the one on which the model was trained. This implies that these models are picking up the distinctive cultural features of the countries.

The gap between identical models trained on different countries is more pronounced when evaluating on the gold standard US responses than on the Indian responses. As English is not the primary language used by Indians, online users may implicitly be conforming to Western societal norms.

We also noticed that implicit or common sense aspect assumptions, while appearing in primary positions in AMT responses, were less likely to appear in the blog data, and sometime did not occur at all. For instance, *faithful* is a top AMT descriptor for *wife* and *husband*, but occurs very infrequently in the blog text. We also see this for *educated* with *professor* and *creative* with *musician.* Blog data often contained aspects that were actually antonyms of the actions

[4] Results for word association tasks are traditionally low, and our results are within the same range as previous word association research [21].

and descriptors provided as answers by the respondents. For example, *corrupt* is among the most frequent descriptors linked to *policeman*, as is *estranged* with *wife* and *unwed* with *mother*. This shows that commonsense knowledge is often not expressed in text, as humans tend not to state the obvious. Consequently, in the blog genre, one tends to express anomalous behavior as it pertains to roles.

6 Conclusion

In this paper we introduced a new data set of social roles and the associations they trigger in terms of actions and descriptors in two cultures (US and India). We showed that there are differences in the perceptions associated with the roles, with actions showcasing less variability and descriptors exhibiting a wider variation. Furthermore we analyzed the way roles are associated with various sentiment and emotional dimensions. We further used the data set we collected to conduct pilot evaluations focused on predicting social roles. Both our corpus-statistics and embedding dependency-based models show a stronger predictive ability when the train and test set culture match, indicating that there are indeed cultural differences that can be automatically accounted for in our models. In the future, we aim to devise more refined predictive models that could be used in a variety of settings, from predicting roles, to predicting aspects, to ultimately tackling model bias. The dataset introduced in this paper is publicly available at http://lit.eecs.umich.edu/downloads.html.

Acknowledgments. This material is based in part upon work supported by the Michigan Institute for Data Science, by the National Science Foundation (grant #1815291), and by the John Templeton Foundation (grant #61156). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the Michigan Institute for Data Science, the National Science Foundation, or the John Templeton Foundation.

References

1. Andrews, M., Vigliocco, G., Vinson, D.: Integrating experiential and distributional data to learn semantic representations. Psychol. Rev. **116**(3), 463 (2009)
2. Ashforth, B.E., Kreiner, G.E., Fugate, M.: All in a day's work: boundaries and micro role transitions. Acad. Manag. Rev. **25**(3), 472–491 (2000)
3. Bamman, D., Dyer, C., Smith, N.A.: Distributed representations of geographically situated language. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 828–834 (2014). http://www.aclweb.org/anthology/P/P14/P14-2134
4. Bamman, D., O'Connor, B., Smith, N.A.: Learning latent personas of film characters, pp. 352–361. http://www.aclweb.org/anthology/P13-1035
5. Bengio, Y., Ducharme, R., Vincent, P., Jauvin, C.: A neural probabilistic language model. J. Mach. Learn. Res. **3**(Feb), 1137–1155 (2003)
6. Berinsky, A.J., Huber, G.A., Lenz, G.S.: Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. Political Anal. **20**(3), 351–368 (2012)

7. Biddle, B.J.: Recent developments in role theory. Ann. Rev. Sociol. **12**(1), 67–92 (1986)
8. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
9. Brown, P.F., Desouza, P.V., Mercer, R.L., Pietra, V.J.D., Lai, J.C.: Class-based n-gram models of natural language. Computat. Linguist. **18**(4), 467–479 (1992)
10. Bruni, E., Boleda, G., Baroni, M., Tran, N.K.: Distributional semantics in technicolor. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1, pp. 136–145. Association for Computational Linguistics (2012)
11. Chaudhari, D.L., Damani, O.P., Laxman, S.: Lexical co-occurrence, statistical significance, and word association. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 1058–1068. Association for Computational Linguistics (2011)
12. Cialdini, R.B., Kallgren, C.A., Reno, R.R.: A focus theory of normative conduct: a theoretical refinement and reevaluation of the role of norms in human behavior (1991)
13. Clifford, S., Jewell, R.M., Waggoner, P.D.: Are samples drawn from mechanical turk valid for research on political ideology? Res. Polit. **2**(4), 2053168015622072 (2015)
14. Collobert, R., Weston, J.: A unified architecture for natural language processing: deep neural networks with multitask learning. In: Proceedings of the 25th International Conference on Machine Learning, pp. 160–167. ACM (2008)
15. Daga, S.S., Raval, V.V., Raj, S.P.: Maternal meta-emotion and child socioemotional functioning in immigrant Indian and white American families. Asian Am. J. Psychol. **6**(3), 233 (2015)
16. De Deyne, S., Navarro, D.J., Storms, G.: Better explanations of lexical and semantic cognition using networks derived from continued rather than single-word associations. Behav. Res. Methods **45**(2), 480–498 (2013)
17. Deerwester, S., Dumais, S.T., Furnas, G.W., Landauer, T.K., Harshman, R.: Indexing by latent semantic analysis. J. Am. Soc. Inf. Sci. **41**(6), 391–407 (1990)
18. Eagly, A.H., Karau, S.J.: Role congruity theory of prejudice toward female leaders. Psychol. Rev. **109**(3), 573 (2002)
19. Esuli, A., Sebastiani, F.: SentiWordNet: a publicly available lexical resource for opinion mining. In: Proceedings of the 5th Conference on Language Resources and Evaluation (LREC 2006), Genova, IT (2006)
20. Feng, Y., Lapata, M.: Visual information in semantic representation. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, pp. 91–99. Association for Computational Linguistics (2010)
21. Garimella, A., Banea, C., Mihalcea, R.: Demographic-aware word associations. In: Proceedings of the International Conference on Empirical Methods in Natural Language Processing (EMNLP 2017), Copenhagen, Denmark (2017)
22. Garimella, A., Mihalcea, R., Pennebaker, J.: Identifying cross-cultural differences in word usage. In: Proceedings of the International Conference on Computational Linguistics (COLING 2016), Japan (2016)
23. Gupta, A., Boleda, G., Baroni, M., Padó, S.: Distributional vectors encode referential attributes. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 12–21. Association for Computational Linguistics, Lisbon, September 2015. https://doi.org/10.18653/v1/D15-1002. https://www.aclweb.org/anthology/D15-1002

