

Interpersonal Coordination of Head Motion in Distressed Couples

Zakia Hammal, Jeffrey F. Cohn, and David T. George

Abstract—In automatic emotional expression analysis, head motion has been considered mostly a nuisance variable, something to control when extracting features for action unit or expression detection. As an initial step toward understanding the contribution of head motion to emotion communication, we investigated the interpersonal coordination of rigid head motion in intimate couples with a history of interpersonal violence. Episodes of conflict and non-conflict were elicited in dyadic interaction tasks and validated using linguistic criteria. Head motion parameters were analyzed using Student's paired t-tests; actor-partner analyses to model mutual influence within couples; and windowed cross-correlation to reveal dynamics of change in direction of influence over time. Partners' RMS angular displacement for yaw and RMS angular velocity for pitch and yaw each demonstrated strong mutual influence between partners. Partners' RMS angular displacement for pitch was higher during conflict. In both conflict and non-conflict, head angular displacement and angular velocity for pitch and yaw were strongly correlated, with frequent shifts in lead-lag relationships. The overall amount of coordination between partners' head movement was more highly correlated during non-conflict compared with conflict interaction. While conflict increased head motion, it served to attenuate interpersonal coordination.

Index Terms—Interpersonal coordination, head motion, windowed cross-correlation, actor-partner analysis

1 INTRODUCTION

THE last two decades have witnessed significant progress in automatic analysis and recognition of human emotion. Until recently, most research in affective computing has focused on facial behavior [68] where the confounding effects of head movement are removed prior to analysis [61]. Although in everyday experience we encounter associations between head movement and emotional expression [29], [13], the communicative functions of head movements in relation to emotion have been fairly unexplored in affective computing. Some notable exceptions are Paterson et al. [45] and Gunes and Pantic [24], [55], who found meaningful correlations between head movement and perceived dimensions of emotion in point-light animations [45] and in intelligent virtual agents [55], Busso and et al. [8] who considered the relationship between head motion and emotional speech in facial animation, De Silva and Bianchi-Berthouze [17] and el Kaliouby and Robinson [20] who used head movement among other features to recognize emotional expressions or mental states, and more recently Xiao et al. [66] who used head movement to categorize a subset of communicative labels in the context of psychotherapy interviews. In all but a few cases, such as Sidner et al. [59] and Morency et al. [42]

who considered robot-human coordination of head motion, interpersonal coordination of head motion in social interactions has not been considered in affective computing.

By contrast, the literature on paralinguistic communication and social interaction emphasizes the importance of head movement in communication. Head movements regulate turn-taking [19], serve back-channeling functions [31], and communicate messages such as agreement or disagreement and interpersonal interaction. Head movement also provides useful supplementary information for perceiving speech [43], differentiating statements from questions [22], [44], and determining word emphasis in phrases [52]. Head movement during speaking, singing, and piano performances conveys a range of meanings and emotions and allows discriminating between often subtle expressive intentions [16]. When Duchenne smiles are coordinated with downward head pitch and head turn, they communicate embarrassment rather than enjoyment [13].

Individual differences, such as gender and expressiveness, influence head nods and turns [7]. For example, women use more active head motion among each other than do men when they talk with each other [6]. A person speaking with a woman tends to nod his or her head more than when speaking with a man [2]. Given the importance of head motion in human communication, it is critical to include the dynamics of head movement and interpersonal coordination in automatic facial expression analysis.

Efforts in this direction have begun. One approach is to observe a participant interacting with a virtual agent or animated avatar via an audio-visual link. Pfeiffer et al. [48] used an interactive eye-tracking paradigm in which participants interacted with a virtual agent. The virtual agent's gaze was manipulated so that it was either averted from or coordinated with that of the participant. The authors investigated the influences of these variations on whether

- Z. Hammal is with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15260. E-mail: zakia_hammal@yahoo.fr.
- J.F. Cohn is with the Robotics Institute, Carnegie Mellon University and the Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260. E-mail: jeffc@cs.cmu.edu.
- D.T. George is with the Division of Intramural Clinical and Biological Research, National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Bethesda, MD. E-mail: tedg@lcs.niaaa.nih.gov.

Manuscript received 31 Dec. 2013; revised 13 May 2014; accepted 14 May 2014. Date of publication 29 May 2014; date of current version 23 July 2014.

Recommended for acceptance by U. Hess.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.

Digital Object Identifier no. 10.1109/TAFFC.2014.2326408

participants believed that another person or a computer program controlled the virtual agent. Bailenson and Yee [3] used a controlled virtual environment to study the influence of mimicry on cooperative behavior. Von der Pütten et al. [64] investigated whether participants' belief that they were interacting with a virtual representation of a human or an agent led to different social behaviors. Using an approach developed by Theobald et al. [60], Boker et al. [7] used motion-tracked synthesized avatars to randomly assign apparent identity and gender during free conversation in a videoconference, thereby dissociating sex from motion dynamics. They tested the relative contribution of apparent identity and head motion dynamics on social perception and behavior. All these studies involved communication between previously unacquainted participants over an audio-visual computer link.

A more challenging task is to study direct face-to-face interaction between participants in relatively natural communication. Three groups have addressed this problem by observing direct face-to-face interaction between previously unacquainted participants or between mothers and their infants. Boker et al. [5] used motion-capture technology to investigate the influence of auditory noise on interpersonal coordination of head motion velocity in adult face-to-face interaction. Messinger et al. [39] used active appearance models and automated measurement of expression intensity to model the face-to-face dynamics of mothers and their infants. Hammal et al. [25] used the results of automatic head tracking to investigate the contribution of head motion to emotion communication and its dynamic coordination in a face-to-face social interaction between parents and infants.

In continuity with these initial efforts and as an initial step toward understanding the contribution of head motion to emotion communication, we investigated interpersonal coordination of rigid head motion in face-to-face interactions between intimate partners during conflict and non-conflict. A well-validated dyadic interaction task was used to elicit conditions of conflict and non-conflict [51] in intimate couples that had a history of interpersonal violence. One of the partners had recently begun inpatient treatment for anger control in relation to their partner and substance dependence. Analyzing pre-existing couples of this sort, instead of solitary participants interacting with a virtual agent or previously unacquainted participant, posed both advantages and challenges. First, intimate partners are more likely to express their emotions in an uninhibited manner [51]. In the couples with a history of violence toward each other that we studied, extreme anger and violence were ever-present possibilities. We hypothesized that head displacement and velocity would increase in conflict compared with non-conflict. Second, because the behaviors are interpersonal, each participant's behavior is both cause and consequence of the other's behavior. Data analytic approaches that are appropriate for mutual influence are needed. Actor-Partner analysis and windowed cross-correlation (WCC) were used to measure the reciprocal relationship between the partners. We hypothesized that patterns of symmetry and asymmetry may shift rapidly as one or the other partner becomes momentarily dominant.

2 METHODS

The following sections describe the participants, the observational and recording procedures, the linguistic coding of the audiovisual recordings, and the procedure used to automatically measure head pose.

2.1 Participants

Twenty-seven (10 women and 17 men) heterosexual intimate partners participated in a dyadic interaction task. One member of each couple was enrolled in inpatient treatment for anger control and alcohol dependence [23]. These members, referred to below as inpatient participants, had committed physical aggression toward their partners on at least two separate occasions within the past year, at least one of which was under the influence of alcohol, and had scored 3 or higher on the physical violence subscale of the Straus Conflict Tactics Scales, a well-validated instrument for assessing intimate conflict [58]. Informed consents were obtained from both partners to be video-recorded during a dyadic interaction task. The average age of participants and their intimate partners was 40.8 years (range: 23-57). Fifty-two percent self-identified as White, 44 percent as Black, 4 percent as Hispanic, and 4 percent as other.

2.2 Observational Procedure

The inpatient participants and their partners participated in a dyadic interaction task within two weeks of the beginning of the inpatient participant treatment. They were seated on opposite ends of a couch with an experimenter seated across the room in front of them [23]. The experimenter instructed them to engage in two tasks. First, they were asked to discuss something neutral, such as how the partner had traveled to the treatment center (low conflict). Next, they were asked to "discuss something that has been difficult for the two of you" (high conflict). The order of the two tasks was the same for all the dyads. Couples were asked to address each other and to ignore the experimenter. The interactions were recorded using two hardware-synchronized VGA cameras and a microphone that were input to a special effects generator that created a split-screen recording. The sessions lasted an average of 16.68 min (range: 6.07-46.54 min). The video captured their faces and upper chest and shoulders. Fig. 1 shows an example of the obtained split-screen for one of the 27-couples. To protect the identity of the participants, their faces have been digitally masked.

