Classification of Children's Social Dominance in Group Interactions with Robots

Sarah Strohkorb
Yale University
51 Prospect Street
New Haven, CT 06521 USA
sarah.strohkorb@yale.edu

Iolanda Leite
Yale University
51 Prospect Street
New Haven, CT 06521 USA
iolanda.leite@yale.edu

Natalie Warren
Yale University
51 Prospect Street
New Haven, CT 06521 USA
natalie.warren@yale.edu

Brian Scassellati
Yale University
51 Prospect Street
New Haven, CT 06521 USA
brian.scassellati@yale.edu

ABSTRACT

As social robots become more widespread in educational environments, their ability to understand group dynamics and engage multiple children in social interactions is crucial. Social dominance is a highly influential factor in social interactions, expressed through both verbal and nonverbal behaviors. In this paper, we present a method for determining whether a participant is high or low in social dominance in a group interaction with children and robots. We investigated the correlation between many verbal and nonverbal behavioral features with social dominance levels collected through teacher surveys. We additionally implemented Logistic Regression and Support Vector Machines models with classification accuracies of 81% and 89%, respectively, showing that using a small subset of nonverbal behavioral features, these models can successfully classify children's social dominance level. Our approach for classifying social dominance is novel not only for its application to children, but also for achieving high classification accuracies using a reduced set of nonverbal features that, in future work, can be automatically extracted with current sensing technology.

Categories and Subject Descriptors

I.2.6 [**Artificial Intelligence**]: Learning

Keywords

Social dominance; children; classification; groups; human-robot interaction.

ICMI 2015, November 9–13, 2015, Seattle, WA, USA.
© 2015 ACM. ISBN 978-1-4503-3912-4/15/11 ...$15.00.
DOI: http://dx.doi.org/10.1145/XXX.2818346.2820735.

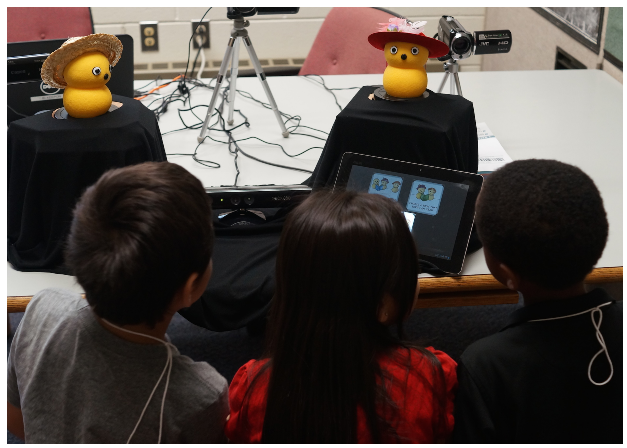

Figure 1: Interactive narrative scenario with social robots Leo and Berry (Robot 1 and Robot 2, respectively) used for data collection.

1. INTRODUCTION

Social robots are becoming increasingly common tools in education [19, 26]. As they are deployed in unstructured real-world environments such as museums and schools, their ability to interact with groups of users (children in particular) acquires a fundamental importance. In addition to evident decreases in cost, time, and space, research suggests that interactions between groups of children and robots share some of the positive outcomes of traditional learning experiences in groups [17]. However, the dynamics of group social interactions introduce additional challenges that are nonexistent in typical one-on-one human-robot interactions. For example, some children might dominate the interaction, giving other group members fewer opportunities to contribute. If robots could detect such situations, they could mediate the interaction so that all children can benefit from the group learning experience.

One of the most important constructs in group dynamics is the notion of social dominance, defined as "a relational, behavioral and interacting state that reflects the actual achievement of influence or control over another via communicative actions" [5]. Social dominance literature suggests that high dominant individuals benefit from social attention that low dominant ones do not [10], which can potentially impact the performance of low dominant children in group learning environments. Since it plays such an important role in interpersonal group relationships, social dominance has been conceptualized and studied from many different perspectives, from evolutionary theories [9] to psychological approaches [22]. Across disciplines, most authors agree that social dominance differs from the related concepts of power and status in that dominance is a means of achieving these through creating expressive, relationally-based strategies or actions [5].

The communicative nature of social dominance offers an opportunity for machines to automatically recognize dominant behaviors in social interactions. In fact, several efforts have been successful at automatically classifying dominance levels in adults based on their behavior [12, 13, 14]. While there have also been attempts at understanding the most common behaviors in dominant children [15, 24], none of these works focused on automatic classification using those behaviors.

