Individual and joint body movement assessed by wearable sensing as a predictor of attraction in speed dates

Vargas Quiros, Jose David; Kapcak, Oyku; Hung, Hayley; Cabrera-Quiros, Laura

DOI
10.1109/TAFFC.2021.3138349

Publication date
2023

Document Version
Final published version

Published in
IEEE Transactions on Affective Computing

Individual and Joint Body Movement Assessed by Wearable Sensing as a Predictor of Attraction in Speed Dates

Jose Vargas-Quiros, Öykü Kapcak, Hayley Hung, *Member, IEEE*, and Laura Cabrera-Quiros

Abstract—Interpersonal attraction is known to motivate behavioral responses in the person experiencing this subjective phenomenon. Such responses may involve the imitation of behavior, as in mirroring or mimicry of postures or gestures, which have been found to be associated with the desire to be liked by an interlocutor. Speed dating provides a unique opportunity for the study of such behavioral manifestations of interpersonal attraction through the elimination of barriers to initiating communication, while maintaining significant ecological validity. In this paper we investigate the relationship between body movement, measured via accelerometer sensors, and self-reports or ratings of attraction and affiliation in a dataset of 399 speed dates between 72 subjects. Through machine learning experiments, we found that both features derived from a single individual's body movement and features designed to measure aspects of synchrony and convergence of the couple's body movement signals were predictive of different attraction ratings. Our statistical analysis revealed that the overall increase or decrease in an individual's body movement throughout an interaction is a potential indicator of friendly intentions, possibly related to the desire to affiliate.

Index Terms—Attraction, body movement, speed dates, synchrony, convergence, non-verbal behavior

1 Introduction

Increased eye contact, smiling, laughter. It's not hard to find these behaviors portrayed as manifestations of attraction in popular culture. Research has shown that it is with good reason, as many of these behaviors, associated also with communicating trust, have been related by meta-analyses to self-reported attraction [1]. Less prevalent in popular culture but similarly researched throughout decades in social psychology are the phenomena of synchrony and mimicry as manifestations of attraction. Recently, computational social science has contributed it's share of research in these areas [2].

A complete computational study of the manifestations of attraction in human behavior must necessarily encompass multiple layers, starting with the definition of the phenomenon, including the collection or procurement of suitable measurements, and the selection and interpretation of a computational model. As with many studies interested in such hypothetical constructs, subjectivity and interpersonal differences in the understanding of a phenomenon necessarily play a role in the analysis and interpretation of results. The use of machine learning models adds statistical power, normally at the expense of interpretability, and specially so for very high-dimensional data.

The advances in sensing technologies and the possibilities of sensing human behavior have brought interest in the automatic assessment of human behavior in the social signal processing community [3] originated in computer science. Many of the computational studies of attraction have been motivated by this goal. One reason is the possibility of building tools that can help people modify their behavior in their relationships via automatic feedback. Modern wearable devices make possible the measurement and provision of real-time feedback during interactions. Behavioral insights are also applicable in the development of more human-like virtual agents or robots and in science, in the development of tools that improve the time and possibly quality of psychological and sociological research.

Our line of work aims to investigate how we can automatically estimate interpersonal attraction by quantifying the body movement of the subjects involved, using wearable sensors. In a previous paper predicting the outcome of speed dates using joint body movement features [4], we have shown that it is possible to do so above chance level using features calculated using both participants' body movement. We proposed interpretable movement and coordination features inspired in previous literature that can be extracted from a single body-worn accelerometer.

In this paper we take a broader approach by comparing, through statistical tests and machine learning experiments, the predictive power of individual body movement features

• *Jose Vargas-Quiros and Hayley Hung are with the Department of Intelligent Systems at TU Delft, 2628 Delft, South Holland, The Netherlands. E-mail: {j.d.vargasquiros, h.hung}@tudelft.nl.*
• *Öykü Kapcak is with the Intelligent Systems, Delft University of Technology, 2628 Delft, South Holland, The Netherlands. E-mail: kapcakoyku@gmail.com.*
• *Laura Cabrera-Quiros is with the Escuela de Ingenieria Electronica at the Instituto Tecnologico de Costa Rica, Cartago 30101, Costa Rica. E-mail: L.C.CabreraQuiros@tudelft.nl.*

Manuscript received 11 Mar. 2021; revised 30 Nov. 2021; accepted 8 Dec. 2021. Date of publication 24 Dec. 2021; date of current version 13 Sept. 2023.
This work was supported in part by the Netherlands Organization for Scientific Research (NWO) under Grant 639.022.606.
(Corresponding author: Jose Vargas-Quiros.)
Recommended for acceptance by C. C. Lee.
Digital Object Identifier no. 10.1109/TAFFC.2021.3138349

(derived from a single person's movement) with that of joint movement features (derived from both people in the interaction) for the prediction of the attraction self-reports in our dataset. We hypothesize and test whether mean intensity and mean changes in intensity of a person's body movement (increase or decrease throughout the interaction) significantly correlate with our attraction labels. Features obtained from a single individual's movement are compared with the previously proposed joint features, designed to capture different aspects of interpersonal coordination, to assess the predictive power of individual and joint body movement. Furthermore, we significantly expand the background literature that supports our joint body movement features, and test for the occurrence of convergence or divergence of body movement in our short date interactions, with the purpose of determining whether this phenomenon could be observed at all in our subjects. Finally, we performed an ablation study with the purpose of understanding the relative importance of the different types of features when used in isolation.

In Section 2, we start by presenting our review of attraction literature, as well as literature about interpersonal body movement phenomena including synchrony, mimicry and convergence. In Section 3 we present the dataset used to test our hypotheses, as well as the individual and joint body movement features proposed. Finally, we present our results and discussion on the relationship between body movement and individual attraction. We test the hypothesis that an increase or decrease in overall body movement throughout a short interaction can be related to the self-reported attraction scores. In a computational stage, we used our individual movement features to directly predict the ratings of attraction. We also investigated the automatic prediction of joint attraction using *match* labels extracted from the individual ratings. In this case we used joint features obtained from the acceleration signals of both interactants.

2 Attraction and Body Movement

The following sections review works in both psychology and computer science that address attraction and the phenomena of synchrony, mimicry and convergence, with a focus on body movement; and its possible role as manifestations of attraction in face-to-face interactions.

2.1 Interpersonal Attraction

Despite the large body of work in the subject, attraction remains notoriously hard to define. The way attraction is treated in recent research does not deviate greatly from the situation in 1969 [5], where most research considers attraction as an attitude, defined as a "readiness to respond toward a particular object in a favourable or unfavourable manner", or a "tendency or predisposition to evaluate an object in a certain way". Attraction is thus generally conflated with positive attitude, and the most common technique to assess an individual's attitude remains self-report. The lack of consensus is not limited to the question of how to define and measure attraction. Montoya [6] lists several other contentious topics which have resulted in a "fragmented field, one that proceeds without a unifying theoretical model".

Multiple works have explored the possibility of attraction as a multi-dimensional phenomenon [5], [6], [7] that cannot be summarized in a scale from negative to positive attitude. Montoya [6] present a two-dimensional model of attraction, with an affective and a behavioral component that are the consequence of an assessment of a target's willingness and capacity to facilitate the individual's goals and interests. The affective component reflects the "quality of one's emotional response towards another", while the behavioral component "reflects one's tendency to act in a particular way toward another". Although in many cases both components are said to align, there are occasions in which they diverge. Attraction is said to differ from love, friendship, attachment and other related constructs in that it is an "immediate evaluation of a target person", that characterizes interpersonal experiences in general.