2.3 Segmentation of Conflict and Non-Conflict Segments

Despite the instructions to the couples, high conflict could and did occur during the nominally low-conflict task. To assess this possibility, manual linguistic coding was used to identify segments of actual low and high conflict throughout the recording period. The linguistic criteria were inspired by previous research. Dunn and Munn [18] used behavioral and linguistic criteria to investigate conflict episodes in sibling interaction. Caughy et al. [9] did the same in mother-child interactions. Because our goal was to understand nonverbal behavior during conflict episodes, we used linguistic criteria to segment conflict and non-conflict



Fig. 1. Example of split-screen video. Images are de-identified to protect participants' identity.

segments and analyzed head movement to characterize corresponding behavior.

2.3.1 Segmentation

Conflict was defined as one partner opposing or contradicting the other. Examples included disagreement, hostile behavior, and deliberate attempts to contradict, upset, or provoke the partner, such as: “no”, “that’s a lie”, “you are wrong”, “never”. Further examples included contradicting statements, such as: “I was not drunk, you had 12 beers”; “Amy was there, she had left already”. Conflict segments ended when a) one member of the dyad submitted to the other, b) the couple reached consensus or changed topics, c) the interviewer intervened between them, or d) one of the partners directed their remarks toward the interviewer using third-person language. Excluded from conflict were segments in which one partner initiated a conflict but the other complied or ignored without reciprocating or exhibiting negative affect.

Non-conflict segments were defined as segments in which the couple engaged in conversation with each other and used second-person language when they referred to each other.

Experimenter segments were those in which either the experimenter addressed one or both partners or one or both partners spoke to the experimenter.

2.3.2 Inter-Observer Reliability

All coding was performed by a single trained annotator. To assess the reliability of the coding, 15 percent of the videos were randomly sampled and independently labelled by another trained annotator. Agreement was measured using coefficient kappa. The mean kappa for interobserver agreement was 0.75, which suggests substantial inter-observer agreement [35], [63].

2.4 Automatic Head Tracking

A cylinder-based 3D head tracker (CSIRO tracker) [15] was used to model the 6 degrees of freedom of rigid head motion (i.e., pitch, yaw, roll, scale, translation y (up/down), and translation x (left/right)), which refers to head movement in a three-dimensional space. For each participant, the tracker was initialized on a manually selected near-frontal image prior to tracking. To assess the concurrent validity of the tracker for the tracking of head orientation, it was compared with two alternative methods on an independent data set. One method was Xiao, Kanade, and Cohn’s generic cylindrical head tracker [65]; the other was a person-specific 2D+3D AAM [27]. Test data were 18 minutes of video from three participants in the Sayette Group Formation Task (GFT) data set [53]. The GFT participants were recorded while seated around a circular table, talking, and consuming drinks containing alcohol or a placebo beverage. Moderate to large head motion and occlusion occurred frequently over the course of these spontaneous social interactions. The CSIRO tracker demonstrated high concurrent validity with both of the other trackers. In comparison with [27], concurrent correlations for roll, pitch, and yaw were 0.78, 0.96, 0.93, respectively. In comparison with [65], the corresponding values were 0.77, 0.93, and 0.94. In a recent behavioral study, the CSIRO tracker has differentiated normal and perturbed parent-infant interaction [25]. In the current study, the tracker covered the wide range of head pose variation that can occur in social interaction within couples and proved robust to both minor occlusion and rapid changes in pose. An example of the tracker’s performance is shown in Fig. 2.

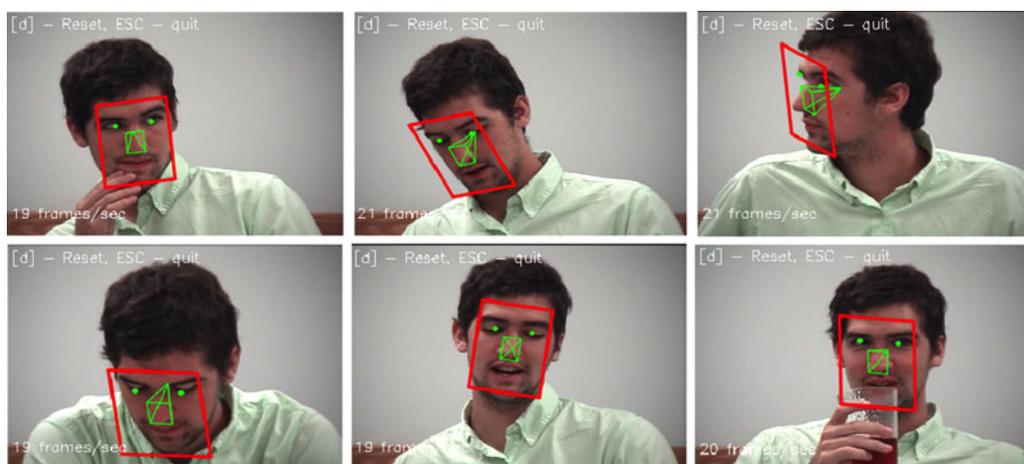


Fig. 2. Examples of head tracking during moderate to large head motion and occlusion. The examples are from Sayette Group Formation Task data set [53].

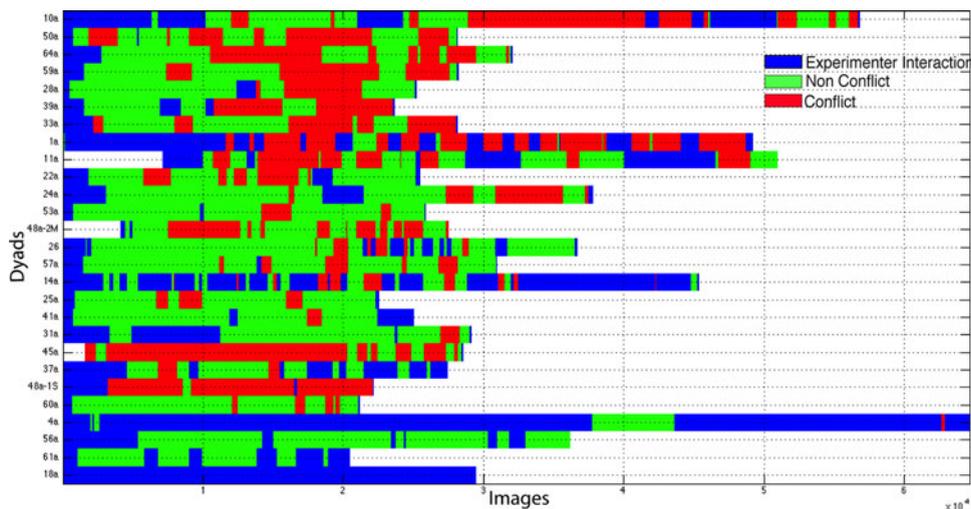


Fig. 3. Interaction map based on linguistic coding. Each row corresponds to one couple. Columns correspond to the results of linguistic coding across frames. Blue delimits interaction with the experimenter, and red and green delimit conflict and non-conflict, respectively.

3 DATA SELECTION AND REDUCTION

In the following sections we first present results for linguistic coding and automatic head tracking. We then describe how angular displacement and angular velocity of head motion were computed.

3.1 Linguistic Coding Results

Fig. 3 shows the interaction map obtained by linguistic coding (see Section 2.3). For all couples, interactions began in conversation with the experimenter (delimited in blue on Fig. 3). In a few cases, couples never progressed beyond this. As the goal of our study was to compare the head movements in conflict and non-conflict segments, segments involving the experimenter were not analyzed. In most cases, segments with and without conflict (delimited in red and green, respectively on Fig. 3) alternated throughout the interaction regardless of the experimental instructions. This pattern highlights the importance of using independent criteria to identify episodes of conflict and non-conflict instead of relying upon the experimenter's instructions. Conflict was interspersed frequently throughout the sessions regardless of instructions. The ratio of conflict to non-conflict was about 2:1. Conflict occurred 49.63 percent of the time, and non-conflict occurred 25.66 percent of the time ($t = 2.94$, $df = 26$, $p = 0.006$).

3.2 Head Tracking Results

For each video frame, the tracker output six degrees of freedom of rigid head motion (i.e. pitch, yaw, roll, translation x ,

translation y , and scale) or a failure message when a frame could not be tracked (see Fig. 2). 4.39 percent of the total frames could not be tracked. To evaluate the quality of the tracking, we visually reviewed the tracking results overlaid on the video (Fig. 2). In 17 percent of the tracked video frames, visual review indicated errors. These frames were excluded from further analyses. Table 1 reports the proportions of tracked frames that met visual review for each participant separately and for each couple (valid tracking for both partners in the same video frame) for the three different interaction episodes. Valid tracking was highest for the experimenter interaction. The proportion of valid frames in conflict and non-conflict did not vary whether for individual participants or for couples. Fig. 4 shows the distribution of valid tracking for each couple.

3.3 Data Selection

Within couples, the proportion of valid frames for conflict and non-conflict varied and for some couples was low. So that imbalance between conflict and non-conflict or low base rate would not bias test statistics, two criteria for data selection were followed. One, the same duration was sampled for both conflict and non-conflict (Fig. 5 for an example). Two, only segments greater than 30 seconds were considered. "Thin slices" less than this amount are believed to be insufficient to permit valid inferences about personality-based characteristics [1]. Using the criteria of a minimum duration of half a minute for both conflict and non-conflict, 17 dyads corresponding to 34 participants had complete data (see Fig. 6). We report on their data in the following.

