2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
Macau, China, November 4-8, 2019

Identifying Opportunities for Relationship-Focused Robotic Interventions in Strained Hierarchical Relationships*

Michael J. Pettinati, *Graduate Student Member, IEEE* and Ronald C. Arkin, *Fellow IEEE*

***Abstract*— When disagreements arise in hierarchical relationships, relationship members sometimes prefer conflict management strategies that avoid or quickly end the overt conflict even if the relationship is left in a state of dissatisfaction. Our lab has proposed that a peripheral robotic agent may be able to support these types of relationships during conflict. In this paper, we present the results of an IRB-approved human-robot interaction study that examines how the members of a hierarchical relationship involved in conflict respond to the presence of an unengaged robot. This study serves as a baseline for additional studies. The unengaged robot appears to have a minimal influence on the interaction. The observed conflicts followed the patterns typically described in mediation literature. Our lab previously proposed a computational model to identify weakness and alienation in these relationships. We discuss a partial implementation of this model, and its ability to recognize problems in certain relationships within the data collected. Based on our observations, and the performance of the model's partial implementation, we suggest considerations that need to be made for an intervening robotic agent.**

I. Introduction

A Parkinson's patient sits quietly as her caregiver wonders aloud if she needs more substantive support and says that she seems depressed. The patient's true internal state is masked by a condition associated with the disease that limits nonverbal communication (see [26]). The patient has been working hard to stick to her treatment regime, feels good about her ability to be independent and does not feel depressed. The caregiver is making her feel uncomfortable.

Although overt conflict does not linger in the relationship described above, the relationship members are dissatisfied. The caregiver feels unable to reach or help the patient to whom she has responsibilities. She is alienated from the patient. The caregiver is not listening to the patient's experience. Similarly, the patient is weak. She feels unable to satisfy the expectations of the caregiver who is tasked with helping her. She is alienated; she may accept what is being presented to her, but she is not agreeable to it and does not necessarily understand why the caregiver feels this way. This mutual weakness and alienation is how conflict is described in transformative mediation literature [4, 5].

Relationships with power differentials often have these types of problematic conflict processes. The lower-power member wants to avoid conflict with or placate the higher-power member. The higher-power member can command the lower-power member to do something she does not want to.

Our lab developed a computational architecture to support relationships with power differentials experiencing interpersonal issues using relationship-focused interventions (see [21]). The agent must identify problematic states in the relationship (the computational model) and then trigger appropriate interventions to support moves of empowerment and recognition (the action-selection mechanism).

This paper presents some of the research being conducted to validate the computational architecture. The second section of the paper discusses related work. It goes over transformative mediation, the framework that provides a foundation for the problematic relationship states identified in our computational model as well as the intervention strategies in our action-selection mechanism. It discusses work examining how technology can support relationships in conflict and how our work is novel in this domain.

Section three presents an HRI study that examines how an unengaged robot is perceived by members of a hierarchical relationship in conflict. The unengaged robot made periodic and subtle movements throughout the interaction that were not in any way correlated with the discussion. The data collected in this study serves as a baseline, which will be compared to results collected with an intervening agent. The data collected provides examples of the problematic relationship states described in our model. The fourth section provides an overview of our model (originally presented in [21]) and discusses the successes and failures of a partial implementation at identifying problematic relationship states.

II. Related Work

A. Relationship-Focused Transformative Mediation

Transformative mediation frames conflict in terms of a breakdown in interpersonal communication [4]. One person feels unable to successfully handle the interaction she is having with another. She is dissatisfied with what the other is saying or doing, and she does not have the capability to change it. The other person in the interaction becomes evidence of a lack of capability, and the relationship member wants to withdraw from this other person to avoid experiencing shame. This alienation signals dissatisfaction to the other who is weakened. The other person in the relationship sees the original individual as evidence of a lack of her capability and withdraws. This cycle leads to mutual weakness (shame) and complete alienation from one another.

The goal in the transformative mediation process is to support relationship members such that they are able to have

*Research supported by the National Science Foundation under Grant #IIS 1317214. This work is being done in collaboration with Linda Tickle-Degnen and Matthias Scheutz at Tufts University.

M. J. Pettinati (corresponding author, e-mail: mpettinati3@gatech.edu) and R. C. Arkin (e-mail: arkin@cc.gatech.edu) are with the School of Interactive Computing at Georgia Tech, Atlanta, GA 30332 USA.

the interaction that they want to have and reach a mutual understanding [4, 5]. It is not about getting agreement on any issue or forcing certain interaction strategies. The relationship members should present themselves as they want, hear each other clearly, and engage each other directly.

A robot task with supporting relationships needs to intervene when states of weakness and alienation are present in the relationship. It can be very difficult to identify these states in relationships because cues are highly context dependent [20, 22]. One must often depend on a constellation of these signals to be confident that the relationship is strained. Once the robot has identified the presence of these states, it must orient the parties to each other and empower the individuals so that they can open up to one another.