In this paper, we provide a first investigation on the automatic classification of children's dominance level (high vs. low) using verbal and nonverbal behavioral features. The motivation for addressing automatic classification of social dominance in groups of children, given similar efforts in adults, is twofold. First, theory suggests that dominant behaviors in children are more related to evolutionary approaches of resource acquisition and conflict. As stated by Hawley [9], social dominance in children is "grounded in the differential ability to acquire resources in the social group, regardless of the means by which this is done." Second, previous research on automatic analysis of social behavior in groups of children shows that adult rules do not necessarily apply to children [16]. If dominant behaviors in children differ from adult behaviors, one cannot expect that the adult computational models will generalize well when applied to children.

This work is situated within a project whose overarching goal is to use socially assistive robots to help children build their emotional intelligence skills through interactive role-playing. The effective acquisition of emotional skills is crucial for children's social development [3, 21, 27], but it requires constant practice in diverse social situations. One way for developing these skills is by active learning techniques such as role-playing scenarios. However, in emotionally charged domains (e.g., bullying prevention), having children take an active part in the role-play may cause undesirable consequences, but observing the interaction might adequately serve as a learning experience. Here, robots offer an inexpensive alternative to human actors because they can display controlled behavior across interventions while having more social presence than virtual characters on a screen [18, 20].

To investigate whether high and low dominant children can be automatically classified using verbal and nonverbal behaviors, we started by collecting a data corpus of 21 children (ages 6 to 8 years) interacting with social robots in the context of interactive educational narratives described above. Our ground truth was based on teacher evaluations of each student using an Interpersonal Dominance Scale [5]. After analyzing and annotating the videos in our data corpus, a set of verbal and nonverbal features were extracted and used as the input of two machine learning binary classifiers (Logistic Regression and Support Vector Machines). Our results suggest that high/low children's dominance level can be successfully classified with accuracy values of 81% and 89% for the Logistic Regression and Support Vector Machines methods, respectively. The fact that most of the relevant features in our model can be extracted automatically with current state-of-the-art perception systems strengthens the contribution of our approach, given the possibility of applying the obtained models in real-time to groups of children interacting with robots.

2. RELATED WORK

This section describes previous research on social dominance in both adults and children. We give special attention to works that investigate which verbal and nonverbal behaviors are relevant for characterizing dominance across the lifespan.

2.1 Social Dominance in Adults

Most of the work on automatic classification of social dominance has been conducted with adult participants. The Augmented Multi-Party Interaction (AMI) corpus is a main source of interaction data for many researchers interested in computational modeling of this construct [12, 13, 14]. The AMI corpus consists of videos of meetings of 4-5 people with prescribed individual roles working towards a common goal. While this corpus has contributed significantly to advances in social dominance research in adults, for the reasons enumerated earlier it is not expected that the outcomes of this research can generalize well to groups of children.

Previous work has identified a set of multimodal features relevant for the automatic classification of social dominance in adults, including verbal cues like speech energy [13, 14] and interruptions [8, 14, 25, 28], visual cues like eye gaze [12, 23] and looking at another while talking [8], as well as kinesic cues like body posture [8] and gestures [8, 13]. In particular, there is a good deal of evidence that speaking time is a very strong cue for dominance in adults [14, 25]. Jayagopi et al. [13] obtained a high classification accuracy using an unsupervised model with only a total speaking length feature. The total number of speaking turns (with short utterances removed) performed equally well in an unsupervised model, and a combination of audio features showed over 90% accuracy in a supervised model [13]. Gaze and looking-time measures have also been used in many adult social dominance studies. The Visual Dominance Ratio introduced by Dovidio and Ellyson [7] and the amount of time others look at the participant in question have proven especially significant in this domain [8, 12].

2.2 Social Dominance in Children

The smaller body of work investigating cues to social dominance in children is different from work with adults in several important ways. First, children undergo a shift in the ways they manifest dominance around the age of five. Younger children tend to display dominance in terms of aggression and resource-directness, but this gives way to more prosocial behaviors as they age [9]. Hawley's work examining dominance expression supports this claim, showing that on average, socially dominant children ages 4-7 exhibited prosocial behavior twice as often as they exhibited coercive behaviors, while non-socially dominant children engaged in both behaviors equally [10]. This explains why there are many studies exploring the link between aggression and social dominance in young children, but not in older children or adults. In research on dominance levels of 6- to 8-year-old children, it is reasonable to analyze features both found in adults and younger children, in order to capture dominance-cueing behaviors on both sides of this developmental shift.