TABLE 1
Proportion of Valid Tracking during Experimenter Interaction, Non-Conflict and Conflict Interaction

	Proportion of good tracking		
	Experimenter Interaction	Non-Conflict	Conflict
Participants	69.34% ^a	65.81% ^{ab}	62.79% ^b
Couples	56.45% ^a	50.73% ^{ab}	38.03% ^b

Note: Within rows, means with different superscripts differ at $p \leq 0.05$.

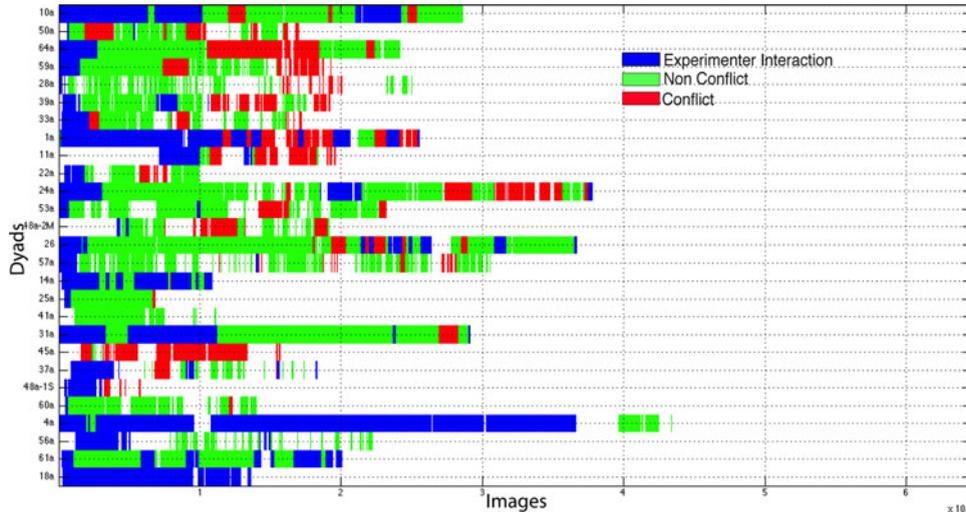


Fig. 4. Interaction map for good tracked frames. Each row corresponds to one couple and the corresponding columns correspond to the time of the interaction.

3.4 Data Reduction: Head Movement Measurement

Angles of the head in the horizontal and vertical directions were selected to measure head movement. These directions correspond to the meaningful motion of head nods (i.e., pitch), and head turns (i.e., yaw), respectively (see Fig. 7). Head angles were converted into angular displacement and angular velocity. For both pitch and yaw, angular displacement was computed by subtracting the overall mean head angle from each observed head angle within each valid segment (i.e., consecutive valid frames). We used the overall mean head angle, which afforded an estimate of the overall head position for each partner in each condition and each segment. Similarly, for both pitch and yaw, angular velocity was computed as the derivative of angular displacement, measuring the speed of changes of head movement from one frame to the next. So that missing data would not bias measurements, the angular displacement and angular velocity for each episode and each subject were computed for each consecutive valid segment separately. The root mean square (RMS) was then used to measure the magnitude of variation of the angular displacement and the angular velocity, respectively. The RMS value of the angular displacement and angular velocity was computed as the square root of the mean value of the squared values of the quantity taken over a segment. The RMS of the horizontal (RMS–yaw), and vertical (RMS–pitch) angular displacement and angular velocity were then calculated for each condition of dyadic interaction for each separate partner.

To minimize missing data, the RMSs of the angular displacement and angular velocity were computed for each selected valid segment and normalized by the duration of the segment. They are denoted nRMS. Then for each dyad the mean of the consecutive nRMSs over the selected segments (for angular displacement and angular velocity, respectively) was computed and used for analyses. Fig. 8 shows an example of selected conflict segments taken from Fig. 5. In this example, to measure the RMS of head movement (displacement and velocity) in the conflict condition we first computed the nRMSs of angular displacement (respectively angular velocity) for each consecutive segment: nRMS1, nRMS2 and nRMS3. The mean RMS of the angular displacement (respectively angular velocity), denoted mRMS, of the cumulative segments is then computed as the mean of these three normalized RMSs as: $mRMS = (nRMS1 + nRMS2 + nRMS3)/3$. The obtained mRMSs (for angular displacement and angular velocity, respectively) are used as the characteristic measure of head movement and are used in all the subsequent analyses. In the following sections mRMS are noted RMS for simplicity.

4 DATA ANALYSES

We first evaluated whether changes (increase or decrease) of head angular displacement and angular velocity are characteristic of conflict compared with non-conflict for men and for women and for inpatient participants and their partners. We then used Actor-Partner analysis to evaluate how the partners affect each other’s head movements. Finally, time series analysis was used to measure the pattern of synchrony in head movement between the partners and whether this pattern changes between conflict and non-conflict.

4.1 Do Head Movements of Distressed Couples Differ During Conflict and Non-Conflict?

Because of the repeated-measures nature of the data, Student’s paired t-tests were used to test differences in head movement between conflict and non-conflict. The RMSs of the angular displacement and angular velocity for pitch and

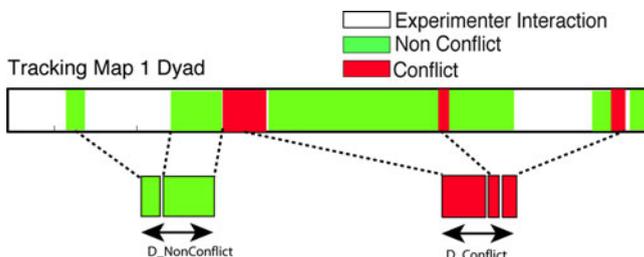


Fig. 5. Event selection after tracking. D_Conflict, corresponds to the duration of the concatenation of the conflict segments; D_NonConflict is equal to D_Conflict and corresponds to the duration of the concatenation of consecutive non-conflict segments to equal D_Conflict.

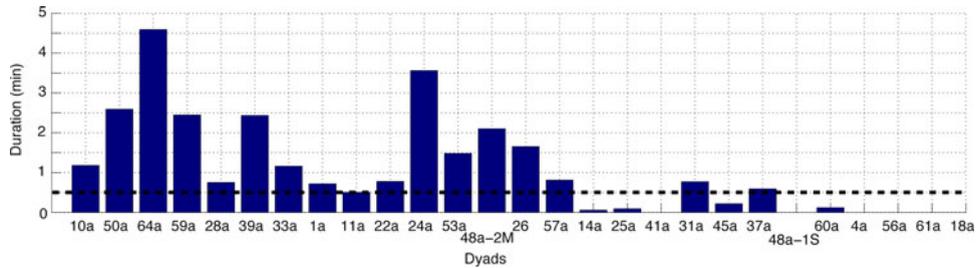


Fig. 6. Maximum duration per dyad after the selection process. The dashed line corresponds to the threshold over which the data was used for subsequent analyses.

yaw were the primary dependent measures. First, within-subject comparisons evaluated the variability of each participant's angular displacement and angular velocity of head movement in conflict compared with non-conflict. A Student's paired *t*-test on the RMS angular displacement and RMS angular velocity of pitch and yaw was computed between conflict and non-conflict, separately for men and women. Second, within-dyad comparisons evaluated differences between inpatient participants' and partners' RMS angular displacement and RMS angular velocity of pitch and yaw separately for conflict and non-conflict.

4.2 How Do Head Movements of Men and Women, Inpatient Participants and Partners Compare?

We expect that each inpatient participant affected the head movements of their partner as well as their partner affected the head movements of the inpatient participant. Given these bi-directional feedback effects, the data need to be analyzed while taking into account both participants in the dyad simultaneously. Each participant's head movements are thus both a predictor and an outcome variable.

To put both partners in a dyad into the same analysis, we used a variant of Actor-Partner analysis [30], [7]. For example, for the analysis of the RMS-pitch angular displacement, we put both participants RMS' pitch angular displacement into the same column in the data matrix; and used a second column as a dummy code gender to identify whether the data in the angular RMS angular displacement column came from a man or a woman. In a third column, we put the RMS-pitch angular displacement from the other participant. We then used the terminology Actor and Partner to distinguish which variable is the predictor and which is the outcome for a selected row in the data matrix. If gender = 1, then the man is the Actor and the woman is the Partner in that row of the data matrix. If gender = 0, then the woman is the Actor and the man is the Partner. The RMS-pitch

angular displacement of the Partner is used as a continuous predictor variable. Binary variables were used to code the partners' status (1 = inpatient participant, 0 = partner) and the conflict condition (1 = conflict, 0 = non-conflict) [30], [7]. We also added an interaction variable (partner status X conflict condition) that allows us to account for an overall effect of the partner status (inpatient participant or their partner) as well as the condition of the interaction (conflict or non-conflict) between the partners. Then models were fit using R (R Core Team, 2013) [49] and *nlme* package [47] to perform linear mixed effects analysis of the relationship between the head movements of the partners. To do so, we used the R generic function `lme()` that fits a linear mixed effects model in the formulation described in Laird and Ware [32]. In our case this model can be expressed as:

$$y_j = b_{j0} + b_1 S_j + b_2 PR_j + b_3 I_j + b_4 PR_j I_j + b_5 PY_j, \quad (1)$$

where y_j is the outcome variable (either RMS-pitch, or RMS-yaw angular displacement or velocity) for dyad j . The other predictor variables are the sex of the Actor S_j , whether the Actor is the inpatient participant PR_j , the interaction condition I_j (conflict or non-conflict), and the interaction between the partner status and the conflict condition $PR_j I_j$, the RMS-pitch, or RMS-yaw angular displacement or angular velocity of the partner PY_j .