B. Technology Supporting Human-Human Relationships

There has been little work done with respect to technology trying to support human-human relationships. Hoffman [10] tested a robotic lamp that was meant to encourage empathy in situations where couples were arguing. The lamp shivered when the couples raised their voices beyond a certain threshold. The lamp seemed to be able to break the tension in certain relationships and was not a distraction to the couples. The use of shivering, however, seems to discourage the couples from raising their voices; this is a prescriptive intervention and not relationship-focused [4, 5]. Loudness is not sufficient for identifying the weakness and alienation that transformative mediation says is damaging to relationships [20, 22]. Our focus is on developing an agent that can support interactions with relationship-focused interventions.

Jung et al. [13] showed a robot could help to repair one individual's impression of another after a personal attack. In this work, a confederate used a scripted personal attack toward a participant. When the robot made a repair statement (chastised the confederate) after the attack, the participant had a better impression of the confederate than when the robot made a general statement about the task. The robot's intervention in this case were prescriptive, and the responses were to single scripted insults; there was not understanding of the relationship state. Jung [14] found that the same dynamics that make for lasting marriages also were predictive of successful student engineering teams (a balance of positive and negative affect and limited hostile affect). He suggests the potential for feedback systems to balance positive and negative affect in task-based situations. There was not an implementation of such a system in these papers.

Shen et al. [24] showed how a robot could enable children to better identify and have more productive conflict processes. Here productive was defined by a mutually agreed upon resolution being reached, so it was not relationship-focused interventions, and the robot was controlled by experimenters. Costa et al. [6] found that they could help calm and empower participants by playing their altered voices back to them during conflict. These studies were not done with participants face-to-face and in one case participants were not interacting with a live human. There was also no model of the relationship state.

Figure 1. Robokind's R25 Platform.

III. HUMAN-ROBOT INTERACTION STUDY

This experiment is the first half of a larger work that is concerned with understanding if an intervening robot can support healthier interactions between two people in a hierarchical relationship. The robot used in the study is Robokind's R25[1] (Figure 1). This study has two between-subject conditions. In one condition, a participant is assigned a lower-power role in a discussion with an experiment confederate where there is conflict. In the other condition, the participant is assigned a higher-power role.

In this experiment, a confederate and a participant are generating an argument about a recent law change in Georgia that divided the student population on the Georgia Tech campus. Specifically, they construct an argument where they choose and defend one side of the issue: "I feel less safe on the Georgia Tech campus with it being legal for permit holders to conceal carry a gun, and it should not be allowed".

The pair is told that they have fifteen minutes to discuss the issue and generate a single argument before the experimenter enters back into the room to hear a final argument. They are told that this final argument will be made by one individual member of the pair, and that the final argument will determine compensation for both of them. It is allegedly assessed on provided guidelines.

The higher-power member in the relationship is the member who makes the final argument. This relationship member has direct control over the compensation; she is explicitly told that she can take the fifteen-minute discussion into account when arguing or not. The higher-power member's only responsibility is to make the strongest possible final argument. The lower-power individual is told to use the fifteen minutes to present her ideas, but she will have to remain silent during the final argument. The responsibility of the lower-power member is to help the make the final argument as strong as possible. Half of the participants are assigned the higher-power role, and half are assigned the lower-power role.

For clarity in this discussion, let us assume that the confederate is a male (which he was) and the subject is female (participants from both genders were included in this study). The confederate, regardless of his role, helps to drive conflict, using provided guidelines, such that the participant is put in a weakened state and becomes alienated from him early in the interaction. After the participant has the opportunity to make a point for her side of the issue, the

[1] https://robots4autism.com/milo/

confederate uses the language identified in the transformative mediation literature (e.g. [19]) that shows a refusal to recognize the other party's thoughts and feelings After dismissing the participant's viewpoint, the confederate argues his side with language from the literature that shows weakness. Finally, the confederate makes a remark that requests recognition from the participant. The participant will see someone who does not seem to respect her thoughts and seems unable to change his mind. Alienation from the confederate drives weakness in the participant [4].

After driving this type of conflict early in the interaction, the confederate mirrors the responsiveness of the participant. The confederate requests recognition (using the language of Moen et al. [19]) at the conclusion of each of his arguments. These statements say, "I want you to understand me", and require acknowledgment of the legitimacy of the confederate's feelings. If recognition is not offered, weakness continues. The confederate continues using responses that follow the pattern described above. If recognition is given, it leads to strength and recognition.

Showing strength means speaking clearly and without the language of frustration or uncertainty. Showing recognition means being attentive and open to the participant when the participant is speaking. It also means acknowledging the validity of the other's feelings and respecting the other's thoughts about the topic (e.g. [4]). The confederate's new pattern (until the participant refuses recognition again) is showing emerging understanding, making a point on his side of the issue using stronger language, and making another request for recognition. To show recognition, the confederate explicitly says something that shows he is considering the participant's viewpoint, and he is willing to engage the participant directly even if he disagrees with the participant's viewpoint. The confederate decides whether or not to respond with strength and recognition or weakness and alienation based on whether the participant gives recognition.