Many of the dominance cues observed in younger children are actively social or agonistic in nature – for example, peer visual regard [24], hostility and aggression [10], and reactions to initiated agonism [29]. A few studies have investigated the connection between dominance and resource utilization ratios, comparing time children spent engaged in an enjoyable individual activity (watching a movie on a viewing machine that must be hand-cranked by someone other than the viewer), to time spent helping others enjoy the activity [15]. Pellegrini et al. [24] studied naturalistic social interactions focusing on aggressive bouts. These methods yielded extremely interesting results, but were extremely time-intensive and focused on the idea of resource utilization rather than general social dominance. A main limitation of these studies is that, with the notable exception of peer visual regard, many of these cues are highly domain-specific. That is, the cues are only present or are only relevant within the specific social groups and environments set up by the researchers. Another limitation is that some of these cues, such as hostility, are potentially as hard to capture automatically as the notion of social dominance itself.

2.3 Our Approach

In light of the limitations of prior research in this area, our work attempts to establish reliable cues to social dominance in children which, like the features explored in adults, can be easily extracted over a short stretch of typical group interactions. The included features not only directly correspond to intuitive ideas of dominance, but also show high potential for good results in computational modeling. The use of machine learning models, which to our knowledge have not yet been applied to children's data in this domain, provides an opportunity to accurately classify dominance behaviors at an age where children's manifestations of dominance are radically shifting.

3. INTERACTIVE NARRATIVES FOR EMOTIONAL LITERACY

Previous research has shown that emotional intelligence skills are crucial for guiding decision-making, attention, and behavioral responses, and are necessary for children to engage in the social world [3, 21]. The case study used in this work includes two MyKeepon Robots (see Figure 1) that play out an interactive story about inclusion. One of the robots, Leo (hereafter referred to as Robot 2), is new at school. At specific moments in the narrative, children can influence the storyline by choosing from a set of actions that Berry (hereafter referred to as Robot 1) can take to make Leo feel included. Children can pick among the story options using a tablet. In other words, they can tell Berry what to do and see the impact of the selected actions on the course of the story. The robots can display different animations during the interaction, such as speaking, idling (while they are waiting for children's choices or listening to the other robot) or bouncing (moving up and down and to the sides). Pre-recorded adult utterances, with modified pitch signal to make them more childlike, were used for the robots' voices.

In the data collection experiment reported in this paper, the only perception the robots had from the environment was the children's story choice input from the tablet. However, our ultimate goal is to endow the robots with social awareness mechanisms that will likely improve the interaction and learning experience. One such mechanism is an automatic system for classifying children's dominance levels in this setting. With this knowledge about the group, the robots could prevent socially dominant children from controlling the interaction and prompt less dominant children to intervene more often.

4. DATA COLLECTION

We collected a corpus of video and audio data of groups of children interacting with robots using the interactive narrative scenario presented in the previous section.

4.1 Participants

Our data corpus consists of 21 children (13 female, 8 male), with ages between 6 and 8 years old ($M = 7.53$, $SD = 0.51$), interacting in small groups of 3 (7 groups in total) with two social robots. One group had females only and all the other groups were gender balanced (at least one male or female). Participants were first and second graders from an elementary school in the East Coast region of the United States. Each group contained students from the same class and our participant pool included 5 different classes. The data collection took place in a small meeting room of the elementary school.

Ethnicity, as reported by guardians, was 17.5% African American, 17.5% Caucasian, 25% Hispanic; 27.5% reported more than one ethnicity, and 12.5% did not report. The annual income reported by guardians was as follows: 30% in \$0-\$20,000, 42.5% in \$20,000-\$50,000, and 10% in the \$50,000-\$100,000 range; 17.5% did not report.

4.2 Procedure

One experimenter was present in the small meeting room for the entire data collection session. The experimenter started by introducing the participants to the two robots, Berry (Robot 1) and Leo (Robot 2), and explaining that the robots would play out a story, and then when the story stopped, they could decide what action Berry would take next from the options that appeared on the tablet. Participants were told that they would have to choose the next story option as a group, and then one of them would select their choice on the tablet.

The interactive story about inclusion used in this data collection contained an introductory scene and three different options that participants could then freely explore. The interaction ended when participants explored all three story options. The average interaction time, from the moment when participants selected the first story option until the robots played all the possible scenes, was 4 minutes 36 seconds ($SD = 39$ seconds).

4.3 Interaction Data

Three HD cameras were used to record the interaction. Each camera captured mainly the upper body posture and face of one participant. Log files containing the content of the robots' actions (speech and nonverbal behaviors) and the story choices made by the children were generated. The logs contained timestamps to allow future synchronization with the remaining data.