4.3 Is Head Movement between Partners Synchronous?

To investigate synchrony, we used windowed cross-correlation and peak picking. Windowed cross-correlation [33], [34] is a method of time-series analysis appropriate when synchrony between two series varies over time [2]. Peak picking is applied to WCC to better reveal pattern of change in lead-lag relationship between series [5].

4.3.1 Windowed Cross-Correlation

The windowed cross-correlation estimates time varying correlations between signals [33], [34]. It produces positive or

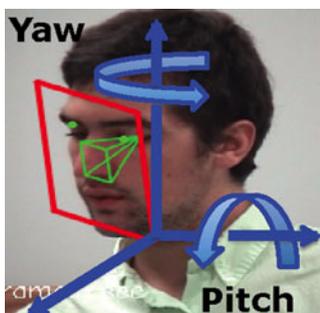


Fig. 7. Illustration of pitch and yaw angles.

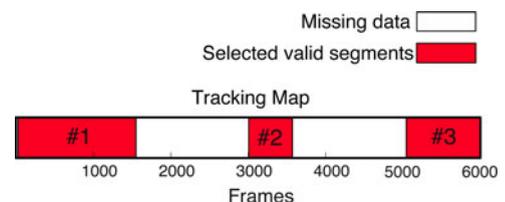


Fig. 8. Red segments correspond to valid conflict segments; white segments correspond to missing data or tracking failure.

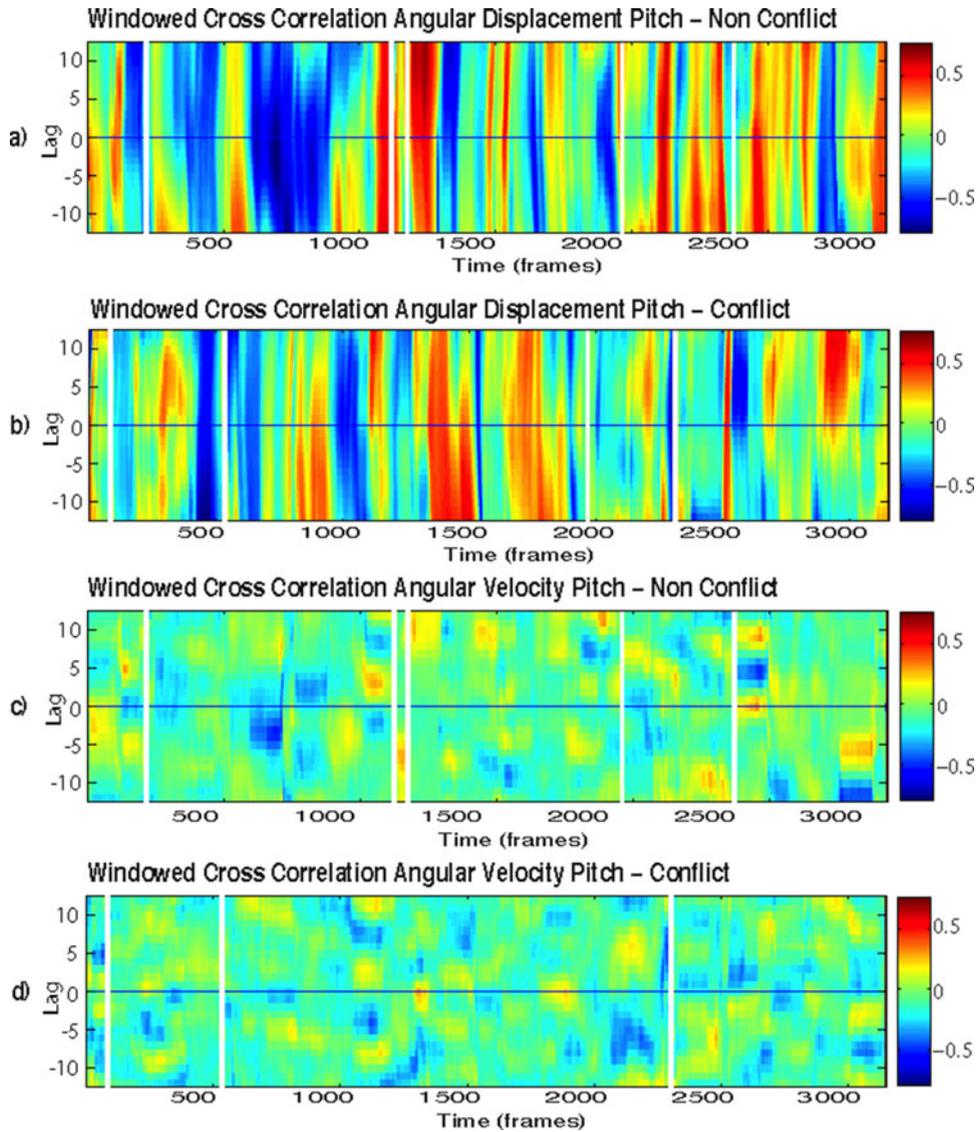


Fig. 9. Examples of windowed cross-correlation matrices for pitch angular displacement (a-b) and angular velocity (c-d) for conflict and non-conflict.

negative correlation values between two signals over time for each (time, lag) pair of values. The windowed cross-correlation assumes that the mean and variance of time series and patterns of dependency between them may change over time. It uses a temporally defined window to calculate successive local zero-order correlations over the course of an interaction. In Boker et al. [5], for example, the windowed cross-correlation was used to compare the frequency of head and hand motion and their synchrony between conversing dyads with and without street noise. In Messinger et al. [39], the windowed cross-correlation was used to measure local correlation between infant and mother during smiling activity over time. Following these previous studies, windowed cross-correlation was used to measure the associations between previous, current, and future head movements of partners. The windowed cross-correlation consists in estimating the successive correlation values and the corresponding time lag from the beginning to the end of the interaction (see Fig. 9). It requests the definition of a set of parameters: the sliding windows size (W_{max}), the window increment (W_{inc}), maximum lag (t_{max}) and the lag increment (t_{inc}). The results matrix will have a number of columns equal to

$(t_{max} * 2) + 1$ and a number of rows equal to the largest integer less than $(N - W_{max} - t_{max}) / W_{inc}$. The choice of parameter values was guided by previous studies and by preliminary inspection of the data. Previous works [5], [50] showed that plus or minus 2s is meaningful for the production and perception of head nods (i.e. pitch) and turns (i.e. yaw). We therefore set window size $W_{max} = 4$ seconds. The maximum lag t_{max} is the maximum interval of time that separates the beginning of the two windows. The maximum lag should allow the capture of the possible lags of interest and the exclusion of lags of no interest. After exploring different lag values, we chose a maximum lag $t_{max} = 1/2$ seconds $\times 30$ sample/sec = 15 samples. Using these parameter values, the windowed cross-correlation was computed for the angular amplitude and the angular velocity of pitch and yaw, for all the dyads during conflict and non-conflict, and for each consecutive valid segment (see Section 3.4).

4.3.2 Peak Picking

To analyze patterns of change in the obtained cross-correlation matrix, Boker et al. [5] proposed a peak selection algorithm that selects the peak correlation at each elapsed

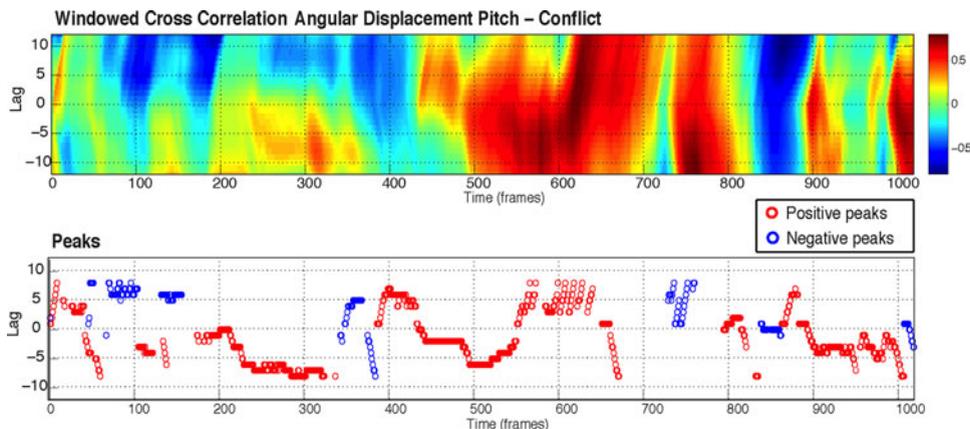


Fig. 10. For the windowed cross-correlation plot in the upper panel, we computed correlation peaks using the peak-picking algorithm. The peaks are shown in the lower panel. Red indicates positive correlation; blue indicates negative correlation.

time according to flexible criteria. The peaks are defined so that each selected peak has the maximum value of cross-correlation centered in a local region where the other values are monotonically decreasing through each side of the peak. The second row of Fig. 10 shows an example.