Recognition is given if: the participant explicitly acknowledges the validity of the argument the confederate has just made, explicitly says he has made a point that is important for the argument, considers the confederate's argument in more detail, e.g. by saying she would appreciate that argument more "if…" or by asking questions regarding the argument, or if the participant admits that she is thinking about the topic in a new way. The participant dictates whether or not mutual recognition is given in the interaction.

A. Participants

There was a total of 31 participants tested during this study. The participants were all Georgia Tech students who confirmed that they held a strong opinion on the controversial topic introduced above. We restricted the study to this population because the topic is meaningful to Georgia Tech students and, perhaps, is less meaningful to those outside of that community because they are not regularly attending events on campus. If there was no emotional investment in the topic, then the dynamics being studied may not emerge.

During the 31 trials, there was 1 trial where the robot malfunctioned and did not move at all, and there were 2 participants who claimed that they knew the confederate was a member of the study team. These 3 trials were excluded from analysis. Within the 28 remaining trials, there were 14 participants who had the lower-power role, and there were 14 participants who had the higher-power role.

There were 11 males and 3 females who had the lower-power role. There were 10 males and 4 females who had the higher-power role. The average age of the participants with the lower-power role was 21.57. The average age of participants with the higher-power role was 20.93.

Participants completed self-report measures (Likert-style measures ranging from 1 to 7) at the outset of the study that related to their perceived dominance; their views on interpersonal conflict, and their views on robots helping in interpersonal relationships. There were no significant differences ($p > .05$) between groups on these measures.

B. Unengaged Robotic Behavior

In both conditions, the robot maintained an unengaged behavior for the entire fifteen-minute conversation. The unengaged robot makes certain movements periodically. The periods of the movements are long to prevent the participant from recognizing that the robot is repeating movements. The robot blinked, twisted its head slightly and slowly side to side, tipped its head up and down slightly and slowly, swung one or both arms, and stretched its hips slightly. These movements show no indication of the robot following the discussion. They were programmed to make the robot not stationary during the interaction.

C. Study Procedure

The experiment begins with the experimenter and confederate leaving the research building at the time the participant is scheduled to arrive. The confederate leaves through the back of the building and enters through the front doors to appear to be just another participant in the study. The experimenter goes directly to the lobby to greet the confederate and the participant. After both have arrived, the experimenter leads them to two cleared off desks just outside of the office where the discussion takes place. The pair complete the consent form for the study as well as the demographics survey (discussed above). The confederate pretends to fill out these forms until the participant is done.

After these forms have been completed, the participant and confederate are led into the office where the discussion takes place and asked to have a seat in their respective locations. The pair have identical sensors affixed. They were instructed to take three minutes to clear their heads (without interacting) to relax them before the discussion.

Once the three-minute time period ends, the experimenter sets up the interaction. He begins by telling the dyad that he is going to be asking them to work together to form an argument for or against a controversial topic. It is the goal of the dyad to form the best argument they are able to within a fifteen-minute period. They are explicitly told that they are forming a single argument. The experimenter tells the dyad that the strength of the argument is assessed using the guidelines defined on a page that is given to each dyad member. The dyad members are told that the strength of their argument will determine their compensation. This small act

of deception encourages the participant to have a stake in making a strong argument and helps to create weakness in the participant when the confederate begins to drive the conflict.

The experimenter asks the participant whether or not she feels less safe with concealed carry of a gun being allowed on the Georgia Tech campus. Once she has given an answer, the confederate is asked. The confederate will always choose the position opposite that of the participant.

The dyad members are also told that the interaction is structured to allow for a fair and equitable discussion. The dyad members will each have uninterrupted one to two-minute time periods to respond to what their partner has said and to make their next point. After one dyad member has had a one to two-minute period, he/she should yield to the partner for one to two minutes. The experimenter then assigns the participants their roles in the interaction, as described above, to establish the relationship hierarchy.

Finally, the robot is introduced to the dyad as something to support their communication. They are told it knows nothing about the topic at hand, and it may or may not intervene in the discussion. The experimenter then leaves the parties to have the discussion. The confederate follows the guidelines described above during the discussion. The confederate tries to appear just like another participant.

After the fifteen minutes conclude, the experimenter enters the room and requests the final argument from the higher-power dyad member. If the confederate is the higher-power dyad member, he uses a scripted argument. After the final argument is given, the experimenter tells the dyad that they will be compensated after they complete the final surveys. The dyad is shown to desks outside of the office where they complete questionnaires.

D. Measures

Objective, self-report, and video-rating measures were collected as part of this study. The objective measures included: the average utterance length by the participant, the total speaking time of the participant, the count of unique arguments presented by the participant, the count of mentions of emotion by the participant, and counts of the mentions of the robot and glances toward the robot by the participant. These were coded by one of the experimenters.

The self-report measures included a couple of previously validated Likert-style measures used with permission. Participants completed Jehn's [11, 12] intragroup conflict scales; they gave their impressions of the level of emotional and task-related conflict during the discussion. The participants also completed Burgoon and Hale's [2, 3] Relationship Communication Scale. The participants rated dimensions of the confederate's communication, including: involvement and affection, receptivity and trust, similarity, inclusion and depth, dominance, composure versus arousal, formality, and social versus task-orientation.