A coding scheme including verbal and nonverbal behaviors was developed based on the social dominance literature reported in section 2. Using the videos collected during the interactions, five annotators coded the start and end times of specified participant behaviors defined in the coding scheme. The verbal behaviors included **utterance type** – *making a demand*, *making a suggestion or prompt*, *helping others*, *thinking aloud*, *insulting* or *other* for all the talking behaviors that did not fit in the earlier categories, **utterance addressee** – *to colleagues*, *to robots*, *to experimenter*, or *thinking aloud*, and **interruptions** – *successful* or *unsuccessful*. The nonverbal behaviors defined in the coding scheme were **gaze** – *looking at the robots*, *looking at other participants*, *looking elsewhere* and *unclear*, **gestures** – *intrusive* (e.g., gestures that take others' space), *illustrative* gestures while talking and *adaptors* (e.g., nervous habits, or fidgeting), and the presence of **physical coercion**. Annotations were done using the ELAN annotation tool [4]. To avoid issues with intra-coder reliability, each coding category was coded by only one annotator.

4.4 Social Dominance Ground Truth

Our ground truth values for social dominance were based on teacher evaluations of participants collected during an individual interview with each teacher. We adapted the Interpersonal Dominance Scale developed by Burgoon [5], which has been previously used in HCI research [2, 11]. Out of the 32 items comprising Burgoon's original Interpersonal Dominance Scale, we selected the ones that seemed more relevant to group interactions between children, ending up with 11 items. We asked the teachers, for each study participant in his/her class, how often this child (1) takes charge of conversations, (2) talks more than listens, (3) influences others, (4) is very expressive during interactions, (5) wins arguments against his/her peers, (6) is concerned with what others think of him/her, (7) is successful at persuading others, (8) exhibits "bossy" behaviors, (9) wants to be in charge, (10) is the main focus of attention and (11) has influence in choosing group activities. Teachers were asked to rate each one of these assertions on a scale of 1 to 5, 1 meaning "almost never" and 5 meaning "almost always."

Our social dominance scale had excellent internal consistency (Chronbach's $\alpha = 0.96$). We averaged the teacher scores on all 11 scale items to obtain an absolute dominance score for each participant between the range of 1.0 and 5.0 ($M = 2.86, SD = 1.27$). The least dominant child had an average score of 1.27, while the most dominant was scored 4.73. The distribution of the scores is shown in Figure 2.

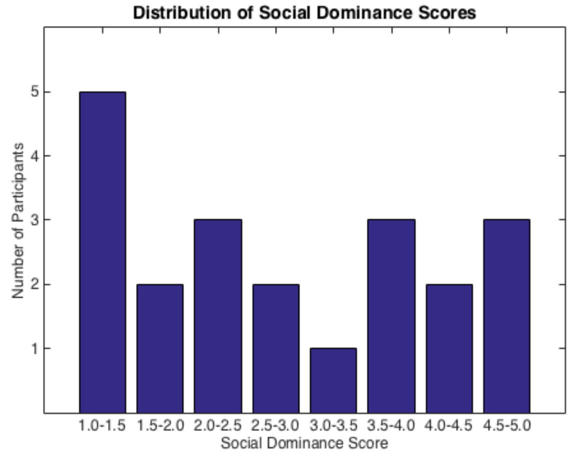

Figure 2: The distribution of the social dominance scores as evaluated by the teachers of the participants ($M = 2.86, SD = 1.27$).

5. FEATURE EXTRACTION & ANALYSIS

From the annotations described in section 4.3, we extracted a set of domain-independent verbal, nonverbal and combined (verbal and nonverbal) features for further analysis. In addition to these, gender was also included as a feature given the controversial results as to whether or not it has an effect on children's dominance [9, 15].

The feature analysis described here and the generated models reported in the next section rely on hand-annotated data. This deliberate choice was made because we wanted to distinguish the adequacy of our feature set to classify social dominance from the adequacy of particular feature detectors. However, our future goal is to replace the human-annotated features with autonomous perception systems, such that we can run a real-time implementation of the developed models in our robots.

5.1 Verbal Behaviors

We examined the following verbal behaviors that characterize participants' utterances: total talking time, interruptions, utterance addressee and utterance type. The extracted features from all these behaviors represent the total duration that each participant expressed that behavior, normalized by the length of the interaction.

Total talking time. We recorded the total talking time of each participant during the interaction.

Interruptions. This feature encodes the total time a participant spends interrupting other participants. An interruption is defined as any time a participant starts to talk while another participant is already talking. Interruptions are noted as either successful or unsuccessful. A successful interruption is marked for participant A when they start talking while participant B is talking and participant B stops talking before participant A does.