Among computational studies attraction has been conflated with interest. Gatica-Perez defines the term interest as "people's internal states related to the degree of engagement displayed, consciously or not, during social interaction" [8]. He also notes that this engagement may arise from different factors such as interest in the topic of a conversation, attraction to the other person or social rapport. In this work we make use of the terms *attraction* and *interest* interchangeably, as expressing a desire to maintain or increase contact with another person, and encompassing friendly, romantic and sexual intentions.

A good portion of the work on attraction has conducted experiments in speed date settings, where self-reported attraction can be obtained from questionnaires filled-in by participants [9]. Previous work investigated romantic, friendly and business interest between partners by extracting four types of social signal measures from audio: activity, engagement, emphasis and mirroring and successfully predicted each type of interest using these features [10]. Prosodic, dialogue, and lexical features extracted from audio recordings have also been used to predict both flirtation intention and perception [11].

Research also has explored the different mechanisms and strategies used when it comes to searching for short-term and long-term partners [12], which unsurprisingly differ between men and women. It has been noted that men tend to relax their standards further than women when seeking short-term mates and tend to have higher preferences for physical attractiveness in short-term than long-term mates [13]. Courtship behavior such as flipping of the hair and moving the shoulders has been observed more particularly in women, while men tended to cross and uncross their legs more often [14].

Previous work [15] found that positional features extracted from video such as position, distance, movement and synchrony are indicators of attraction. Their results also indicated that separating male and female training data increased the task performance. Cabrera-Quiros *et al.* attempted to classify attraction levels between participants using statistical features extracted from accelerometer data [16]. For them, separating male and female data did not improve prediction performance.

2.2 Individual Body Movement and Attraction

Numerous factors determine our body movement during an interaction. While some of them can be related to variables accessible to measurement, like our own speech output [17], [18] or environmental stimuli like music, many are understood to be modulated by our own internal states.

Although to the best of our knowledge the direct relationship between attraction and intensity of body movement has not been studied in a speed dating setting, a link between the two can be made through physiological arousal. Arousal levels have been studied as a correlate of attraction with significant results. Most studies in this area manipulate arousal via physical activity [19], [20] or by startling subjects [21], finding that increased physiological arousal resulted in higher attraction ratings compared with baseline arousal. While these results would suggest that arousal is the cause of increased attraction, and not conversely, the direction of the relationship is not important as it relates to predictive performance.

2.3 Synchrony, Mimicry, Convergence and Their Role in Attraction

The behavior of our interlocutor is another factor that clearly influences our own body movement in an interaction [22]. Numerous terms have been used in literature to refer to the dependence in the behavioral signals of dyadic partners, such as synchrony [2], [23], mimicry [24], coordination [25], [26], [27] and chameleon effect [22].

Delaherche defines synchrony as the ”dynamic and reciprocal adaptation of the temporal structure of behaviors between interactive partners”, where the important element is ”the timing, rather than the nature of the behaviors” [2]. Interactional mimicry, on the other hand, has a slightly more precise definition: ”when a behavior is repeated by an interaction partner within a short window of time” , typically no longer than three to five seconds” [22], [28].

However, there is clearly no consensus for the previous or any definition of synchrony. Bernieri defines coordination as ”the degree to which the behaviors in an interaction are nonrandom, patterned or synchronized in both form and timing” [29], where synchrony describes the ”timing” dimension. Other authors however, have followed even less inclusive definitions. Paxton defines synchrony as a special case of coordination, where the same behavior is performed at the same time, thus conflating it with behavioral mimicry [30].

Although mimicry may be of speech, facial expressions, head movement, laughter, emotional responses and other ”observables” (ie. the behavior we observe in others) [31], [32], [33], [34], [35], [36], some of which cannot be easily delimited in time, we are interested in body movement mimicry, also termed ”behavioral mimicry”, ”behavioral matching” or ”chameleon effect” [37]. This includes the repetition of the same gestures (eg. hair touching), or movement of the trunk (eg. leaning forward), and the use of similar postures.

We abide by the definition by Delaherche [2], and consider mimicry to overlap with synchrony; and coordination to be an umbrella term including both phenomena and referring to all ”nonrandom and patterned behaviors during a social interaction” [25], [26]. Although episodes of body movement mimicry can be considered episodes of synchrony under the definition presented, insofar as repetition of the same action implies some degree of synchrony, this repetition might be performed in a highly uncoordinated manner (eg. waiting too long or too little to reciprocate a handshake may be perceived as awkward). We consider that the measurement of the kind of coordination that facilitates social interaction requires access to contextual variables, and cannot be agnostic to the nature of the actions. Like most empirical studies, we adopt a more functional approach with measures of coordination that include aspects of both synchrony and mimicry, and can be defined for behavioral time series, such as mutual information.

Synchrony has been studied specially in its link to affect, where a positive association has been found [23]. Previous work found that temporal coordination of same-sex dyads changed depending on if they described liking, disliking, or being unacquainted to each other [38]. Synchrony has been found to relate to multiple individual outcomes like reduced anxiety and tendency to self-identify in terms of relationships with others; as well as interpersonal outcomes like increased harmonious feelings and prosocial behavior [26]. Other studies have found that synchrony could relate to communication competence [39]; that synchrony decreases significantly during arguments [40], that more synchronous groups are perceived as more united [41] and that synchrony occurs in the psychotherapy setting [42] and could positively affect ratings of the bond with the therapist [43].

Mimicry, on the other hand, has been linked repeatedly to rapport and linking, increased mimicry leading to more favorable evaluations from an interaction partner [44] and to higher ratings of smoothness of the interaction [45]. Furthermore, having an affiliation goal was found to increase non-conscious mimicry; and people who unsuccessfully affiliate in an interaction were found to mimic more, providing evidence for mimicry being used as a tool to achieve affiliative goals [37], [44]. Computational studies have estimated team cohesion in meeting settings using audio-visual cues and mimicry features [46], [47] with performance significantly better than random.

In the courtship setting, a meta analysis found mimicry of nonverbal behavior to be associated with self-reported attraction [1]. In a similar context, it has been found that nonverbal mimicry is positively associated with romantic interest in an interlocutor [24], that people who are involved in a romantic relationship mimic an attractive opposite-sex other to a lesser extent than people not in a relationship, and that they mimic less the closer they are to their current partner [48]. Beyond mimicry of nonverbals, similar associations have been found for language similarity between partners [49]. A study with speed dates [50] found that men evaluated the interaction more positively when they were mimicked by their female partner, while also increasing their ratings of the sexual attractiveness of the woman. In a study on four-minute speed dates, authors found no evidence that attraction ratings can be predicted by mimicry of certain coded behaviors (smiling, laughing, head shaking, hand gestures, face touching), although it found evidence that synchrony in physiological signals like heart rate and skin conductance does predict attraction [51], and evidence for physiological synchrony has been found in other contexts [52]. A more recent study [53] found that coupling in body swaying during speed dates predicted interest in a long-term relationship.

2.4 Measuring Synchrony, Mimicry and Convergence

When it comes to measuring synchrony and mimicry, it is clear that it is hard to separate these two phenomena from one another. Microanalysis from videos consists in the fine-grained coding of the timing of particular within-action moments, which can be used to measure differences in timing, related to synchrony. However, this technique is expensive in terms of human effort [27]. The coding of actions or behaviors has been prevalent in the literature as a way of quantifying action imitation or mimicry [24], [48], [50], which also enables the analysis of leading and following behaviors and roles [54]. However, behavioral coding is also expensive and cannot be used for the study of synchrony without fine-grained temporal resolution or lower level annotations (ie. microanalysis). Therefore, many studies have resorted to the use of motion energy analysis [23], [55] from videos, wearable accelerometers or motion tracking methods [40], [56], [57]. All of these methods result in time series that act as proxies for the motion of a particular body part, or as an average of body movement energy.