24. Hovy, D., Purschke, C.: Capturing regional variation with distributed place representations and geographic retrofitting. In: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, pp. 4383–4394 (2018)
25. Jurgens, D., Tsvetkov, Y., Jurafsky, D.: Writer profiling without the writer's text. In: Ciampaglia, G.L., Mashhadi, A., Yasseri, T. (eds.) SocInfo 2017. LNCS, vol. 10540, pp. 537–558. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67256-4_43
26. Katz, D., Kahn, R.L.: The Social Psychology of Organizations, vol. 2. Wiley, New York (1978)
27. Kent, G.H., Rosanoff, A.J.: A study of association in insanity. Am. J. Psychiatry **67**(1), 37–96 (1910)
28. Levy, O., Goldberg, Y.: Dependency-based word embeddings, pp. 302–308 (2014). http://www.aclweb.org/anthology/P14-2050
29. Manning, C.D., Surdeanu, M., Bauer, J., Finkel, J., Bethard, S.J., McClosky, D.: The Stanford CoreNLP natural language processing toolkit. In: Association for Computational Linguistics (ACL) System Demonstrations, pp. 55–60 (2014). http://www.aclweb.org/anthology/P/P14/P14-5010
30. Mikolov, T., Yih, W., Zweig, G.: Linguistic regularities in continuous space word representations. In: NAACL HLT, Atlanta, GA, USA, pp. 746–751 (2013)
31. Miller, G.A.: WordNet: a lexical database for English. Commun. Assoc. Comput. Mach. **38**(11), 39–41 (1995)
32. Mohammad, S.M., Turney, P.D.: Crowdsourcing a word-emotion association lexicon. Comput. Intell. **29**(3), 436–465 (2013)
33. Nelson, D.L., McEvoy, C.L., Schreiber, T.A.: The university of South Florida free association, rhyme, and word fragment norms. Behav. Res. Methods Instrum. Comput. **36**(3), 402–407 (2004)
34. Paul, M., Girju, R.: Cross-cultural analysis of blogs and forums with mixed-collection topic models. In: Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, Singapore, pp. 1408–1417, August 2009. http://www.aclweb.org/anthology/D/D09/D09-1146
35. Pennington, J., Socher, R., Manning, C.: Glove: global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp. 1532–1543 (2014)
36. Plutchik, R.: The Emotions. Random House, New York (1962)
37. Raval, V.V., Raval, P.H., Salvina, J.M., Wilson, S.L., Writer, S.: Mothers' socialization of children's emotion in india and the usa: a cross-and within-culture comparison. Soc. Dev. **22**(3), 467–484 (2013)
38. Ritzer, G., et al.: The McDonaldization of Society. Pine Forge Press, Newbury Park (1992)
39. Rogers, A., Hosur Ananthakrishna, S., Rumshisky, A.: What's in your embedding, and how it predicts task performance. In: Proceedings of the 27th International Conference on Computational Linguistics, Santa Fe, NM, USA, pp. 2690–2703, August 2018
40. Roseman, I.J., Dhawan, N., Rettek, S.I., Naidu, R., Thapa, K.: Cultural differences and cross-cultural similarities in appraisals and emotional responses. J. Cross Cult. Psychol. **26**(1), 23–38 (1995)
41. Salton, G., Lesk, M.: Computer evaluation of indexing and text processing. J. ACM **15**(1), 8–36 (1968). https://doi.org/10.1145/321439.321441. http://portal.acm.org/citation.cfm?doid=321439.321441
42. Shweder, R.A.: Thinking Through Cultures: Expeditions in Cultural Psychology. Harvard University Press, Cambridge (1991)

43. Sunstein, C.R.: Social norms and social roles. Columbia Law Rev. **96**(4), 903–968 (1996)
44. Triandis, H.C., Marin, G., Hui, C.H., Lisansky, J., Ottati, V.: Role perceptions of hispanic young adults. J. Cross Cult. Psychol. **15**(3), 297–320 (1984)
45. Turian, J., Ratinov, L., Bengio, Y.: Word representations: a simple and general method for semi-supervised learning. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics, pp. 384–394. Association for Computational Linguistics (2010)
46. Wang, S., Manning, C.D.: Baselines and bigrams: simple, good sentiment and topic classification. In: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers, vol. 2, pp. 90–94 (2012). http://dl.acm.org/citation.cfm?id=2390688
47. Wilson, T., et al.: OpinionFinder: a system for subjectivity analysis. In: Proceedings of HLT/EMNLP 2005 Interactive Demonstrations (2005)
48. Zafar, L., Afzal, M.T., Ahmed, U.: Exploiting polarity features for developing sentiment analysis tool. In: EMSASW@ ESWC (2017)
49. Zou, W.Y., Socher, R., Cer, D., Manning, C.D.: Bilingual word embeddings for phrase-based machine translation. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 1393–1398 (2013)