Utterance Addressee. We recorded the total time a participant's vocalization is directed to one of the following audiences: (1) to colleagues, (2) to robots, (3) to experimenter, or (4) thinking aloud. Utterance addressee is used to compute some of the combined features described later on (looking while talking and looking while listening).

Utterance Type. The subject of each utterance is recorded in order to evaluate the use of prosocial vs. coercive behaviors. We recorded the total time a participant's utterances fall into one of the following categories: (1) making a demand, (2) making a suggestion or prompt, (3) helping (demonstrate, guide, or help), (4) thinking aloud, (5) insulting, (6) other.

(a) Illustrative Gesture (b) Physical Coercion

Figure 3: Examples of some of our nonverbal behavior features: illustrative gestures and physical coercion.

5.2 Nonverbal Behaviors

We also investigated the following nonverbal behavior features: gestures, physical coercion, and gaze. These features represent the total time each participant exhibits these behaviors and are normalized by the duration of the interaction. Some of these nonverbal features (illustrative gestures and physical coercion) are shown in selected participants in Figure 3.

Gestures. We recorded the total time each a participant expressed either illustrative, intrusive, or adaptor gestures. An illustrative gesture is displayed with the goal of making a point or accentuate what the participant is saying. A intrusive gesture invades the other children's social space, such as pointing somewhere in an aggressive manner. An adaptor gesture is a nervous habit like playing with hair, clothes or putting fingers in the mouth.

Physical Coercion. We recorded the total time participant expressed physical coercion (shoving, grabbing, intentionally bumping into, etc.).

Gaze. We recorded the time each participant looks at one of the following specified targets in the scenario: (1) looking at Robot 1, (2) looking at Robot 2, (3) looking at left participant, (4) looking at middle participant, (5) looking at right participant, (6) looking elsewhere, (7) unclear. This feature is used to calculate some of the combined features of the next subsection (looking while talking and looking while listening) as well as the 'looked at' feature. The 'looked at' feature represents the total time each participant is looked at by their peers (derived from the features looking at left/middle/right participant).

5.3 Verbal/Nonverbal Combined Behaviors

Research has shown that individuals with high social dominance spend more time looking at the person they are talking to while talking, and less time looking at others while listening to them compared to individuals lower in social dominance [7, 8, 12]. Using a combination of the 'utterance addressee' and 'gaze' features, we derived the following features:

Looking while listening to colleagues. The total time that a participant looks at their colleagues while their colleagues are talking. This feature is normalized by the total time their peers are talking during the interaction.

Looking while listening to robots. The total time that a participant looks at the robots while they are talking. This feature is normalized by the duration of the interaction because the robots talk for an equivalent amount of time in each interaction.

Looking while talking to colleagues. The total time that a participant looks at their colleagues when they are talking to them. This feature is normalized by the total time the participant talks during the interaction.

Looking while talking to robots. The total time that a participant looks at the robots when they are talking to them. This feature is also normalized by the total time the participant talks during the interaction.

5.4 Individual Differences

In addition to the verbal, nonverbal, and combined verbal/nonverbal behaviors, we also noted each participant's gender: either male (1) or female (0).

5.5 Feature Correlation Analysis

We calculated the Pearson correlation coefficients (ρ) between children's social dominance scores and all the features we just described. In our feature set, two features were significantly correlated: 'looked at' ($\rho = -0.587$, $p \leq 0.01$) and 'looking at Robot 1' ($\rho = -0.449$, $p \leq 0.05$). Additionally, a visual analysis of all scatter plots revealed marginally significant trends in other features as well, particularly in the ones with higher absolute correlation coefficients. The marginal significance in these features might have been due to the limited duration of the interaction or the moderate number of participants in our data.

From the literature, we would expect that there would be a positive, not negative, correlation between the amount of time a participant is looked at and their social dominance rank [12]. Our observed negative correlation could be due to the changes in the interaction dynamics and the presence of the robots, who are talking for most of the interaction. It is possible that the most dominant agent in the interaction is one of the robots – potentially Robot 1 (Berry), who does most of the talking. Thus, the way that the robots potentially affect the group dynamics could account for this negative correlation.

It is also interesting to observe that the amount of time that a participant looks at Robot 1 (Berry), the robot whose actions the children are choosing, has a negative correlation with the social dominance rank. It is plausible that Robot 1 is the most dominant agent in the interaction, and would be looked at less by the high dominant participants. Previous literature reported in section 2 supports this idea, showing that individuals with high social dominance spend less time looking at others while listening to them compared to individuals lower in social dominance [7, 8, 12].