5 RESULTS

To answer the above questions, three sets of analyses were pursued and presented in the following sections.

5.1 Head Movement Characteristics in Conflict Compared with Non-Conflict

Student's paired *t*-tests revealed that men's RMS angular displacement and RMS angular velocity of pitch and yaw increased significantly in the conflict condition compared with the non-conflict condition (see Tables 2 and 3). For women, RMS angular velocity of pitch and yaw but not amplitude increased in conflict (see Table 3). Overall, both men and women moved their heads faster in conflict than in non-conflict, while only men increased RMS angular displacement of pitch and yaw during conflict. No significant differences were found between inpatient participants and partners for RMS angular displacement or velocity of pitch or yaw during conflict and non-conflict.

5.2 Actor-Partner Results

Table 4 presents the results of the mixed effects random intercept models for the RMS angular displacement of pitch

and yaw, respectively. The intercepts, $7.404e-05$ radian/sec ($p = 0.03$) and $6.894e-05$ radian/sec ($p = 0.047$) correspond to the overall mean vertical RMS angular displacement of both partners for pitch and yaw, respectively. The lines "Partner RMS PitchAmp" and "Partner RMS YawAmp" estimate the reciprocal relationship between the partners' RMS-Pitch and RMS-Yaw angular displacements, respectively. The obtained results show that there was a reciprocal relationship in the partners' RMS-yaw angular displacements. For each 1 radian/sec increase in the partners RMS-yaw, there was a 0.273 radian/sec increase in the RMS-yaw angular displacement of the other partner ($p = 0.029$). No significant reciprocal relationship was found between the RMS-pitch angular displacement of the partners. The lines "Actor is Male" suggest that the males produced RMS-pitch and RMS-yaw angular displacements that were not significantly different from those of females. The lines "Actor is Inpatient participant" suggest that the inpatient participants produced RMS-pitch and RMS-yaw angular displacements that were not significantly different from those of their partners. The lines "Conflict" evaluate if partners angular displacements change in conflict compared with non-conflict conditions. The obtained results suggest that the RMS-pitch angular displacement is $7.671e-05$ radian/sec higher in conflict compared with non-conflict ($p = 0.046$). The lines "Actor is Inpatient participant X Conflict" suggest that there is no significant interaction effect between inpatient participants and the conflict condition. The results of the mixed effects random intercept models

TABLE 2
Within-Subject Effects for Pitch and Yaw RMS Angular Displacements for Men and Women during Conflict and Non-Conflict

	Paired t-test		
	t	df	p
Pitch			
Men	-2.35	16	0.031
Women	-1.49	16	0.155
Yaw			
Men	-2.21	16	0.041
Women	-1.74	16	0.099

Note: *t*: *t*-ratio, *df*: degrees of freedom, *p*: probability.

TABLE 3
Within-Subject Effects for Pitch and Yaw RMS Angular Velocities for Men and Women during Conflict and Non-Conflict

	Paired t-test		
	t	df	p
Pitch			
Men	-2.74	16	0.014
Women	-2.26	16	0.037
Yaw			
Men	-2.78	16	0.013
Women	-2.34	16	0.032

Note: *t*: *t*-ratio, *df*: degrees of freedom, *p*: probability.

TABLE 4
Pitch and Yaw RMS Angular Displacement Predicted Using a Linear Mixed Effects Analysis

Pitch	Value	SE	df	t	p
(Intercept)	7.404e-05	3.326e-05	46	2.226	0.030
Partner RMS PitchAmp	0.098	0.126	46	0.778	0.440
Actor is Male	2.081e-05	2.764e-05	46	0.753	0.455
Actor is Inpatient participant	1.877e-05	3.815e-05	46	0.492	0.625
Conflict	7.671e-05	3.755e-05	46	2.043	0.046
Actor is Inpatient participant X Conflict	-4.517e-05	5.289e-05	46	-0.854	0.397
Observations = 68,	Groups = 17,	AIC = -933.9,	BIC = -916.8.		
Yaw					
(Intercept)	6.894e-05	3.375e-05	46	2.042	0.047
Partner RMS YawAmp	0.273	0.122	46	2.240	0.029
Actor is Male	-1.43e-05	2.951e-05	46	-0.285	0.776
Actor is Inpatient participant	3.694e-05	4.096e-05	46	0.901	0.371
Conflict	5.216e-05	4.048e-05	46	1.288	0.204
Actor is Inpatient participant X Conflict	-1.343e-05	5.640e-05	46	-0.238	0.812
Observations = 68,	Groups = 17,	AIC = -931.3,	BIC = -914.2.		

"Inpatient participant" is the member of the couple that was in treatment for anger control and substance dependence.
Note: RMS: Root Mean Square, SE: Std Error, t: t-ratio, df: degrees of freedom, p: probability.

applied to RMS angular velocity of pitch and yaw respectively are presented in Table 5. The same pattern of reciprocal interaction effect between the partners head movements were found for the RMS-pitch and the RMS-yaw angular velocity. The lines "Partner RMS PitchVel" and "Partner RMS YawVel" show that for each 1 radian/sec increase in the partner's RMS-pitch and RMS-yaw angular velocities there was respectively 0.724 radian/sec and 0.641 radian/sec increase in the RMS-pitch and RMS-yaw angular velocities of the other partner ($p = 0.000$).

Overall, the obtained results highlight a significant reciprocal relationship between the partners' head movements for the angular displacements of pitch and the angular velocities of pitch and yaw. Thus over the dyadic interaction, greater than normal head movement by one partner is associated with greater than normal head movement by the

other partner. This effect could be attributed to symmetry or mimicry between partners. Finally, there was evidence that partners' pitch angular displacements increase when partners are in conflict interaction compared with non-conflict. While pitch angular displacement was not associated with a significant reciprocal increase or decrease between partners' head movements, pitch angular displacement appears to be a characteristic head behavior during conflict interaction.

5.3 Head Movement Coordination in Conflict Compared with Non-Conflict

5.3.1 Windowed Cross-Correlation Patterns

Fig. 9 shows examples of windowed cross-correlation for pitch angular displacement and angular velocity for one

TABLE 5
Pitch and Yaw RMS Angular Velocity Predicted Using a Linear Mixed Effects Analysis

Pitch	Value	SE	df	t	p
(Intercept)	4.60e-06	3.99e-06	46	1.153	0.254
PartnerRMS PitchVel	0.724	0.087	46	8.282	<0.001
Actor is Male	8.00e-06	3.59e-06	46	0.236	0.814
Actor is Inpatient participant	-0.10e-06	4.96e-06	46	-0.026	0.979
Conflict	6.80e-06	4.93e-06	46	1.388	0.171
Actor is Inpatient participant X Conflict	-6.80e-06	6.86e-06	46	-0.997	0.323
Observations = 68,	Groups = 17,	AIC = -1196.1,	BIC = -1179.1.		
Yaw					
(Intercept)	6.5e-06	4.87e-06	46	1.338	0.187
PartnerRMS YawVel	0.641	0.097	46	6.592	<0.001
Actor is Male	-0.500e-06	4.42e-06	46	-0.117	0.907
Actor is Inpatient participant	1.00e-06	6.12e-06	46	0.162	0.871
Conflict	7.9e-06	6.10e-06	46	1.289	0.203
Actor is Inpatient participant X Conflict	-5.60e-06	8.46e-06	46	-0.660	0.512
Observations = 68,	Groups = 17,	AIC = -1169.9,	BIC = -1152.8.		

"Inpatient participant" is the member of the couple that was in treatment for anger control and substance dependence.
Note: RMS: Root Mean Square, SE: Std Error, t: t-ratio, df: degrees of freedom, p: probability.

dyad during conflict and non-conflict. In each of these graphs, the abscissa plots the elapsed time during the dyad, and the ordinate plots the lag of the two windows. The color represents the value of the cross-correlation. For each graph, the area above the midline of the plot ($Lag > 0$) represents the relative magnitude of correlations for which the head angular displacement (respectively head angular velocity) of one partner predicts the head angular displacement (respectively head angular velocity) of the other; the corresponding area below the midline ($Lag < 0$) represents the opposite. The midline ($Lag = 0$) indicates that both partners are changing their head angular displacements and velocities at the same time. Positive correlations (red) convey that head angular displacement (respectively angular velocity) of both partners is changing in phase. Negative correlation (blue) conveys that head angular displacement (respectively velocity) of both partners is changing out of phase.