The participants also completed three Likert-style items (on a scale of 1 to 5) about their impressions of the robot during the discussion. These items were: "The robot was disruptive", "The robot clearly didn't understand how I felt", and "The robot was on my side". The final self-report measure included six Likert-style items again ranging from 1 to 5 that tried to assess participants' level of empowerment and recognition based on questions asked by mediators [9]. Items related to empowerment were: "I had the opportunity to express myself", "I had the opportunity to make my points and ask the questions I had of the other participant", and "This process helped me become clearer about my thoughts and my feelings on the issue". Items related to recognition were: "I gained a better understanding of the other person during the process", "I was able to see the other person's perspectives, views or opinions", and "The other person was able to see my perspective views or opinions".

There were video rating measures included as part of the study. One experimenter and one video coder, who was unaware of the study's purpose and hypotheses rated the videos using guides rooted in transformative mediation literature [4, 5, 19, 22] to identify states of weakness and alienation in the relationship. There is clearly the opportunity for experimenter bias as one of the experimenters was a video rater. The experimenter took precautions to ensure strict adherence to the scales. The video coder trained with the experimenter on six videos to reach a common understanding of the scales. The video coder and experimenter then rated all of the videos independently (including the six training videos). They checked their agreement every four videos to ensure that they could justify their ratings according to the scales (no ratings were changed during these checks).

The video coders gave a binary rating each time the participant spoke and each time the participant listened as to whether that section of video showed a lack of empowerment (weakness) in the participant or not. Each time the participant responded to a response given by the confederate, the participant's response was given a binary rating that indicated whether or not the participant was giving recognition.

Inter-rater reliability was assessed using Cohen's κ; a measure used in related works (e.g. [14]). The agreement between raters was $\kappa = .512$ on the weakness ratings. The agreement between raters was $\kappa = .502$ on recognition ratings. This is moderate agreement [25] and is comparable to the agreement by coders in previous studies (e.g. [14]).

For each participant, a notion of that person's weakness level and alienation level were generated. The experimenter divided the total number of weakness ratings given for the participant by the possible number of weakness ratings for that participant. The proportion of the responses that showed a lack of recognition were also recorded for each participant.

E. Results

To compare the lower and higher-power groups on each of the measures introduced in the previous section, we used a two-sample (unpaired) two-tailed t-test. We used Levene's Test of Equal Variance [16] to ensure that the groups did not violate the equal variance assumption. In the case, they did we ran a t-test where equal variance was not assumed. The t-test is the standard way of evaluating differences between groups in the literature and has been used in similar studies on measures we have incorporated here (e.g. [10, 13, 14]).

On certain measures, the normality assumption of the t-test did not hold. We used Shapiro and Wilk's [23] test of normality to check the normality assumption for each group

used in the t-test. When the data failed this test, we ran the Mann-Whitney U Test (a nonparametric test) [17]. Previous research has used nonparametric methods when the normality assumption of the t-test does not hold (e.g. [14]). To check the internal consistency of scales containing multiple items measuring a single construct, we used Cronbach's α [7]. If the value fell below 0.7, we tested each item independently.

Table 1. Higher-power participants viewed the confederate as a less engaged and colder during the discussion than lower-power participants. This provides evidence of participants internalizing the roles given to them by the experimenter;.

Relational Communication Scale Measures [2, 3] (Likert-scales 1 to 7)	**Two-Sample, Two-Tailed t-test Result**			
	Lower-Power Participant Mean (Standard Deviation) (n = 14)	***Higher-Power Participant Mean (Standard Deviation) (n = 14)***	***t-score***	***p-value***
Intimacy: Involvement / Affection	**5.621 (0.638)**	**4.813 (0.616)**	**t(26) = 3.409**	**p = .002**

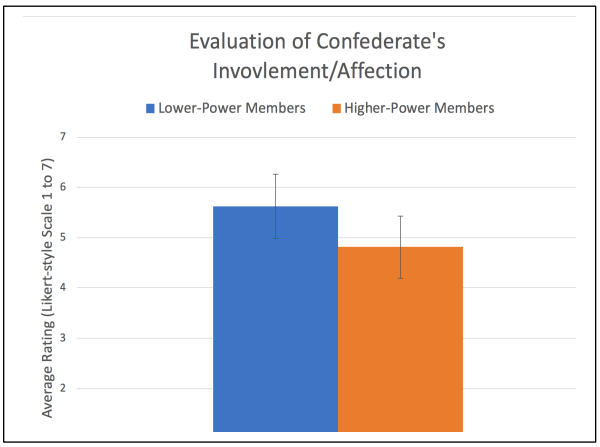

Figure 2. The rating of the confederate's involvement and affection in both the lower and higher-power conditions.

Table 2. Participants responded to the roles assigned by the experimenter for the interaction. Lower-power participants rated the higher-power confederate member as more dominate and social.