6. SOCIAL DOMINANCE CLASSIFICATION

This section describes the experimental procedure followed to automatically classify social dominance based on a selected set of features. We discuss the accuracy of the obtained data-driven models and their ability to classify social dominance in small groups of children.

6.1 High/Low Classification Labels

Considering the number of participants in our data set, we opted for binary classification, grouping participants in either *high* or *low* social dominance categories. Using the social dominance scores collected from the teachers (ranging from 1 to 5), each child was labelled as high in dominance if their total score was greater than or equal to 3.0, and low dominance otherwise. Of the 21 participants, 9 participants (6 females and 3 males) were classified as high ($M = 4.16, SD = 0.50$) and 12 participants (7 females and 5 males) were classified as low ($M = 1.88, SD = 0.58$) in social dominance.

6.2 Models of Social Dominance

For classification, we chose to pursue Logistic Regression [30] because this classification method does not require the features to be independent from each other, and Support Vector Machines (SVMs) [6] due to their effectiveness in dealing with a large set of features without overfitting.

We selected the 4 behavioral features with the highest Pearson correlation (ρ) absolute value for training our models. This final set of behavioral features and their corresponding absolute Pearson correlation values are shown in Table 1.

Behavior Feature	**Abs(ρ)**
Looked at	0.587
Looking at Robot 1	0.449
Physical coercion	0.388
Illustrative gestures	0.369

Table 1: The top 4 ranked behavior features based on the absolute value of the Pearson correlation value (ρ) of each feature with the respective ground truth.

To evaluate the performance of these models, we used a Leave-Pair-Out Cross Validation (LPOCV) procedure [1]. LPOCV has low-variance in both high and low dimensional feature space for small sample sizes, which makes it a good approach for our data. For each iteration, 2 participants were 'left out' as a test set and the remaining 19 participants constituted the training set. The model was then trained and evaluated using the chosen training and test sets. This process was repeated for all $\binom{21}{2} = 210$ combinations of participants. After all combinations were tried, the mean Accuracy and F1 Score were calculated. The Logistic Regression and SVM models used a C value (error penalty) of $1,000$ and the SVM model used a RBF (Radial Basis Function) Kernel.

We were able to achieve an accuracy of 0.890 and F1 Score of 0.871 with our SVM model an an accuracy of 0.807 and F1 Score of 0.764 with our Logistic Regression model. These results allow us to conclude that from a selection of behavioral features, it is possible to accurately predict high/low social dominance in children in group interactions.

7. DISCUSSION

Beyond the mean accuracy measures of the models, we also investigated the individuals that accounted for the majority of the model testing error. The chart in Figure 4 shows the number of misclassifications for each participant. Both the SVM and Logistic Regression models worked very well for most participants, but misclassified the same 2 or 3 participants in the majority of the cross-validation cycles. With the LPOCV (leave-pair-out cross validation) process, each participant is tested 20 times. Examining Figure 4, participants 6 and 13, with social dominance scores of 3.91 and 2.18, are misclassified every time they are tested by both the SVM and Logistic Regression models, accounting for most of the model testing error. From this analysis of where the misclassifications occur, our models seem to very accurately classify the majority of individuals, but there are a few individuals that the models consistently fail to classify correctly.

It is possible that our model misclassified a few participants consistently because we have not captured dimensions in our features that explain the behavior of these participants. These participants could have behaved differently because they had a complex social relationship with the other peers or they behaved in a different manner due to the presence of a very high or low dominant colleague. There are many possibilities because the dynamics of social interactions are so complex and difficult to quantify. In the future, it would be interesting to investigate these factors in a larger data corpus to better understand the main limitations of our models. It would also be interesting to see how additional features encoding friendship relations, physical size, and personality traits among the group members would contribute increase the accuracy of the models.

Contrasting to previous work in adults where verbal behavior plays a major role in the automatic classification of social dominance, the most discriminative features for classifying dominance in groups of children with our model were all nonverbal. Although these results are in line with previous research on children's dominance behavior, one should be careful about generalizing this finding beyond this particular learning domain [10, 24]. The absence of discriminative verbal features could have been due to the fact that, in our dataset, the percentage of talking time during the interaction is low and very similar for all children.