Multiple methods attempt to derive a measure of synchrony from such time series using, for example, windowed correlations between them, possibly with different time lags [57]. It is clear, however, that correlation-based measures capture elements of both synchrony and mimicry, as both the nature and the timing of actions can affect them. The length and delay between windows is critical in this process. Schoenherr [58] compared different such time series analysis methods present in literature, including global (whole time-series) Pearson correlations and windowed correlations. The authors experimented with different ways of summarizing these outputs into scalar synchrony measures, and found that these measures were only partially correlated with each other. Furthermore, they did not find evidence of a common factor, concluding that these measures capture different aspects of synchrony.

Some recent studies using acceleration signals have made use of cross-recurrence quantification analysis (CRQA) [59], [60]. This method allows researchers to measure the extent in which two streams of information exhibit similar patterns in time, while answering questions about the characteristic time-lags in the interaction [30]. Computational methods for the discovery of mimicry episodes have also been presented [61].

Datasets have been created for the study of mimicry, although in very different and specific settings like political discussions, role playing games and negotiations [62], [63].

Somewhat more clear is the definition of the *interpersonal convergence*. We abide by its most common definition as an increase in similarity, according to some measure of similarity between features of interest [64], [65]. A study with conversations lasting between 15 and 20 minutes found evidence for the occurrence of pitch convergence and its relation to perceived attractiveness, likability and conversation quality [66], [67]. Convergence has also been observed in the amount of laughter in a conversation [68] and the use of iconic gestures [69]. Ogata [70] coined the term *coevolution* to refer to joint changes in body movement, and found it to be more prevalent in face-to-face than in non-face-to-face interaction. A similar study used the term ”synchrony” [71].

In the speech community, the related phenomenon of ”entrainment”, which can be understood to include both synchrony and convergence, has been established and studied in different acoustic-prosodic features such as intensity, pitch and jitter [64], [72], as well as turn-taking features [73] and gap lengths [65] while being related to different social outcomes [74].

Synchrony relates to convergence in that it can be the object of convergence [35], that is, individuals may become more coupled in time as an interaction progresses. Convergence is certainly not limited to synchrony, as it can affect the nature of the behaviors as well (i.e., mimicry) or modulate the way they are performed (e.g., their intensity). In some cases such as entrainment to external stimuli [75], synchrony and convergence may be tightly linked.

Moulder [76] wrote about the importance of using surrogate data when establishing the occurrence of synchrony, to avoid observing pseudo-synchrony, the amount of spurious synchrony expected between two individuals who are not interacting. A simple surrogate data generation method may consist in calculating synchrony between non-interacting pairs to serve as a baseline or control. These ideas are necessary in studies of synchrony [23], [55] and further apply to study of convergence.

In conclusion, there is enough evidence in previous literature to support a link between attraction and body movement, possibly mediated by the known link between mimicry and rapport. It is however unclear whether this link is limited to mimicry or if features capturing more general coordination or convergence phenomena may also be informative. The role of individual body movement in isolation as an indication of being attracted to the conversational partner also remains unexplored. Furthermore, previous work does not elucidate what kinds of attraction can be predicted from wearable body movement signals and little is known about gender differences in the link between overall body movement (as measured by wearables) and attraction.

3 Dataset and Methods

In our experiments, we made use of the *MatchNMingle* dataset, a multimodal and multi-sensor dataset recorded to be used in research about automatic analysis of social signals and interactions for both social and data sciences [16]. The data was collected in an indoor in-the-wild setting. It was attempted to keep the social interactions between participants as natural as possible.

3.1 Experiment Context

The *MatchNMingle* dataset was recorded over three days in a local bar. Each day had different participants. The event started with a speed dating round where participants of opposite sex had a three-minute date with each other, followed by a mingling event. In this study, only the data from

Fig. 1. Speed dating participants wearing accelerometer devices sat opposite to each other during speed dates [77].

the speed dating part of the event was used. Fig. 1 shows several pictures of the speed daters.

Participants were recruited from a university, fitting the criteria of being single, heterosexual and between the ages of 18 and 30. In total of 92 participants attended the event, with equal number of men and women. The majority of the participants did not know each other. Before the event, participants were asked to wear sensors around their necks to record tri-axial acceleration and proximity, as a requisite for participation. The accelerometers recorded at a frequency of 20Hz. Participants were also made aware that they were being recorded via cameras installed on a frame above the interaction area. The recorded video data is not used in this study.

After each three-minute date with a participant of the opposite sex, participants were given 1 minute to fill a booklet with a questionnaire about their date partner indicating their interest in each other. Responses for these questionnaires constitute the ground truth for the tasks in this study.

The collection of the MatchNMingle dataset took place over three days. 16 males and 16 female subjects participated in the first day, each involved in 14 speed dates. In the second and third days, 15 males and 15 females took part each day, with each person participating in 15 dates. This resulted in a total of 674 speed dates. However, due to malfunctioning wearable devices, some participants did not have valid acceleration data and the data from their speed date interactions had to be discarded. From the 92 participants in the event 72 had valid data. Furthermore, a smaller number of interactions were removed because booklet responses were unreadable. This reduced the number of speed dates in the dataset from 674 to 399. In the final dataset, each subject is present in 11.1 speed dates on average, with a minimum of 9 and maximum of 14 speed dates for any one subject. Each of these dates became an example in our dataset.

3.2 Defining the Ground Truth

The questionnaire that participants filled after their dates consisted of following questions with responses on a 7-point Likert scale (low = 1, high = 7):

1) How much would you like to see this person again?
2) How would you rate this person as a potential friend?
3) How would you rate this person as a a short term sexual partner?
4) How would you rate this person as a long term romantic partner?

These questions were chosen because, in line with a general notion of attraction as *interest in the interlocutor* in a goal-oriented manner, they cover most common ways in which subjects may be interested in each other in the context of an informal speed date. Concretely, the first question captures a general notion of interest by wanting to see the other person at least one more time. This interest could be towards any of the three goals implicit in the next three questions. Question 2 explicitly asks for interest in a friendship. This type of interest has been linked to rapport, with it incorporating feelings of friendship and caring, and the notion of being in-sync [78]. Romantic and Sexual ratings, on the other hand, are directly related to partner choice, where a range of factors like similarity, reciprocity, physical attractiveness and security offered by the partner are known to play a role in the assessments [79].

In Fig. 2, we show the correlations between the raw Likert-scale ratings of the same interaction, where the goal was to understand overall gender-related differences in the way males and females treated the ratings, given that large gender-based differences in partner choice are reported in literature [79]. The first plot shows correlation between the four different ratings (questions) given by males for the same interaction; the second between ratings given by females, and the third between the ratings of the males and the ratings of the females (ie. a positive value means that men and women tended to agree in their ratings of how much they liked each other; a negative value that ratings were often opposite). Males made a big distinction between the Friendly label and the rest of the labels, but SeeAgain, Romantic and Sexual have similarly higher levels of correlation. Females, on the other hand, tended to form two clusters, with Friendly and SeeAgain ratings being one (labeled similarly) and Romantic and Sexual labels being another.

Correlations between male and female responses are low, highlighting the importance of analyzing attraction first as an individual construct, as there is seldom agreement on attraction. Interestingly, only correlations involving the Sexual rating were significant. Male *Sexual* ratings correlate negatively with all female ratings except for the Friendly intention. For females, their Sexual ratings correlate negatively with male Sexual and Friendly ratings.