Relational Communication Scale Items [2, 3] (Likert-scales 1 to 7)	**Two-Tailed Mann-Whitney U Test Result**			
	Lower-Power Participant Median (Standard Deviation) (n = 14)	***Higher-Power Participant Median (Standard Deviation) (n = 14)***	***U***	***p-value***
"Your partner in this interaction dominated the conversation"	**4.0 (1.139)**	**2.0 (0.917)**	**U= 41.5**	**p = .010**
"Your partner in this interaction was as interested in building a good relationship as in completing the task at hand"	**6.0 (1.399)**	**3.5 (1.703)**	**U = 47.0**	**p = .020**

Jehn's [11, 12] emotional conflict scale had a Cronbach α = .381. We ran three independent tests for each question. Jehn's task-related conflict scale had a Cronbach α = .888. Therefore, we averaged the results of the three questions for each participant. There was no significant differences between the groups on these measures ($p > .05$). The perceived emotional conflict was not different between groups. There was very low amount of emotional conflict according to the measure (averages between 1 and 2 on a scale of 5). Participants indicated a moderate amount of task-related conflict in both groups (average 3 out of 5).

On the empowerment and recognition self-report measures, the Cronbach α values were .661 and .371 respectively. Each question was analyzed independently. There was not a significant difference between groups ($p > .05$) on all six questions. The participants reported feeling empowered and that they were able to give and receive recognition (averages between 4 and 5 on a scale of 5).

The Cronbach α value for the involvement and affection on the Burgoon and Hale Relational Communication Scale [2, 3] was 0.8536. Therefore, we averaged the items for each participant for this scale. There was a significant difference between the group with the participant in the lower-power role and the group with the participant in the higher-power role ($p < .05$). Participants in the lower-power position rated the confederate as being more engaged and warmer towards them. This result is summarized in Table 1 and Figure 2.

Another Relational Communication Scale measurement assessed the dominance of the confederate (how much control he exerted in the conversation). The Cronbach α value for this scale was 0.0752. The 21 items that composed the scale were analyzed individually using the same process as what has been described. Of the 21 questions, 20 were not significant at a $p = .05$ level ($p > .05$). The only item that was significant at $p = .05$ ($p < .05$) stated directly that the confederate "dominated the conversation". See Table 2.

There were 21 items related to dominance answered by each participant; it is common practice to use correction and lower the level of significance to $p = .0025$ [18] when being conservative. This suggests no significant difference between groups on the dominance scale. It is worth noting the effect size of this item was medium to large (z-score = 2.57307, r = .486). We speculate participants in the lower-power role may have perceived the confederate to have taken over the conversation more so than those in the higher-power role.

The measure related to the confederate's social versus task-oriented nature had a Cronbach's α value of .133168. The four items were analyzed independently. Three of the four items were not significant at $p = .05$ level ($p > .05$). The item that was significant at $p = .05$ ($p < .05$) stated the confederate "was as interested in building a good relationship as in completing the task at hand". See Table 2. Again there is the question of using correction, which would make the result not significant at $p = .0125$. Again, the effect size is medium to large (z-score = -2.32, r = .438). We speculate that participants in lower-power roles saw the confederate as more social-oriented than those in the higher-power roles.

The measures related to receptivity/trust, similarity/ inclusion/depth, arousal, and formality from the Relational

Communication Scale had Cronbach's α values of .7366, .7682, .8354 and .762 respectively. We averaged the items of these scales for each participant. There was not a significant difference between the two groups on any scale ($p > .05$).

There were no significant differences between groups ($p > .05$) on the three questions regarding the participants' impressions of the robot. Participants in both groups disagreed that the robot was disruptive (average of approximately 2 out of 5), and the participants in both groups were neutral when deciding if the robot understood how they felt or was on their side (averages of 3 out of 5).

On all of the objective measures, there was not a significant difference between groups ($p > .05$). It is also worth mentioning that none of the participants in either condition mentioned their feelings in the moment (e.g. being upset, off put by the confederate, etc.).

On the video rating measures, there was not a significant difference between groups. These results are summarized in Table 3 and Figure 3. For an average participant, nearly half of the video segments show weakness and alienation.

Table 3. The average participant in both groups had signs of weakness and alienation in nearly half of the video segments.

Proportion of Labelled Video Segments Showing Weakness/Alienation	**Two-Sample, Two-Tailed t-test Result**			
	Lower-Power Participant Mean (Standard Deviation) (n = 14)	***Higher-Power Participant Mean (Standard Deviation) (n = 14)***	***t-score***	***p-value***
Weakness	0.402 (0.249)	0.492 (0.255)	t(26) = -0.95	p = .351
Alienation	0.496 (0.237)	0.515 (0.259)	t(26) = -0.20	p = .844