With regards to our classification of participants as either 'high' or 'low' in social dominance, this division might not not necessarily be the only or best labeling scheme. For our population size of 21 participants, we opted for a binary labeling criterion. Depending on the context and application of social dominance classification, a multiclass classification

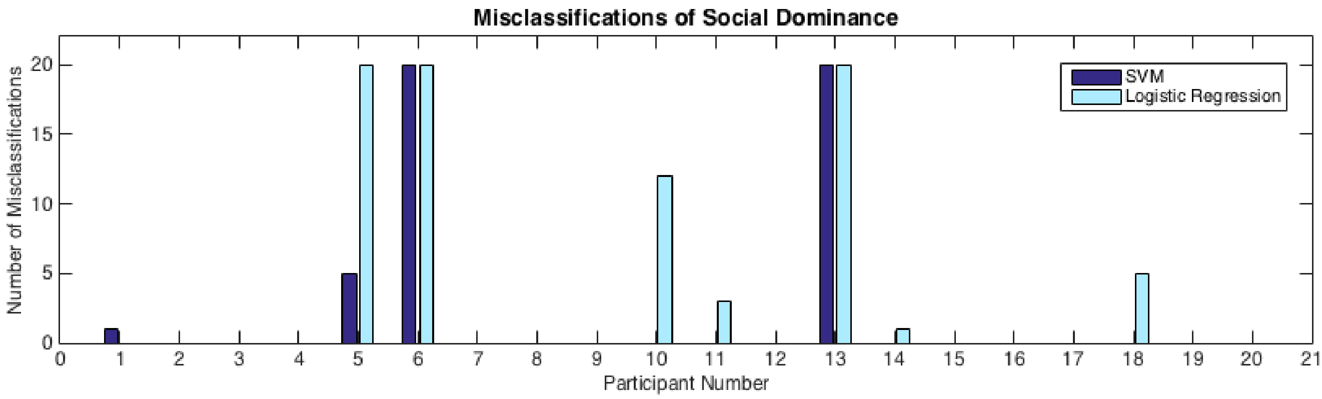

Figure 4: Number of misclassifications per participant in both SVM (dark blue) and Logistic Regression models (light blue). With the LPOCV (leave-par-out cross validation) process, each participant has 20 predictions. So the participants that has 20 misclassifications were misclassified every time they were tested. From this bar graph, it is clear that a few participants are responsible for the majority of our prediction errors.

criterion could be more appropriate. Based on our results, predicting the social dominance level automatically based on a set of behavioral features works for a high/low labeling and could work for other labelings as well.

8. CONCLUSION

The use of robots in educational settings is very promising because robots can provide children with much-needed individual attention. As robots enter real-world environments, they need to gain the necessary skills for interacting with groups of children, so they can provide teachers and students with more possibilities for personalized instruction.

In this paper, we investigated data-driven models to automatically classify social dominance levels of children (high vs. low) while interacting with robots in small groups. We trained Logistic Regression and SVM models with different feature sizes using multimodal data. Our Logistic Regression model had a mean accuracy of 0.807 and a F1 Score of 0.764, while our Support Vector Machines model had a mean accuracy of 0.890 and a F1 Score of 0.871. These results were obtained using hand-annotated nonverbal and verbal features. In the future, these features could feasibly be automated with the use of audio processing and visual perception tools, and be used in a real-time implementation. Such a classifier would allow a robot to be in a better position to mediate a group learning task.

Our results show that using a fairly reduced subset of nonverbal behaviors (eye gaze, physical coercion and illustrative gestures), we can automatically classify children's social dominance levels in this child-robot interaction domain. Moreover, the models seem robust independent of the dominance level of the other group elements. Further research is needed to ensure that similar results also apply when predicting social dominance in children using other group sizes, age groups and interaction contexts. Nevertheless, these findings are promising not only for human-robot interaction, but also for other types of interactive technology involving groups of children.

9. ACKNOWLEDGMENTS

This work is supported by NSF grant 1139078.

10. REFERENCES

[1] A. Airola, T. Pahikkala, W. Waegeman, B. De Baets, and T. Salakoski. A comparison of AUC estimators in small-sample studies. In *Proceedings of the 3rd International workshop on Machine Learning in Systems Biology*, pages 15–23, 2009.

[2] J. P. Birnholtz, T. Grossman, C. Mak, and R. Balakrishnan. An exploratory study of input configuration and group process in a negotiation task using a large display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 91–100. ACM, 2007.

[3] M. A. Brackett, S. E. Rivers, and P. Salovey. Emotional intelligence: Implications for personal, social, academic, and workplace success. *Social and Personality Psychology Compass*, 5(1):88–103, 2011.

[4] H. Brugman, A. Russel, and X. Nijmegen. Annotating multi-media/multi-modal resources with elan. In *LREC*, 2004.

[5] J. K. Burgoon, M. L. Johnson, and P. T. Koch. The nature and measurement of interpersonal dominance. *Communications Monographs*, 65(4):308–335, 1998.