Each of these ratings was used to define different tasks for the interest prediction problem as *See Again*, *Friendly*, *Sexual* or *Romantic*, which consist in predicting the corresponding label. For a more straightforward interpretation of the results, we treated the classification problem as a binary one. Responses to one question were binarized by assigning a positive label to the ratings equal or above the median (per gender) of all ratings given for that question, and a negative label otherwise. In other words, the median of the ratings was used as the threshold for binarization. The threshold was different per gender because in the experiments we also predicted separately for males and females and the distributions of scores were very different between them. Fig. 3 shows the distribution of booklet responses and the the corresponding median thresholds used for binarization. Additionally, interactions were labeled as a *match* when both speed daters had a positive label for the interaction.

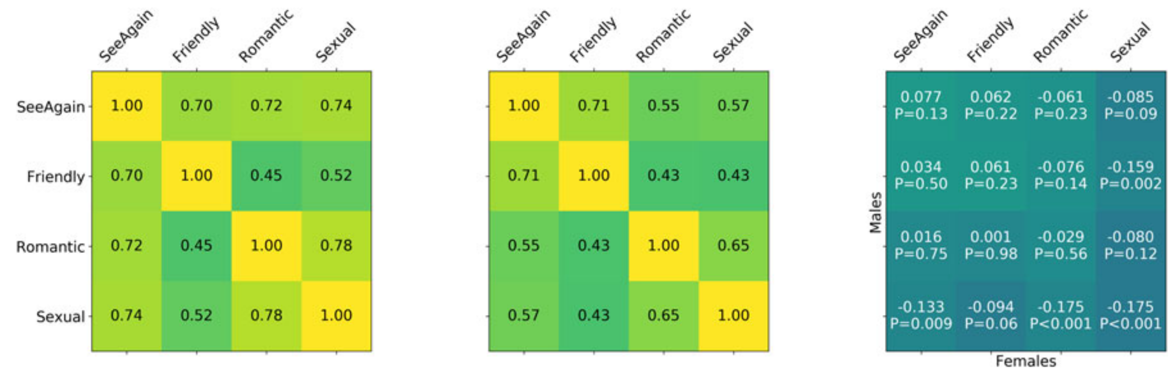

Fig. 2. Spearman correlations of speed date responses (Likert scale from 1 to 7). (Left): Male ratings. (Center): Female ratings. (Right): Male and female ratings.

3.3 Feature Extraction

Our method aims to model the coordination of behavior between two people in an interaction using nonverbal behavioral features extracted from accelerometer readings. We describe the feature extraction process in detail below.

3.3.1 *Preprocessing*

The accelerometer data consists of 3-dimensional readings recorded at 20 Hz with the X axis capturing the left-right movements; the Y axis up-down movements and Z axis forward-backward movements. Initially each axis of each person's recordings is normalized by subtracting its mean and dividing by its standard deviation. This is done to reduce the effect of gravity and interpersonal differences of movement intensity in the sensor readings, and follows previous work [80]. These three normalized signals are augmented with the absolute value signal of each axis, and the magnitude of the acceleration computed as $\sqrt{(x^2 + y^2 + z^2)}$ for a total of 7 signals.

Each of these 7 signals was divided into n-second windows using a sliding-window approach, with $n/2$ second shifts between each window. Since the optimal window size to capture relevant behavior is not known, we chose to extract windows for multiple values of n: 1, 3, 5 and 10 seconds; all of which are included.

Similar to [80], statistical (mean, variance) and spectral (power spectral density) features are extracted from each window. Power spectral density (PSD) per window is computed using 6 logarithmically-spaced bins between 0-10 Hz, to increase the resolution at low frequencies, which contain most of the energy of human movement.

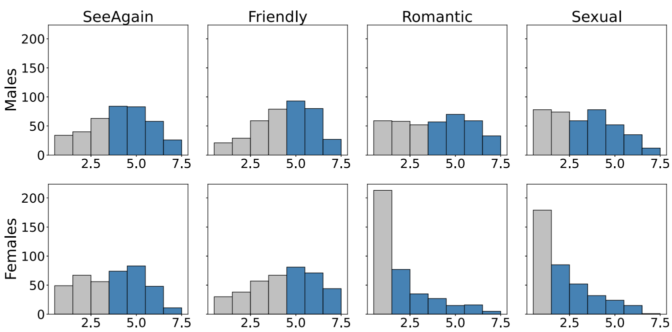

Fig. 3. Distribution of the speed date responses for the four questions asked, for the 399 interactions in the dataset.

Each bin gives information about the characteristic of behavior of the person at that time window, therefore each bin is treated as a single feature. Combining these features results in 8 feature dimensions per window.

Computing these 8 features for each 7 types of signal mentioned earlier and for 4 different window-sizes results in 224 low-level signals that will further be used to extract behavioral coordination features that are explained in the following subsection.

An illustration of the pre-processing steps is shown in Fig. 4.

The aforementioned low-level signals are used to extract more complex body movement features that are grouped into two categories: individual and pairwise features.

3.3.2 *Individual Features*

For experiments using the body movement of a single individual as input, we made use of two simple features that quantify how low-level body movement signals change during the course of the interaction.

Time-correlation. One time-correlation feature was computed as the Pearson correlation coefficient (Pearson's r) computed between one of the low-level signals (eg. PSD bin 3 of the X axis) and time. These capture the general direction of change of the low-level signal throughout the interaction. A positive coefficient for the mean of the magnitude of acceleration, for example, would indicate an increase in body movement intensity throughout the interaction.

Split difference. One split difference feature was computed as the difference between the mean of the low-level signal in the last third of the interaction and the mean in the first third of the interaction. These features similarly capture changes in the underlying low-level signals, by comparing them at the beginning and end of the interaction.

3.3.3 *Pairwise Features*

The following measures aim to quantify body movement behavior between two subjects. The first three measures were created to capture different types of coordination between the movement of the two people in the dyad, especially synchrony and convergence. The next two features were designed to measure convergence (or divergence), the tendency of body movement to become more or less similar during the course of the interaction. Note that, as for the individual features, all of

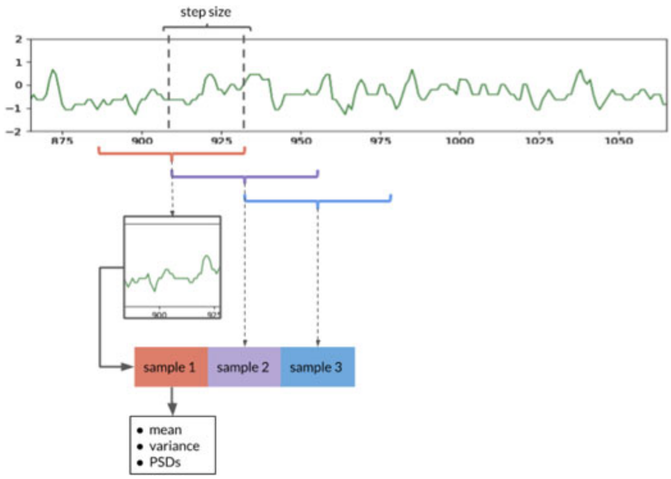

Fig. 4. Pre-processing step: Using a sliding window approach, the signal is divided into samples from which the statistical and spectral features are extracted.

the following joint features are computed for the 224 multi-scale low-level signals (see Section 3.3.1. When present in this section, X and Y refer to a corresponding low level signal (eg. the mean of the X axis of acceleration, calculated using a sliding window of 3 seconds); X for one subject, and Y for the other subject in the interaction.