The obtained cross-correlation graphs highlight vertical slices of positive (red plots) and negative (blue plots) correlations between the angular displacements of head movement of partners (Figs. 9a and 9b) and local and sparse periods of positive and negative correlations between the angular velocities of head movement of partners (Figs. 9c and 9d). Visual examination of Fig. 9 suggests that in both non-conflict and conflict tasks, head angular displacements and velocities were strongly correlated between partners with instability and dynamic exchange of the correlation signs (appearance of correlations with alternative positive and negative lags). This dynamic change of the direction of the correlations highlights the non-stationarity of the interaction between the partners. The interaction between partners changes over time (presence of correlation peaks at both negative and positive lag over time), consistent with the view that interpersonal communication is a dynamic system. Overall, at any given time men and women form a co-regulated dyadic head pose system. Each partner's head movements predict the other in time, and men to women influence is mirrored by women to men influence.

5.3.2 Peaks Cross-Correlation Patterns

The cross-correlation peak selections were run separately on the selected segments for conflict and non-conflict conditions. The second row of Fig. 10 shows an example of the peaks of correlation (red represents positive peaks and blue represents negative peaks) obtained using the peak picking algorithm on the windowed cross-correlation matrix reported on the first row of Fig. 10. The visual inspection of the obtained peaks shows the same pattern described above. The direction of the peaks correlation changes dynamically over time with frequent changes in which partner is leading the other. To compare coordination between partners' head movements in conflict and non-conflict conditions, the mean of the obtained correlation peak values were calculated for each dyad within each condition for each separate valid segment and normalized by the duration of the segment. In order to avoid bias due to low correlations, only peaks correlation values greater than 0.3 were considered for analyses. Correlations at this threshold or higher represent effect sizes of medium to large [12].

TABLE 6
Within-Dyad Effects for Normalized Mean of Peaks of Correlation Values for Angular Displacement and Angular Velocity of Pitch and Yaw during Non-Conflict and Conflict

Angular displacement	Paired t-test		
	t	df	p
Pitch	1.91	16	0.07
Yaw	1.84	16	0.08
Angular velocity			
Pitch	2.38	16	0.02
Yaw	2.15	16	0.04

Note: *t*: *t*-ratio, *df*: degrees of freedom, *p*: probability.

Student' paired t-tests were used to evaluate the variability of the correlations of the angular displacement and the angular velocity of head movement in conflict compared with non-conflict separately for each dyad. To do so, the obtained means of correlation were first Fischer *r*-to-*z* transformed and then used as the primary dependent measures in the Student' paired t-tests. The obtained results show that for the angular displacement, the means of peak correlation decreased marginally in conflict interaction for pitch and yaw (see Table 6). For the angular velocities, the means of peak correlation decreased significantly in conflict compared with non-conflict for pitch and yaw (see Table 6). Thus, partners' head movements were more highly correlated during non-conflict compared with conflict.

6 DISCUSSION

We used a dyadic interaction task paradigm to investigate the dynamics of head movement between intimate partners during conflict and non-conflict. In preliminary analyses, we found that both men and women responded with increased RMS angular velocity to conflict. Additionally, men responded with increased RMS angular displacement, as well. Actor-partner analyses revealed patterns of mutual influence and their relation to conflict. This study represents the first time that automated head tracking has been used to investigate mutual influence in intimate couples.

We found a significant reciprocal relationship between partners' head movements. During both conflict and non-conflict episodes, greater and faster head movements by one partner were associated with greater and faster head movements by the other partner. This finding of strong mutual influence for head movement was consistent with previous findings using manual measurement of perceived emotion in couples [51]. Levenson and Gottman [36], for instance, found that the display of more negative emotions from one spouse lead to the display of more negative emotions from the other spouse. A provocative exception in this literature is a recent study by Madhyastha et al. [37]. While most of the couples in their study showed mutual influence, a few were more influenced by their own initial state and inertia. Several factors might account for this difference between our findings and theirs. Madhyastha et al. [37] analyzed positive and negative valence rather than head motion, considered only conflict rather than both conflict

and non-conflict, and used separate models for each couple rather than a multilevel model. The modal pattern in their study, moreover, was mutual influence, which has been found by most investigators using manual measurement. Using automated measurement, mutual influence was our finding, as well.

Related literature has investigated mutual influence in unacquainted partners [2], [6]. Using motion capture devices and actor-partner analysis, Boker et al. found strong evidence of sex differences in head movement; whereas our findings for head movement were mixed. In same-sex dyads or groups of previously unacquainted people, women are more expressive and do more back channeling than do men in same-sex groups. We found similar sex effects for t-tests but not for more rigorous actor-partner analysis. A factor may be that people respond differently with intimate partners than with a stranger. Most studies of sex differences in social behavior have been conducted on unfamiliar persons. Both relationship history and data analytic strategy are important to consider.

We found an increase in Partners' RMS angular displacement of head pitch during conflict relative to non-conflict. This increase was not different in inpatient participants and their partners. This result suggests that it is not just the inpatient participant who is differing between conflict and non-conflict. Both partners are working together to behave differently. This is consistent with Caughlin's work on the demand-withdraw (DW) interaction pattern which suggests that demand and withdrawal are not pre-established roles between partners. From Caughlin's perspective, DW is a dynamic pattern in which roles are exchanged over the course of interaction [10], [11]. The results of windowed cross-correlation analyses further support this inference. The direction of influence between partners dynamically varied over time. Had the inpatient participant or spouse assumed only one or the other role (i.e., one demanding and the other withdrawing), then a unidirectional pattern would have emerged. These results are in accordance with a dyadic perspective.

Previous work in emotion communication and social interaction emphasized facial expression. Head movement may not rival facial expression in subtlety of expression, but as described above it conveys much information about the experience of intimate partner. When couples were in conflict their head nods increased compared with non-conflict. The temporal coordination of head movements between partners and reciprocal influence strongly suggests a deep inter-subjectivity between them. The obtained results suggest that head motion is not only a source of noise for automatic facial expression analysis, but a potentially rich feature set to include in emotion and relationship detection. Body language, such as head movement, warrants our attention for automatic affect recognition.

The current work opens new directions for future investigations of face-to-face interaction and emotion communication. One direction is the study of the coordination between head movement and emotion exchanges in different contexts of social interaction (e.g., intimate partners, parent-infants). Another direction is to explicitly include head motion as a feature to inform action

unit and emotional expression detection. Finally, further work is needed to integrate temporal coordination of facial and vocal expression with head motion for emotional expression recognition. We aim at testing these hypotheses in other dyadic social interactions such as parents-infants interaction.

Yet another contribution to the literature on nonverbal behavior is to inform the design of intelligent virtual agents and social robots in the context of human-robot and human-agent interactions. Head movement is involved in behavioral mirroring and mimicry among people and qualifies the communicative function of facial actions (e.g., smile is correlated with head pitch in embarrassment [13] and anger with head pitch and yaw in conflict between intimate partners). Social robots that can interact naturally with humans in different contexts will need the ability to use appropriately this important channel of communication. Social robots as well as virtual human agents would benefit from further research into human-human interaction with respect to head motion and other body language. Work in this area is just beginning. Recent work with intelligent virtual agents and social robots has explored individual differences in personality and expressiveness [46], [57]. To achieve even more human-like capabilities, designers might broaden their attention by considering the dynamics of head movements between humans, such as intimate partners or strangers, and by the understanding of the causes of differences, instability, and changes in the direction of lead-lag relationships according to the context. This would make possible greater synchrony, shared intention, and goal-corrected outcomes.

7 CONCLUSION

We proposed an analysis of the interpersonal timing of head movements in dyadic interaction. Videos of intimate couples with a history of violence toward each other were analyzed in two conditions: conflict and non-conflict. The root mean square of horizontal and vertical angular displacement and angular velocity was used to evaluate how animated each participant was during each condition. A variant of actor-partner analysis was used to detect differences between men and women and inpatient participants in angular displacement and velocity of head pitch and yaw. Windowed cross-correlation analysis was used to examine moment-to-moment dynamics. Consistent with previous findings on reciprocal interaction and emotion exchange in dyadic interaction between intimate partners, our results extend these findings to head movements. Partners' head movements highlighted strong mutual influence between partners. Moreover, in both conflict and non-conflict, head angular displacements and velocities were strongly correlated between partners, with frequent shifts in lead-lag relationships between them. The overall amount of coordination between partners' head movement was more highly correlated during non-conflict compared with conflict interaction. While high conflict increased head motion, it served to attenuate interpersonal coordination. In summary, we found that head movement is a powerful signal in dyadic social interaction between intimate partners that varies with conflict.