[6] C. Cortes and V. Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[7] J. F. Dovidio and S. L. Ellyson. Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly*, 45(2):106–113, 1982.

[8] N. E. Dunbar and J. K. Burgoon. Perceptions of power and interactional dominance in interpersonal relationships. *Journal of Social and Personal Relationships*, 22(2):207–233, 2005.

[9] P. H. Hawley. The ontogenesis of social dominance: A strategy-based evolutionary perspective. *Developmental Review*, 19(1):97–132, 1999.

[10] P. H. Hawley. Social dominance and prosocial and coercive strategies of resource control in preschoolers. *International Journal of Behavioral Development*, 26(2):167–176, 2002.

[11] W. Huang, J. S. Olson, and G. M. Olson. Camera angle affects dominance in video-mediated

communication. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, pages 716–717. ACM, 2002.

[12] H. Hung, D. B. Jayagopi, S. Ba, J.-M. Odobez, and D. Gatica-Perez. Investigating automatic dominance estimation in groups from visual attention and speaking activity. In *Proceedings of the 10th International Conference on Multimodal Interfaces*, pages 233–236. ACM, 2008.

[13] D. B. Jayagopi, H. Hung, C. Yeo, and D. Gatica-Perez. Modeling dominance in group conversations using nonverbal activity cues. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(3):501–513, 2009.

[14] C. Jie and P. Peng. Recognize the most dominant person in multi-party meetings using nontraditional features. In *2010 IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS)*, volume 1, pages 312–316. IEEE, 2010.

[15] P. J. La Freniere and W. R. Charlesworth. Effects of friendship and dominance status on preschooler's resource utilization in a cooperative/competitive situation. *International Journal of Behavioral Development*, 10(3):345–358, 1987.

[16] I. Leite, H. Hajishirzi, S. Andrist, and J. Lehman. Managing chaos: Models of turn-taking in character-multichild interactions. In *Proceedings of the 15th International Conference on Multimodal Interaction*, ICMI '13, pages 43–50. ACM, 2013.

[17] I. Leite, M. McCoy, M. Lohani, D. Ullman, N. Salomons, C. Stokes, S. Rivers, and B. Scassellati. Emotional storytelling in the classroom: Individual versus group interaction between children and robots. In *Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction*, HRI '15. ACM, 2015 (to appear).

[18] I. Leite, A. Pereira, C. Martinho, and A. Paiva. Are emotional robots more fun to play with? In *The 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, pages 77–82. IEEE, 2008.

[19] D. Leyzberg, S. Spaulding, and B. Scassellati. Personalizing robot tutors to individuals' learning differences. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '14, pages 423–430. ACM, 2014.

[20] D. Leyzberg, S. Spaulding, M. Toneva, and B. Scassellati. The physical presence of a robot tutor increases cognitive learning gains. In *Proceedings of the 34th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society*, 2012.

[21] P. N. Lopes, P. Salovey, S. Coté, and M. Beers. Emotion regulation abilities and the quality of social interaction. *Emotion*, 5(1):113–118, 2005.

[22] H. A. Murray. *Explorations in Personality*. Oxford University Press, 1938.

[23] Y. Nakano and Y. Fukuhara. Estimating conversational dominance in multiparty interaction. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction*, ICMI '12, pages 77–84, New York, NY, USA, 2012. ACM.

[24] A. D. Pellegrini, C. J. Roseth, S. Mliner, C. M. Bohn, M. Van Ryzin, N. Vance, C. L. Cheatham, and A. Tarullo. Social dominance in preschool classrooms. *Journal of Comparative Psychology*, 121(1):54–64, 2007.

[25] R. Rienks and D. Heylen. Dominance detection in meetings using easily obtainable features. In *Machine Learning for Multimodal Interaction*, pages 76–86. Springer, 2006.

[26] M. Saerbeck, T. Schut, C. Bartneck, and M. D. Janse. Expressive robots in education: varying the degree of social supportive behavior of a robotic tutor. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1613–1622. ACM, 2010.

[27] P. Salovey and J. D. Mayer. Emotional intelligence. *Imagination, cognition and personality*, 9:185–211, 1989.

[28] L. Smith-Lovin and C. Brody. Interruptions in group discussions: The effects of gender and group composition. *American Sociological Review*, 54(3):424–435, 1989.

[29] F. Strayer and J. Strayer. An ethological analysis of social agonism and dominance relations among preschool children. *Child Development*, 47(4):980–989, 1976.

[30] H. Yu, F. Huang, and C. Lin. Dual coordinate descent methods for logistic regression and maximum entropy models. *Machine Learning*, 85(1-2):41–75, 2011.