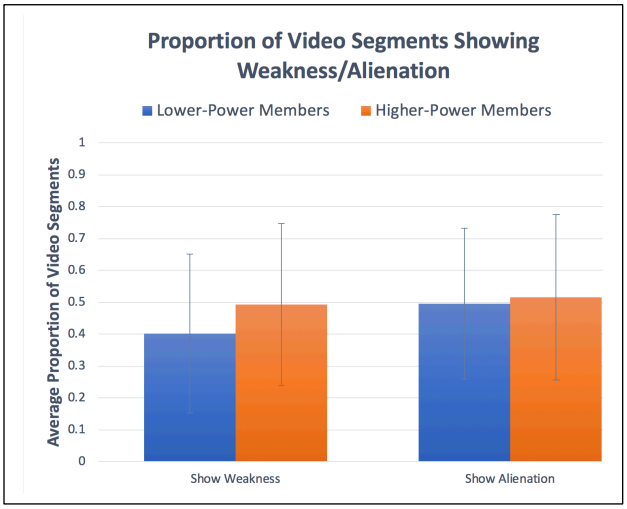

Figure 3. The proportion of video segments that show weakness and alienation for both the higher- and lower- power groups.

F. *Discussion*

The manipulation to set up a power dynamic within the relationship seemed to be successful. The higher-power participants saw the confederate as less engaged and colder compared to lower-power members. It makes sense that this is the case. In cases where the confederate was a lower power relationship member, he had no reason to be combative. He was explicitly told to help generate the strongest argument possible, and he immediately began as disagreeable (dismissing the ideas of the person tasked with actually making the argument). When the participants were lower-power relationship members, they were tasked with helping the confederate. He did not have to listen to their ideas. Their ideas were shown recognition if they showed recognition to the confederate at points later in the discussion.

The unengaged robot did not stop typical relationship dynamics from emerging in the hierarchical relationships. The lack of significant differences between groups on the empowerment/recognition measures, Jehn's intragroup conflict measure, and the video ratings suggests that the conflict was experienced the same for the higher-power member as the lower-power member. This is consistent with the literature in transformative mediation [4, 5]. The video ratings showed that almost half of the statements by the average participant showed weakness and a lack of recognition. This indicates that the confederate was able to generate the type of conflict described in the transformative mediation literature, and the mere presence of an unengaged robot did not disrupt this cycle of weakness and alienation.

There were no significant differences between how the lower- and higher-power participant members perceived the unengaged robot. The higher-power members did not view the agent as something they were responsible for. They did not give the agent additional attention. Qualitative remarks in both groups indicate that they were unsure what the robot's role was in the interaction and largely ignored the robot.

IV. The Relationship Model

In order to successfully support healthier communication in human-human relationships, a robot needs to be able to identify problematic states in the relationship and intervene when these states are active. The previous section introduced a study in which we collected examples of hierarchical relationships that had conflict present. In order to assess if a robot could help in such relationships, we need to compare the results from this previous study with results from a study in which a robot appropriately intervenes in hierarchical relationships with problematic conflict dynamics. This section explores work we have done to develop an autonomous robot that appropriately intervenes in strained relationships. We are currently conducting a study with an autonomous robot that intervenes to support the strained relationships using this work. The results from this next study will be compared to the study presented above.

Our lab developed a computational model to identify six states of weakness and alienation in hierarchical relationships [21]. The six states of weakness and alienation correspond to six robotic behaviors that trigger interventions to support the relationship. See our previous publication [21] for the full computational architecture. This computational model has a set of n sensors (S) that continually write readings to circular buffers. The readings within these circular buffers are abstracted into m percepts by m independent functions (F). These percepts are then mapped by a single function g to a vector indicating which states are active or inactive (X). The data flow of the computational model appears in Figure 4.

This paper is focused on identifying one specific state of weakness and one specific state of alienation in hierarchical relationships. The state of weakness is named negativity. This is where the lower-power relationship member withdraws from the higher-power member or aggressively pushes the higher-power member away (showing high degrees of shame or frustration). The state of alienation is named insensitivity. This state is active if the higher-power relationship member is inattentive toward or demonstrates aggression toward the lower-power member (e.g. raising her voice).

In the experiment described in the previous section, the experimenter and the video coder used guides to identify negativity and insensitivity within the trials. They gave a binary rating of negativity for each speaking part of the lower-power participants. They gave a binary rating for insensitivity for the speaking and listening parts in higher-power participants. As above, Cohen's κ was used to assess inter-rater reliability. The κ value for the negativity ratings was .526. The κ value for insensitivity was .551. Again, this signifies moderate agreement between raters [25] and is similar to reliability estimates in previous studies (e.g. [14]).

The experimenter implemented these two states. He used the labelled recordings to understand if the robot could label the recordings in the same way as the human raters.

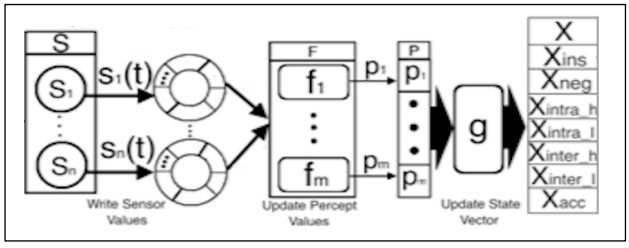

Figure 4. The data flow of the model to identify problematic relationship states. These states correspond to robotic behaviors that trigger interventions to support the relationship.