3.3.3.1 Correlation Linear correlation scores have been used in the literature as a measure of similarity of overall body motion as well as motion of specific body parts such as the hands or head of two people [81], [82], [83], [84], [85].

The linear correlation between two person's body movement signals is expected to result in a score closer to 1.0 the more similar the movement of the two people, hence capturing mimicry in particular but also being affected by the precise timing of the behavior.

Correlation with a time lag has also been used to measure the linear relationship between a follower and a leader's movement [81], [84]. The following computes the correlation between X and Y signals at a positive lag of τ samples:

$$\rho_{xy} = \frac{\sum_{i=1}^{N-\tau}(x_i - \mu_x)(y_{i+\tau} - \mu_y)}{\sigma(X)\sigma(Y)} \tag{1}$$

where x_i and y_i are corresponding samples, μ_x and μ_y the means of the signals and $\sigma(X)$ is the standard deviation of X.

Using time lags enables capturing the leader-follower relationship of two people in a conversation. In an example case of measuring the correlation between persons A and B's movement, if a higher score is obtained when person B's signal is positively lagged, this indicates that person B is leading the interaction.

Following the literature, we use +/- 1 time step lags, and no lag for direct correlations.

3.3.3.2 Distance. This movement similarity measure is inspired by the work of Nanninga [47] and adapted for movement data.

The goal is to capture when one person imitates their partner's behavior. Fig. 5 illustrates how this feature is computed. Each sample window of Person A's signal is compared with the consecutive window of Person B's signal. To compare these windows, the distance between low-level features of these windows are computed, resulting in distance scores $D = [d_0, d_1, \ldots d_n]$ for the entire interaction.

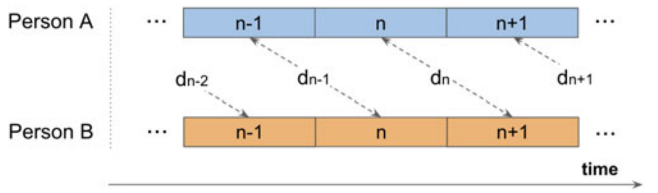

Fig. 5. Distance features. Each time sample is compared with the other signal's preceding time sample.

From these distance scores, minimum ($min(D)$), maximum ($max(D)$), mean ($mean(D)$) and variance ($var(D)$) are computed and used as features. Since this feature is asymmetrical, the reverse is also computed.

3.3.3.3 (Normalized) Mutual Information. Mutual information computed between the random variables corresponding to two movement signals has also been used in the literature to capture the dependence between two people's behavior [80], [86]. In our case it captures the dependence of two people's behavior on each other. It quantifies how much information can be obtained about one variable by observing the other variable. Mutual information is calculated as follows:

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{2}$$

where $H(X)$ and $H(Y)$ represent the entropy of random variables X and Y and $H(X,Y)$ represents their joint entropy. As the calculation of entropy requires knowledge of the underlying probability distributions, we approximated $P(X)$, $P(Y)$ and $P(X,Y)$ using categorical distributions by calculating 10 bin histograms for the marginal distributions, and a 10×10 histogram for the joint distribution.

Additionally, normalized mutual information is computed by dividing by $\sqrt{H(X)H(Y)}$ to obtain a score between 0 and 1. A higher score is expected when two people have an influence on each other's behavior.

While the three previous features attempt to measure elements of coordination, the next two sections describe features that aim to capture the degree of convergence or divergence of body characteristics during the short interaction.

3.3.3.4 Time-correlation. Time-correlation features try to capture if the difference between two people's behavior increases or decreases over time [47], [66]. In order to compute it, corresponding windows of two participants' signals are compared with each other. To measure the similarity at each time step, the distance between these corresponding samples' low-level features are computed as illustrated in Fig. 6, resulting in distance scores $D = [d_1, d_2, \ldots, d_n]$, for each sample. After that, the correlation of these scores with time is computed to understand if they increase or decrease using Pearson correlation formula (Eq. (1)) and a correlation coefficient is obtained. Since the goal is to capture convergence, a decreasing distance indicates converging behavior. Therefore, the correlation coefficient is expected to be more negative for converging interactions where participants show similar behavior over the interaction.

We further incorporated a second type of time-correlation feature inspired in previous work [47], where they

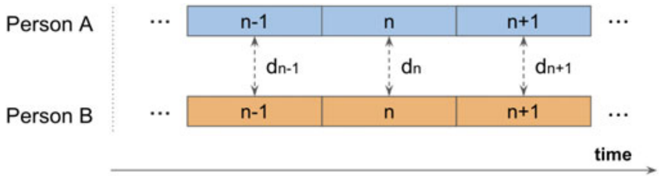

Fig. 6. Time-correlation feature. Each time sample is compared with the other person's corresponding time sample. These distance scores are further correlated with time to extract a convergence score.

were found to be effective at measuring para-linguistic mimicry in meetings. In this case the first two minutes of the date interaction are taken as a "learning period" in which the baseline level of one participant is modeled and the last one minute of the second participant (analysis window) is compared to this learned baseline. To understand if the second person's behavior converges to the behavior exhibited by the first person during the learning period, the N low-level features in the analysis window are compared to the learning period's low-level features. We compared features via subtracting their means, resulting in distance scores $D = [d_1, d_2, \ldots, d_N]$, for each window in the last one minute of interaction as illustrated in Fig. 7. The correlation of scores D with time was then computed using Pearson correlation. A negative correlation coefficient indicates behavior that becomes more similar to that of the other person's baseline. Since this feature is asymmetrical, it was computed for both possible combinations.

The rationale for including these features is the capturing of a baseline level of body movement of one participant for a long period of time (the 2 min "learning period") compared to other features (which compare individual windows) to measure the tendency of the other participant to approach or reject this baseline level.

3.3.3.5 Split-difference. Split-difference features are inspired by the work of [66]. The idea is to measure the similarity of two people's behavior in the beginning and at the end of their date interaction and compare these similarities. It is expected that the behavior will be more similar at the end of the interaction when convergence occurs. To capture this, the first and second half of the signals are taken as illustrated in Fig. 8. The similarity d_0 between the first half's features of the two persons is computed. An equivalent similarity d_0 is calculated for the second half. One feature corresponds difference between these similarities: $c = d_1 - d_0$. This difference is expected to be negative when convergence occurs.

Table 1 summarizes all the features that are used in our experiments, along with their dimensionality. Joint features are separated in those measuring coordination and those measuring convergence of behavior as explained in this section.

3.4 Dimensionality Reduction

After extracting the features, they were processed with the objective of reducing the dimensionality of the feature space. We applied principal component analysis (PCA) and the top principal components preserving 95% of the variance were kept. Features where then normalized to have zero-mean and unit standard deviation, as is standard practice for classification.

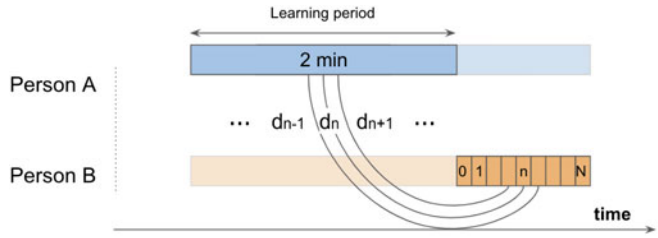

Fig. 7. Convergence features with a learning period. Each window in the last 1 minute period was compared with the other person's first 2 minutes by computing a distance score between mean sample features.