ACKNOWLEDGMENTS

Research reported in this publication was supported in part by the National Institute of Mental Health of the National Institutes of Health under Award Number MH096951. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- [1] N. Ambady and R. Rosenthal, "Half a minute: Predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness," *J. Personality Social Psychol.*, vol. 64, no. 3, pp. 431–441, 1993.
- [2] K. T. Ashenfelder, S. M. Boker, J. R. Waddell, and N. Vitanov, "Spatiotemporal symmetry and multifractal structure of head movements during dyadic conversation," *J. Exp. Psychol.: Human Percept. Perform.*, vol. 35, no. 4, pp. 1072–1091, 2009.
- [3] J. N. Bailenson and N. Yee, "A longitudinal study of task performance, head movements, subjective report, simulator sickness, and transformed social interaction in collaborative virtual environments," *Presence*, vol. 15, no. 6, pp. 699–716, 2006.
- [4] L. Brent and S. C. Marsella, "Emotionally expressive head and body movement during gaze shifts," in *Proc. 7th Int. Conf. Intell. Virtual Agents*, 2007, pp. 72–85.
- [5] S. M. Boker, J. L. Rotondo, M. Xu, and K. King, "Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series," *Psychol. Methods*, vol. 7, no. 1, pp. 338–355, 2002.
- [6] S. M. Boker, J. F. Cohn, B. J. Theobald, I. Matthews, T. Brick, and J. Spies, "Effects of damping head movement and facial expression in dyadic conversation using real-time facial expression tracking and synthesized avatars," *Philosoph. Trans. Roy. Soc. London B: Biol. Sci.*, vol. 364, no. 1535, pp. 3485–3495, 2009.
- [7] S. M. Boker, J. F. Cohn, B. J. Theobald, I. Matthews, M. Mangini, J. R. Spies, and T. R. Brick, "Something in the way we move: Motion, not perceived sex, influences nods in conversation," *J. Exp. Psychol.: Human Percept. Perform.*, vol. 37, no. 3, pp. 874–891, 2011.
- [8] C. Busso, Z. Deng, M. Grimm, U. Neumann, and S. Narayanan, "Rigid head motion in expressive speech animation: Analysis and synthesis," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 15, no. 3, pp. 1075–1086, Mar. 2007.
- [9] M. O. B. Caughy, K.-Y. Huang, and J. Lima, "Patterns of conflict interaction in mothertoddler dyads: Differences between depressed and non-depressed mothers," *J. Child Family Stud.*, vol. 18, pp. 10–20, 2009.
- [10] J. P. Caughlin and T. L. Huston, "A contextual analysis of the association between demand/withdraw and marital satisfaction," *Pers. Relationships*, vol. 9, no. 1, pp. 95–119, 2002.
- [11] J. P. Caughlin and A. M. Scott, "Toward a communication theory of the demand/withdraw pattern of interaction in interpersonal relationships," in *New Directions in Interpersonal Communication Research*, S. W. Smith and S. R. Wilson, Eds. Thousand Oaks, CA, USA: Sage, 2010, pp. 180–200.
- [12] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ, USA: Lawrence Erlbaum, 1988.
- [13] J. F. Cohn, L. I. Reed, T. Moriyama, J. Xiao, K. L. Schmidt, and Z. Ambadar, "Multimodal coordination of facial action, head rotation, and eye motion," in *Proc. 6th IEEE Int. Conf. Autom. Face Gesture Recog.*, Seoul, Korea, 2004, pp. 129–138.
- [14] J. F. Cohn and M. A. Sayette, "Spontaneous facial expression in a small group can be automatically measured: An initial demonstration," *Behav. Res. Methods*, vol. 42, no. 4, pp. 1079–1086, 2010.
- [15] M. Cox, J. Nuevo-Chiquero, J. M. Saragih, and S. Lucey, "CSIRO Face Analysis SDK," Brisbane, Qld., Australia, 2013.
- [16] J. W. Davidson, "What type of information is conveyed in the body movements of solo musician performers?" *J. Human Movement Stud.*, vol. 6, pp. 279–301, 2004.
- [17] R. De Silva and N. Bianchi-Berthouze, "Modeling human affective postures: An information theoretic characterization of posture features," *J. Comput. Anim. Virtual World*, vol. 15, nos. 3/4, pp. 269–276, 2004.
- [18] J. Dunn and P. Munn, "Becoming a family member: Family conflict and the development of social understanding in the second year," *Child Develop.*, vol. 56, pp. 480–492, 1985.
- [19] S. Duncan, "Some signals and rules for taking speaking turns in conversations," *J. Personality Social Psychol.*, vol. 23, no. 2, pp. 283–292, 1972.
- [20] R. el Kaliouby, and P. Robinson, "Real-time inference of complex mental states from facial expressions and head gestures," in *Proc. Real-Time Vis. HCI*, 2005, pp. 181–200.
- [21] B. Xiao, P. G. Georgiou, B. Baucom, and S. S. Narayanan, "Data driven modeling of head motion toward analysis of behaviors in couple interactions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Vancouver, BC, Canada, 2013, pp. 3766–3770.
- [22] C. G. Fisher, "The visibility of terminal pitch contour," *J. Speech Hearing Res.*, vol. 12, pp. 379–382, 1969.
- [23] T. D. George, M. J. Phillips, M. Lifshitz, T. A. Lionetti, D. E. Spero, and N. Ghassemzadeh, "Fluoxetine treatment of alcoholic perpetrators of domestic violence: A 12-week, double-blind, randomized placebo-controlled intervention study," *J. Clin. Psychiatry*, vol. 72, no. 1, pp. 60–65, 2011.
- [24] H. Gunes and M. Pantic, "Dimensional emotion prediction from spontaneous head gestures for interaction with sensitive artificial listeners," in *Proc. 10th Int. Conf. Intell. Virtual Agents*, 2010, pp. 371–377.
- [25] Z. Hammal, J. F. Cohn, D. S. Messinger, W. Masson, and M. Mahoor, "Head movement dynamics during normal and perturbed parent-infant interaction," in *Proc. 5th Humaine Assoc. Conf. Affective Comput. Intell. Interact.*, Geneva, Sep. 2013, pp. 276–282.
- [26] D. F. Hay, "Social conflict in early childhood," in *Annals of Child Development*, G. J. Whitehurst, Ed. Greenwich, CT, USA: JAI Press, 1984, pp. 1–44.
- [27] J. S. Jang and T. Kanade, "Robust 3D head tracking by online feature registration," in *Proc. IEEE 8th Int. Conf. Autom. Face Gesture Recog.*, Amsterdam, The Netherlands, Sep. 2008, pp. 1–6.
- [28] D. A. Kashy and D. A. Kenny, "The analysis of data from dyads and groups," in *Handbook of Research Methods in Social Psychology*, I. H. Reis and C. M. Judd, Eds. New York, NY, USA: Cambridge Univ. Press, 2000, pp. 451–477.
- [29] D. Keltner, "Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement and shame," *J. Personality Social Psychol.*, vol. 68, no. 3, pp. 441–454, 1995.
- [30] D. A. Kenny, L. Mannetti, A. Pierro, S. Livi, and D. A. Kashy, "The statistical analysis of data from small groups," *J. Personality Social Psychol.*, vol. 83, pp. 126–137, 2002.
- [31] M. L. Knapp and J. A. Hall, *Nonverbal behavior in human communication*, 7th ed. Boston, MA, USA: Cengage, 2010.
- [32] N. M. Laird and J. H. Ware, "Random-effects models for longitudinal data," *Biometrics*, vol. 38, pp. 963–974, 1982.
- [33] G. Laurent and H. Davidowitz, "Encoding of olfactory information with oscillating neural assemblies," *Science*, vol. 265, pp. 1872–1875, 1994.
- [34] G. Laurent, M. Wehr, and H. Davidowitz, "Temporal representations of odors in an olfactory network," *J. Neurosci.*, vol. 16, no. 12, pp. 3837–3847, 1996.
- [35] J. R. Landis and G. G. Koch, "The measurement of observer agreement for categorical data," *Biometrics*, vol. 33, pp. 159–174, 1977.
- [36] R. W. Levenson and J. M. Gottman, "Marital interaction: Physiological linkage and affective exchange," *J. Personality Social Psychol.*, vol. 45, no. 3, pp. 587–597, 1983.
- [37] T. M. Madhyastha, E. L. Hamaker, and J. M. Gottman, "Investigating spousal influence using moment-to-moment affect data from marital conflict," *J. Family Psychol.*, vol. 25, no. 2, pp. 292–300, 2011.
- [38] I. Matthews, J. Xiao, and S. Baker, "2D vs. 3D deformable face models: Representational power, construction, and real-time fitting," *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 93–113, 2007.
- [39] D. S. Messinger, M. H. Mahoor, S. M. Chow, and J. F. Cohn, "Automated measurement of facial expression in infant-mother interaction: A pilot study," *Infancy*, vol. 14, no. 3, pp. 285–305, 2009.
- [40] D. S. Messinger, W. I. Mattson, M. H. Mahoor, and J. F. Cohn, "The eyes have it: Making positive expressions more positive and negative expressions more negative," *Emotion*, vol. 12, no. 3, pp. 430–436, 2012.
- [41] A. Mignault and A. Chaudhuri, "The many faces of a neutral face: Head tilt and perception of dominance and emotion," *J. Nonverbal Behav.*, vol. 27, no. 2, pp. 111–132, 2003.