A. Partial Model Implementation

Two sensors were used for this implementation, the participant's lapel microphone and a webcam with a clear view of the participant's face. The percepts were: the loudness of the participant's voice, the fragmented nature of the participant's speech patterns, the tempo of the participant's speech, and the orientation of the participant (toward or away from the confederate).

To generate these percepts, we used openSMILE [8] for audio processing and dlib [15] and OpenCV [1] for visual processing. openSMILE provides a loudness measurement that is written to a circular buffer at 100Hz. One can also use the tools in openSMILE to find the duration of utterances in recordings. openSMILE also provides the number of onsets per second during these utterances (from the fundamental frequency), which can be used as an estimate of the number of syllables per second that the participant is speaking. The loudness percept value is determined by comparing loudness over a sliding window of time to a threshold set based on the participant's initial speech. The tempo percept is determined by comparing the tempo over a sliding window of utterances to constant values typical in relaxed conversation. The fragmented percept is determined by comparing the number of utterances per minute and the average utterance length over a sliding window of utterances to constant values typical in relaxed human conversations (the set values were hand tuned using the data gathered in the study discussed above).

To assess the orientation of the participant to the confederate, we used OpenCV to locate the face of the participant and the back of the confederate's head. We used dlib to identify key points on the participant's head and a generic three-dimensional model of a head to solve for the orientation of the participant's head. We wrote at 2Hz whether the participant's head was oriented toward or away from the confederate. The orientation percept was set by comparing the count of the number of times the participant oriented toward or away from the confederate to set values.

It can be difficult to understand if these states are active or not based on a single cue [20, 22]. The function g sets the states active (and the robot intervenes) if multiple percepts indicate the state is active.

B. Results and Conclusions

We played the recordings, and we had the algorithm label the weakness and alienation states. We had the algorithm label negativity as active or inactive during speaking parts in trials where the participant was the lower-power relationship member. The algorithm labeled the insensitivity state as active or inactive during each speaking and listening part when the participant was the higher-power member.

The precision of the algorithm when identifying the weakness state was .42; the recall of the algorithm was .385. The precision of the algorithm when identifying the insensitivity state was .516; the recall was .211. There were three clear sources of error associated with the automatic identification of the two states.

First, there were participants where the percepts were misleading. Participants spoke loudly and quickly because of excitement and were monotone while using language that pushed the confederate away. There were also technical issues. Certain participants positioned themselves such that the webcam had a poor view for identifying their face. The microphone of the participant sometimes picked up the confederate's voice as a participant utterance. Finally, the nonverbal and paralinguistic cues of these states did not always appear simultaneous to the state being labelled as active by the coders. These cues can occur in parallel, consecutively or spaced throughout an interaction [22].

The algorithm was accurate for certain relationships. In the lone case that the video raters agreed the negativity state was inactive the entire discussion, the algorithm labelled all video segments as inactive. In three of the five cases the video raters agreed that over fifty percent of the video segments showed negativity, the algorithm labelled over forty percent of the segments as showing negativity.

The study we are currently running that uses this implementation as part of an autonomous robotic system that intervenes when the relationship between the confederate and participant becomes strained.

V. CONCLUSION

We discussed a baseline human-robot interaction study that examined how participants in strained hierarchical relationships responded to an unengaged robot that was introduced as something to help with communication.

The unengaged robot was perceived similarly by higher and lower-power relationship members. It did not change normal dynamics from developing in the relationship.

- There was not a significant difference between the two groups on self-report measures relating to the robot's level of distraction, comprehension or neutrality.

- There was not a significant difference between the two groups on objective measures relating to attention given to the robot (e.g. gazes, mentions).

- There was similar amounts of weakness and alienation observed in both lower- and higher-power participants.

The higher-power relationship members did not necessarily see the robot an adversary who was meant to monitor them. Several participants in both high and lower-power roles indicated that the robot fell into the background, and they forgot it was there. Participants indicated they were unsure of the role of the robot when it did not intervene. This highlights the importance of interventions being clear and notably different from the unengaged behavior.

Due to the difficulty of recognizing states of weakness and alienation, it is important for early robotic interventions to be about orienting the parties to each other and encouraging speaking and decision-making.

- The precision values for our current algorithm indicates one out of every two to three active labels is accurate.

Early interventions should be positive and not assume anything about the affect of the participants. The robot should wait until there are several instances where it has detected negativity or insensitivity before trying to amplify the emotions (to ensure these emotions are present).

The insights from this work has been incorporated into an autonomous intervening robot that we are using in an ongoing study. This study follows the same procedure as the study presented here, except, instead of an unengaged robot, the robot intervenes when it believes the relationship is strained. The participants still have a conversation with the confederate while the robot is a bystander. The results from this study will be compared to the results presented here to get an understanding if a robot can use relationship-focused interventions to support relationships involved in conflict.

REFERENCES

[1] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library.* O'Reilly Media, Inc. 2008.

[2] J. K. Burgoon and J. L. Hale, "The fundamental topoi of relational communication", *Communication Monographs*, 51.3, 193-214, 1984.