4 Results

Our experiments can be separated in three parts. First we investigate the relationship between body movement intensity and attraction at the individual level, via a correlation. Second, we attempt the automatic prediction of the individual binary attraction levels using a set of convergence features extracted only from individual body movement. Finally, we investigate the automatic prediction of the mutual attraction labels using features designed to capture synchrony and convergence, thus derived from both individuals' time series during these interactions.

4.1 Body Movement and Attraction

We start by investigating the relationship between overall body motion and attraction, starting with a simple hypothesis: the intensity of overall body motion in the interaction is linked to attraction. The magnitude of the accelerometer signal (see 3.3.1) was normalized per participant by dividing by the participant's mean magnitude over all its interactions. This is expected to capture relative changes of individual body movement and remove interpersonal differences in body motion energy.

Table 2 shows the results of correlating the average intensity of the accelerometer readings with the questionnaire responses (7-point scale) for males and females separately. Spearman's r was used to avoid excessive influence from individuals with extreme body movement energies. No significant correlations were found, and in fact all correlation coefficients were negative, suggesting a weak opposite relation.

For the previous calculations, body movement energy was averaged for an interaction, meaning that we did not capture the effect that the interaction had on the body movement intensity of participants, ie. its increasing or decreasing. Our next hypothesis tests whether net increases in body movement indicate heightened interest, possibly

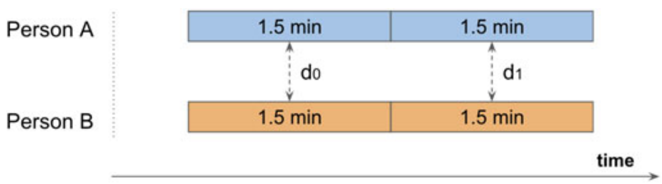

Fig. 8. Split-difference feature. The difference between both persons' features is computed for each half of the interaction.

TABLE 1
Summary of the Individual and Joint Features Used to Predict Attraction Ratings

Type	Category	Feature type	Total
Indiv.	-	Time Correlation [Section 3.3.2.1]	224
		Split-difference [Section 3.3.2.2]	224
Joint	Coordination	Correlation [Section 3.3.3.1]	672
		Distance [Section 3.3.3.2]	1792
		Mutual Information [Section 3.3.3.3]	336
	Convergence	Time Correlation [Section 3.3.3.4]	784
		Split-difference [Section 3.3.3.5]	224

Total indicates the size of each feature vector or number of individual features.

through an increasingly animated conversation. To quantify this we calculated correlations between body movement intensity throughout the interaction, and time. Most correlations were found to be significant ($\alpha = 0.05$), indicating a substantial change in body movement throughout the interaction. For females, from the total of 398 interactions, 32 interactions had a significant increase, and 204 a significant decrease in body movement. For males, 43 coefficients were positive, and 198 negative. The fact that participants were seated and changed seats between interactions is the most likely cause of the very high number of interactions with decreasing movement intensity. Even though the analyzed interactions start a few seconds after participants have seated and greeted each other, it is possible that this moment of higher arousal has an influence on the rest of the interaction, and that participants take more time to reach a state that is closer to their baseline. The same is not true for the end of the interaction, where the recording is ended right before a bell was rang during the event, indicating participants to switch partners.

We used these correlation coefficients as a variable quantifying the effect of the interaction in body movement. Table 3 shows the results of correlations between corresponding r values and speed date responses. In this case three of the correlations were found significant. Interestingly, for all labels correlations are positive for males and negative for females. The strongest significance was found for the *Friendly* and *SeeAgain* labels for both males and females. A possible explanation for this last fact is that high rapport is the driver of these changes in overall body movement. A stronger link of high rapport to the *Friendly* ratings, in comparison with Sexual and Romantic ratings where other aspects like physical attractiveness play a big role, would explain the differences in significance. *SeeAgain* ratings are inherently more ambiguous and the analysis of Section 3.2 indicates that males and females tended towards different interpretations. Note however that all coefficients are below 0.5. The rapport link would imply that high-rapport is associated with increases in male body movement (or less steep decreases given that most of the r values were negative) and with stronger decreases in female body movement throughout the interaction.

4.1.1 Automatic Prediction of Individual Interest

We predicted individual interest based on an individual's accelerometer features (extracted as per Section 3.3.2) and the joint movement features extracted from both speed daters. In these experiments we train a classifier to predict attraction from male to female and from female to male. A logistic regressor (linear model) with L2 regularization was chosen as classifier for the task.

The model was evaluated via 10-fold cross-validation. To avoid having dates from the same subject in train and test sets, the cross-validation split was done via a leave-n-subjects-out approach. When male labels are predicted, the dates from a number of males (three subjects for most folds) are separated as test set in such a way that their dates are not present in the training set. The equivalent happens when female labels are predicted. A nested cross-validation loop within each fold was used to tune the regularization parameter. To obtain a measure that is unaffected by the class imbalance, the Area under the Receiver Operator Characteristic (ROC-AUC) was used as performance measure.

Performances for different attraction type predictions were compared to a random baseline classifier (expected AUC of 0.5), via a statistical test on the 100 classification scores obtained from running 10-fold cross validation 10 times (10x10-fold cross validation). P-values were obtained by using the correction to the paired Student t-test initially proposed by Nadeau and Bengio [87] and recommended [88] for enhancing replicability of the p-values obtained

TABLE 2
Correlations Between Mean Intensity of Body Movement and the Attraction Ratings did not Give any Evidence of Increased or Decreased Body Movement Being a Manifestation of Attraction

	Males		Females	
	Spearman's r	p-value	Spearman's r	p-value
SeeAgain	-0.041	.41	-0.098	.05
Friendly	-0.005	.92	-0.050	.32
Romantic	-0.062	.21	-0.022	.66
Sexual	-0.077	.12	-0.057	.26

TABLE 3
Correlations Between the Individual Time-Correlation Scores and Attraction Labels

	Males		Females	
	Spearman's r	p-value	Spearman's r	p-value
SeeAgain	0.084	.093	-0.107	*.032
Friendly	0.106	*.035	-0.112	*.026
Romantic	0.068	.18	-0.047	.35
Sexual	0.078	.12	-0.034	.49

An asterisk (*) marks significant correlations ($\alpha = 0.05$)

TABLE 4
Mean AUC Scores Obtained in Individual Interest Prediction Tasks via 10x10-Fold Cross-Validation

	Individual Features				Joint Features			
	Males		Females		Males		Females	
Label	AUC	p-value	AUC	p-value	AUC	p-value	AUC	p-value
SeeAgain	0.482	.35	0.588	*.008	0.508	.73	0.584	*.012
Friendly	0.482	.27	0.555	.06	0.510	.76	0.608	*.0002
Romantic	0.493	.71	0.483	.22	0.601	*.005	0.519	.49
Sexual	0.501	.97	0.574	*.011	0.573	.06	0.531	.34

P-values are for the probability of observing more extreme cross-validation scores under a true mean of 0.5 AUC, calculated using the Nadeau and Bengio correction to the paired Student-t test for comparing classifiers [87].

from 10x10-fold cross validation classifier scores. Obtained results are shown in Table 4. Note that AUC scores lower or equal to 0.5 indicate that the classifier was not able to discriminate between the two classes above chance level.

4.2 Joint Body Movement and Attraction

This section focuses on joint movement measures (calculated from both subjects' movement signals) and their relation with mutual ratings of attraction. As before, this is done through both statistical results and classification experiments.

4.2.1 Convergence of Body Movement

Following previous literature which explored the phenomenon of convergence in features of speech in dyadic conversation [66] we investigated whether we can find evidence of convergence of body movement between interacting partners. Previous work found important evidence that several pitch features converge globally over the course of a conversation, independent of the perceived attractiveness or likability of the interlocutor.