- [42] L. P. Morency, C. Sidner, C. Lee, and T. Darrell, "Head gestures for perceptual interfaces: The role of context in improving recognition," *Artif. Intell.*, vol. 171, nos. 8/9, pp. 568–585, Jun. 2007.
- [43] K. G. Munhall, J. A. Jones, D. E. Callan, T. Kuratate, and E. Vatikiotis-Bateson, "Visual prosody and speech intelligibility: Head movement improves auditory speech perception," *Psychol. Sci.*, vol. 30, no. 2, pp. 133–137, 2004.
- [44] K. G. Nicholson, S. Baum, L. Cuddy, and K. G. Munhall, "A case of impaired auditory and visual speech prosody perception after right hemisphere damage," *Neurocase*, vol. 8, pp. 314–322, 2002.
- [45] H. Paterson, F. Pollick, and A. G. Sanford, "The role of velocity in affect discrimination," in *Proc. 23rd Annu. Conf. Cognition Sci. Soc.*, 2001, pp. 1–6.
- [46] C. Pelachaud, "Modelling multimodal expression of emotion in a virtual agent," *Philos. Trans. Roy. Soc. B*, vol. 364, pp. 3539–3548, 2009.
- [47] J. Pinheiro, D. Bates, S. DebRoy, D. Sarkar, and R Core Team, "nlme: Linear and Nonlinear Mixed Effects Models," R Package Version 3.1-109, 2013.
- [48] U. J. Pfeiffer, B. Timmermans, G. Bente, K. Vogeley, and L. Schilbach, "A non-verbal turing test: Differentiating mind from machine in gaze-based social interaction," *PLoS ONE*, vol. 6, no. 11, p. e27591, 2011.
- [49] R Core Team. (2013). R: A language and environment for statistical computing [Online]. Available: <http://www.R-project.org/>
- [50] J. L. Rotondo and S. M. Boker, "Behavioral synchronization in human conversational interaction," in *Mirror Neurons and the Evolution of Brain and Language*, M. Stamenov, and V. Gallese, Eds. Amsterdam, The Netherlands: John Benjamins, 2002.
- [51] N. A. Roberts, J. L. Tsai, and J. A. Coan, "Emotion elicitation using dyadic interaction tasks," in *Handbook of Emotion Elicitation and Assessment*, J. A. Coan, and J. J. B. Allen, Eds. New York, NY, USA: Oxford Univ. Press, 2007, pp.106–123.
- [52] A. Risberg and J. Lubker, "Prosody and speechreading," *Speech Transm. Lab. Q. Progress Report and Status Report*, vol. 4, pp. 1–16, 1978.
- [53] M. A. Sayette, K. G. Creswell, J. D. Dimoff, C. E. Fairbairn, J. F. Cohn, B. W. Heckman, and R. L. Moreland, "Alcohol and group formation: A multimodal investigation of the effects of alcohol on emotion and social bonding," *Psychol. Sci.*, vol. 23, no. 8, pp. 869–878, 2012.
- [54] M. Cox, J. Nuevo-Chiquero, J. M. Saragih, and S. Lucey, "CSIRO Face Analysis SDK," ed. Brisbane, Australia, *10th IEEE Int. Conf. Autom. Face Gesture Recog.* Shanghai, China, Apr. 2013.
- [55] The SEMAINE project. (2008–2010) [Online]. Available: <http://www.semaine-project.eu/>
- [56] M. Stel, R. B. van Baaren, J. Blascovich, E. Van Dijk, C. McCall, and M. M. Pollmann, "Effects of a priori liking on the elicitation of mimicry," *Exp. Psychol.*, vol. 57, no. 6, pp. 412–418, 2010.
- [57] M. Schrder, E. Bevacqua, R. Cowie, F. Eyben, H. Gunes, D. Haylen, M. ter Maat, G. McKeown, S. Pammi, M. Pantic, C. Pelachaud, B. Schuller, E. de Sevin, M. Valstar, and M. Wollmer, "Building autonomous sensitive artificial listeners," *IEEE Trans. Affective Comput.*, vol. 3, no. 2, pp. 165–183, Apr.-Jun. 2012.
- [58] M. A. Straus, "Measuring intrafamily conflict and violence: The conflict tactics (CT) scales," *Phys. Violence Amer. Families*, vol. 41, no. 1, pp. 75–88, 1979.
- [59] C. Sidner, C. Lee, L. p. Morency, and Clifton Forlines, "The effect of head-nod recognition in human-robot conversation," in *Proc. 1st ACM SIGCHI/SIGART Conf. Human-Robot Interact.*, Salt Lake City, Utah, USA, Mar. 2006, pp. 290–296.
- [60] B. J. Theobald, I. Matthews, M. Mangini, J. R. Spies, T. Brick, J. F. Cohn, and S. M. Boker, "Mapping and manipulating facial expression," *Lang. Speech*, vol. 52, nos. 2/3, pp. 369–386, 2009.
- [61] F. De la Torre, and J. F. Cohn, "Visual analysis of humans: Facial expression analysis," in *Visual Analysis of Humans: Looking at People*, T. B. Moeslund, A. Hilton, A. U. Volker Krger, and L. Sigal, Eds. Berlin, Germany: Springer, 2011, pp. 377–410.
- [62] A. G. Van Kleef and S. Cote, "Expressing anger in conflict: When it helps and when it hurts," *J. Appl. Psychol.*, vol. 92, no. 6, pp. 1557–1569, 2007.
- [63] A. J. Viera and J. M. Garrett, "Understanding interobserver agreement: The kappa statistic," *Family Med.*, vol. 37, no. 5, pp. 360–363, 2005.
- [64] A. M. Von der Pten, N. C. Krmer, J. Gratch, and S. Kang, "It doesn't matter what you are: Explaining social effects of agents and avatars," *Comput. Human Behav.*, vol. 26, p. 10, 2010.
- [65] J. Xiao, S. Baker, I. Matthews, and T. Kanade, "Real-time combined 2D+3D active appearance models," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2004, pp. 535–542.
- [66] B. Xiao, P. G. Georgiou, B. Baucom, and S. S. Narayanan, "Data driven modeling of head motion toward analysis of behaviors in couple interactions," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Vancouver, Canada, 2013, pp. 3766–3770.
- [67] H. Yehia, T. Kuratate, and E. Vatikiotis-Bateson, "Linking facial animation, head motion and speech acoustics," *J. Phonetics*, vol. 30, pp. 555–568, 1994.
- [68] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol. 31, no. 1, pp. 39–58, Jan. 2009.



Zakia Hammal received the PhD degree in image and signal processing from the University of Joseph Fourier at Grenoble, France. She is currently a research associate at Robotics Institute at Carnegie Mellon University. Her research interests include computer vision and signal and image processing applied to human-machine interaction, affective computing, and social interaction. She has authored numerous papers in these areas in prestigious journals and conferences. She has served as the organizer of several workshops including CBAR 2012 and CBAR 2013. She is the publication Chair of the ACM ICMI 2014 and served as a reviewer for numerous journals and conferences in her field. Her honors include an outstanding paper award at ACM ICMI 2012.



Jeffrey F. Cohn received the PhD degree in psychology from the University of Massachusetts at Amherst. He is currently a professor of Psychology and Psychiatry at the University of Pittsburgh and Adjunct Professor at the Robotics Institute, Carnegie Mellon University. He has led interdisciplinary and interinstitutional efforts to develop advanced methods of automatic analysis of facial expression and prosody and applied those tools to research in human emotion, interpersonal processes, social development, and psychopathology. He codeveloped influential databases, including Cohn-Kanade, MultiPIE, and Pain Archive, and has cochaired the IEEE International Conference on Automatic Face and Gesture Recognition and ACM International Conference on Multimodal Interaction.



David T. George received the medical degree in 1977 from the Bowman Gray School of Medicine, completed a residency in internal medicine at Henry Ford Hospital, Detroit, Michigan, in 1980, and subsequently completed a psychiatric residency at Duke University, Durham, North Carolina, in 1983. He is currently a chief of the Section of Clinical Assessment and Treatment Evaluation at the National Institute on Alcohol Abuse and Alcoholism of the US National Institutes of Health. His current research interests include alcohol in the etiology of domestic violence and their relation to serotonergic function. He has conducted treatment trials to examine the efficacy of serotonergic, dopaminergic, cannabinoid, and NK-1 agents to decrease alcohol consumption in patients with alcoholism.

▷ For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.