[3] J. K. Burgoon and J. L. Hale, "Validation and measurement of the fundamental themes of relational communication", *Communication Monographs*, 54.1, 19-41, 1987.

[4] R. A. B. Bush and J. P. Folger, "Transformative mediation: Theoretical foundations" in *Transformative Mediation: A Sourcebook. Resources for Conflict Intervention Practitioners and Programs*, J. Folger, R. A. B. Bush & D. J. D. Noce, Ed., Hempstead, NY: Institute for the Study of Conflict Transformation, Inc., 2010, 15-30.

[5] R. A. B. Bush and J. P. Folger, "Transformative mediation: Core Practices." in *Transformative Mediation: A Sourcebook. Resources for Conflict Intervention Practitioners and Programs*, J. Folger, R. A. B. Bush & D. J. D. Noce, Ed., Hempstead, NY: Institute for the Study of Conflict Transformation, Inc., 2010, 31-50.

[6] J. Costa, M. Jung, M. Czerwinski, F. Guimbretiere, T. Le, and T. Choudhury, "Regulating Feelings During Interpersonal Conflicts by Changing Voice Self-perception", *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems,* 631, 2018.

[7] L. J. Cronbach "Coefficient alpha and the internal structure of tests", *psychometrika*, 16.3, 297-334, 1951.

[8] F. Eyben, F. Weninger, F. Gross, and B. Schuller, "Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor", In Proc. ACM Multimedia, ACM 2013.

[9] J. P. Folger, "Transformative Mediation and the Courts: A Glimpse at Programs and Practice." in *Transformative Mediation: A Sourcebook. Resources for Conflict Intervention Practitioners and Programs*, J. Folger, R. A. B. Bush & D. J. D. Noce, Ed., Hempstead, NY: Institute for the Study of Conflict Transformation, Inc., 2010, 165-180.

[10] G. Hoffman, O. Zuckerman, G. Hirschberger, M. Luria and T. Shani Sherman, "Design and evaluation of a peripheral robotic conversation companion", *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 3-10 2015.

[11] K. A. Jehn, "A multimethod examination of the benefits and detriments of intragroup conflict", *Administrative science quarterly*, 256-282, 1995.

[12] K. A., Jehn and E. A. Mannix, "The dynamic nature of conflict: A longitudinal study of intragroup conflict and group performance", *Academy of management journal*, *44*.2, 238-251. 2001.

[13] M.F. Jung, N. Martelaro, and P. J. Hinds, "Using robots to moderate team conflict: the case of repairing violations", *In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 229-236. ACM, 2015.

[14] M. F. Jung, "Coupling interactions and performance: Predicting team performance from thin slices of conflict", *ACM Transactions on Computer-Human Interaction*, 23.3: 18, 2016.

[15] D. E. King, "Dlib-ml: A Machine Learning Toolkit", *Journal of Machine Learning Research*, 10, 1755-1758, 2009.

[16] H. Levene, In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, I. Olkin et al. eds., Stanford University Press, pp. 278-292, 1960.

[17] H. B. Mann and D. R. Whitney, "On a test of whether one of two random variables is stochastically larger than the other", *Annals of Mathematical Statistics*, 18, 50-60, 1947.

[18] R. G. Miller, *Simultaneous Statistical Inference*. Springer, 1966.

[19] J. K. Moen, D. T. Hudson, J. R. Antes, E. O. Jorgensen, and L. H. Hendrikson, "Identifying Opportunities for Empowerment and Recognition in Mediation" in *Designing mediation: Approaches to training and practice within a transformative framework*, J. Folger & R. A. B. Bush, Ed., Hempstead, NY: Institute for the Study of Conflict Transformation, Inc., 2001, 112-132.

[20] D. J. D. Noce, "Shifts, Fractures, and Supports: A Communication Perspective on Conflict Transformation" in *Transformative Mediation: A Sourcebook. Resources for Conflict Intervention Practitioners and Programs*, J. Folger, R. A. B. Bush & D. J. D. Noce, Ed., Hempstead, NY: Institute for the Study of Conflict Transformation, Inc., 2010, 145-159.

[21] M. J. Pettinati and R. C. Arkin. "A Robot to Provide Support in Stigmatizing Patient-Caregiver Relationships." *The Thirty-First International Flairs Conference*. 2018.

[22] S. M. Retzinger, *Violent emotions: Shame and rage in marital quarrels*. Sage Publications, 1991.

[23] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)", *Biometrika*, *52*(3/4), 591-611, 1965.

[24] S. Shen, P, Slovak, and M. F. Jung. "Stop. I See a Conflict Happening.: A Robot Mediator for Young Children's Interpersonal Conflict Resolution." *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, 2018.

[25] J. Simand and C. C. Wright. "The kappa statistic in reliability studies: use, interpretation, and sample size requirements." *Physical therapy* 85.3, 257-268, 2005.

[26] L. Tickle-Degnen and K. D. Lyons, "Practitioners' impressions of patients with Parkinson's disease: the social ecology of the expressive mask", *Social Science & Medicine*, 58(3): 603-614, 2004.