We hypothesized that during the 3-minute dates the participants movement characteristics converge or diverge due to the effect of the social interaction. In order to test our hypothesis we compared the convergence scores of interacting and non-interacting pairs. We created non-interacting feature pairs by randomly matching input signals from males to females who were not conversing together. Convergence scores were calculated for real and artificial non-interacting pairs as described in Section 3.3.3.4. However, for these experiments we used only the time-correlation and split-difference convergence features due to their easy interpretation and because they capture the complete temporal extent of the interaction. We used only the convergence features extracted using windows of 3 seconds because, since convergence features are correlations with time, scores for different window sizes are expected to be highly correlated.

It was clear however that there is no significant difference in convergence of body movement magnitude. Not only did we find no significant difference between the means of interacting and non-interacting pairs ($P = 0.97$), but more significantly converging and diverging interactions were found for randomly matched pairs than in the actual interaction. Most of the convergence behavior can thus be attributed to an overall reduction in body movement rather to the effect of the social interaction.

Given these results, we performed a more complete analysis by using similar one-tailed t-tests ($\alpha = 0.05$) for the rest of the time-correlation and split-difference convergence features, this time for all the Power Spectral Density bins, and variance. However, from the total of 112 features only three of these tests were significant, less than expected by chance. We found thus no evidence of difference in the mean of convergence features between interacting and non-interacting pairs.

4.2.2 Automatic Prediction of Mutual Attraction

In these experiments we train a classifier to predict the mutual attraction or *match* binary labels using the joint movement features presented in Section 3.3.3. The goal here is to test the predictive power of body movement in interactions where both participants rated each other above average in a particular item. Note that because *match* labels where obtained as the intersection (logical *and*) of the individual labels (Section 3.2), the dataset is more unbalanced in these tasks. Furthermore, because *match* labels come from both subjects, we did not perform cross-validation splits at the person level, and instead used a traditional split at the example (speed date) level.

As before, we use a logistic regressor with L2 regularization trained and evaluated via 10x10-fold cross-validation using the AUC score as evaluation metric. Obtained classification scores are shown in Table 5. In this case, although three of the mean scores are above 0.55, more than in the individual tasks for males and females, only predictions of the *Friendly* labels reached significance and *Romantic* attraction was the hardest to predict. We found thus no clear evidence that predicting matches in this way can be done with better performance. This could suggest that mutual attraction is less characteristically expressed in body movement. However, part of the reason for the lower performance

TABLE 5
Mean AUC Scores From 10x10-Fold Cross-Validation for Mutual Interest Prediction Tasks

Label	AUC	p-value
See-Again	0.553 (0.011)	.06
Friendly	0.562 (0.011)	*.02
Romantic	0.495 (0.016)	.88
Sexual	0.551 (0.015)	.12

P-values are for the right tail of the t-distribution. A random classifier has an AUC of 0.5.

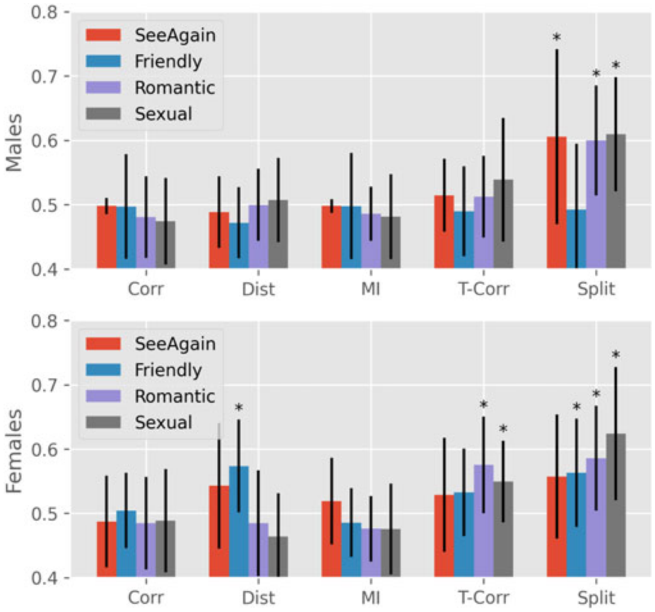

Fig. 9. Results of the ablation study for individual interest prediction tasks with different sets of features. The bars indicate the mean and standard deviation of the AUC scores from 10x10-fold cross-validation. An asterisk indicates performance significantly better than the random baseline classifier.

could come from the lower number of positive labels (25% on average). Imbalance is however hard to avoid as it is a feature of the interactions themselves, where matches are much more rare than one-sided attraction. Experiments with balanced class weights in our Logistic Regressor, a technique which can in some cases offset class imbalance, delivered performance statistically indistinguishable from the results in Table 5 for all four labels.

4.3 Ablation Study: Feature Type Importance

In this section we present the results of an ablation study with the objective of understanding the relative importance of the different types of features (Table 1) in our method. The goal is to understand how different sets of engineered features affected the results in previous sections. We focus on individual interest prediction using joint features, where we had 5 different feature sets designed to capture different aspects of coordination.

The results of the ablation study are shown in Fig. 9. The experimental setup and evaluation was the same as detailed in Section 4.1.1. It stands out from these results that convergence-related features (time-correlation and split difference) were in general the most relevant. These results indicate that features capturing synchrony and mimicry were barely predictive of attraction in isolation, and for males in particular, these coordination features held no discriminative power. Note that it is still possible that interactions between feature are discriminative, but we limited the ablation study to the individual feature sets.

5 Discussion

Our experiments with individual body movement revealed (Table 4) that attraction of a participant can be predicted only by their movement features, with performance significantly better than random guessing. These results suggest that female attraction is more easily revealed by their body movement than male attraction. The statistical analysis (Table 3) suggests a possible explanation: although we found no significant correlation between average acceleration intensity and attraction, women were found to significantly decrease their body movement the more positively they rated their interaction partner in the *SeeAgain* and *Friendly* categories. For men, all correlation coefficients were positive, which suggests that an increase or a less steep decrease in body movement reveals heightened attraction. This relation, opposite to that of females, was however only significant for the *Friendly* rating.

Experiments with joint features designed to capture aspects of synchrony and convergence resulted in better performance in the prediction of individual attraction. Our results indicate that performance in detection of attraction depends not only on the type of attraction but also on the gender of the target subject. In general for females we found stronger evidence that *SeeAgain* and *Friendly* ratings were linked to body movement, and less so for *Sexual* and *Romantic* ratings. For males, the opposite was true in the case of joint body movement (*Romantic* and *Sexual* labels were the better predicted). This separation cuts along the distinction made by participants in their ratings (Fig. 2). Males made a big distinction between the *Friendly* ratings and the rest of the ratings, but *SeeAgain*, *Romantic* and *Sexual* have similarly higher levels of correlation. Females, on the other hand, tended to form two clusters, with *Friendly* and *SeeAgain* ratings being one (labeled similarly) and *Romantic* and *Sexual* ratings being another.

Different interaction dynamics likely play a role in explaining these general trends. Our results suggest that interactions where the female is seeking friendship or the male is seeking romantic or sexual goals have a characteristic signature in body movement. This could be mediated by the interested participant, or possibly both of them, making an effort to affiliate with their partner. Body movement phenomena like mimicry are known to be effective as tools for seeking affiliation and increasing rapport [22], [44].

The better performance of joint features compared to individual ones in predicting individual attraction indicates again that individual experience of attraction has a strong manifestation in the joint interaction, although this general trend could be a result of our particular choice of features.

In attempting to understand the relative importance of the many joint features that we used, the ablation study of Section 4.3 showed convergence features to be the most important, indicating that mimicry and synchrony are less relevant to attraction compared to the less dynamic convergence features. This may appear odd in the light of the results of Section 4.2.1 which established that there was no evidence of convergence taking place above chance levels. However, it is possible that changes in overall body movement levels, or interactions between them captured by the classifier hold the discriminative power. The statistical results of Section 4.2.1 only show that the dyads in our dataset did not converge more often than expected by chance.

Prediction of mutual attraction delivered results significantly better than random for the *Friendly* labels (Table 5). Note that mutual labels have a logical relation to individual labels in that they must both be positive for a positive mutual or *match* label. Therefore, the fact that *Friendly* scores in the joint tasks are between the (low) scores obtained individually for males and the high scores obtained for females (Table 4), would seem to suggest that cases of one-sided female friendliness are easier to detect than when such friendly intentions are mutual. We think however that there is not enough evidence to reach this conclusion, since the greater data imbalance in the mutual tasks could explain having lower results in the mutual tasks.

The fact that no significant difference in convergence could be observed between interacting and non-interacting pairs could be an indication that convergence in overall body movement does not occur over these short timespans, or is much weaker than other factors like the significant average decrease in body movement that we measured during most interactions. However, this evidence is far from conclusive given the simplicity of the sensing modality, that only has access to the acceleration of a single body part (the chest), and is limited to a setting where participants are seated. Another possibility is that convergence manifests itself as increase in the time-synchrony of behavior (ie. is tightly-linked to synchrony), and not in the intensity or style of the movements. This would not be captured by the Time-correlation and Split-difference features, which perform a rough aggregation over the complete interaction.

An analysis directly correlating different joint features with the label of each task revealed that the types of features with the highest correlation coefficients vary with different tasks. Correlation features computed over the *Z*axis were found to be often negatively correlated with *Friendly* attraction as opposed to the expectation of positive correlation that would indicate mimicry. Because the *Z*axis of the accelerometers captured the forward-backward acceleration of the body, low feature values can be produced by a person's backward and partner's forward movement occurring simultaneously. This could indicate that a different kind of synchrony is at play. On the other hand, most of the correlation features extracted from PSD bins had significant positive correlations with the *Friendly* and *Sexual* attraction ratings, indicating that coupling in the frequency of movement could be a correlate of these ratings.

It was also found that *Mutual Information* features tended to have high positive correlation with only the *SeeAgain* and *Friendly* labels whereas the *Mimicry* features correlated more often with the *Romantic* and *Sexual* tasks, offering a possible explanation for the differences in the computational results.

The fact that we found no common features correlating significantly across all of the four ratings tends to indicate that different types of attraction manifest in different behavioral characteristics.

In conclusion, our computational analysis showed that it is possible to predict speed date ratings and the derived matches using individual and joint behavioral coordination features derived from a single body-worn accelerometer. Features engineered to capture synchrony and convergence characteristics, succeeded in predicting three of the mutual attraction levels and distinct individual attraction labels for males and females. Our results indicate that subtle social manifestations of attraction can be captured by wearable devices. This calls for similar studies using more complete body movement sensing. More complex wearable sensors, however, risk interfering with the interactions or limiting body movement. Alternative setups such as video recordings followed by joint position estimation algorithms are worth consideration.

Another limitation of our study is the treatment of the labels, since the combination of the ratings of both partners can have a large effect in the dynamics of the interaction. An interaction where both partners have friendly intentions, for example, can be very different from one where one of them has sexual intentions instead. Not looking at the interaction between labels can therefore be limiting. Classifying label combinations rather than single labels is however impractical with our relatively small dataset.

The development of the computational study of phenomena such as synchrony and convergence via proxies, and their relation with constructs like attraction faces the fundamental problem of lack of suitable, large-scale, ecologically valid datasets. The dataset used in this study is a step in the right direction, but we believe larger wearable sensing or video datasets would allow to more conclusively answer questions related to interpersonal gender, age, and culture-related differences in the manifestation of attraction.

References

[1] R. M. Montoya, C. Kershaw, and J. L. Prosser, "A meta-analytic investigation of the relation between interpersonal attraction and enacted behavior," *Psychol. Bull.*, vol. 144, no. 7, pp. 673–709, 2018.

[2] E. Delaherche, M. Chetouani, A. Mahdhaoui, C. Saint-Georges, S. Viaux, and D. Cohen, "Interpersonal synchrony: A survey of evaluation methods across disciplines," *IEEE Trans. Affective Comput.*, vol. 3, no. 3, pp. 349–365, Jul.–Sep. 2012.

[3] A. Vinciarelli, M. Pantic, and H. Bourlard, "Social signal processing: Survey of an emerging domain," *Image Vis. Comput.*, vol. 27, no. 12, pp. 1743–1759, 2009.

[4] O. Kapcak, J. Vargas-Quiros, and H. Hung, "Estimating romantic, social, and sexual attraction by quantifying bodily coordination using wearable sensors," in *Proc. 8th Int. Conf. Affect. Comput. Intell. Interact. Workshops Demos* 2019, pp. 154–160.

[5] E. Berscheid and E. H. Walster, *Interpersonal Attraction*. Reading, MA, USA: Addison-Wesley, 1969.

[6] R. M. Montoya and R. S. Horton, "A two-dimensional model for the study of interpersonal attraction," *Pers. Soc. Psychol. Rev.*, vol. 18, no. 1, pp. 59–86, 2014.

[7] M. J. Rosenberg and C. I. Hovland, "Cognitive, affective, and behavioral components of attitudes," in *Attitude Organization and Change: An Analysis of Consistency Among Attitude Components*, M. J. Rosenberg Ed., New Haven, CT, USA: Yale Univ. Press, 1960, pp. 1–14.

[8] D. Gatica-Perez, "Automatic nonverbal analysis of social interaction in small groups: A review," *Image Vis. Comput.*, vol. 27, no. 12, pp. 1775–1787, 2009.

[9] E. J. Finkel, P. W. Eastwick, and J. Matthews, "Speed-dating as an invaluable tool for studying romantic attraction: A methodological primer," *Pers. Relationships*, vol. 14, no. 1, pp. 149–166, 2007.

[10] A. Madan, R. Caneel, and A. Pentland, "Voices of attraction," 2004.

[11] R. Ranganath, D. Jurafsky, and D. McFarland, "It's not you, it's me: Detecting flirting and its misperception in speed-dates," in *Proc. Conf. Empir. Methods Natural Lang. Process.*, 2009, vol. 1, pp. 334–342.

[12] R. Fisman, S. S. Iyengar, E. Kamenica, and I. Simonson, "Gender differences in mate selection : Evidence from a speed dating experiment," Oxford Univ. Press, vol. 121, no. 2, pp. 673–697, 2013.

[13] D. Buss and D. Schmitt, "Sexual strategies theory: An evolutionary perspective on human mating," *Psychol. Rev.*, vol. 100, pp. 204–32, 1993.

[14] K. Grammer, M. Honda, A. Juette, and A. Schmitt, "Fuzziness of nonverbal courtship communication: Unblurred by motion energy detection," *J. Pers. Soc. Psychol.*, vol. 77, no. 3, pp. 487–508, 1999.

[15] A. Veenstra and H. Hung, "Do they like me? using video cues to predict desires during speed-dates," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, 2011, pp. 838–845.

[16] L. Cabrera-Quiros, A. Demetriou, E. Gedik, L. van der Meij, and H. Hung, "The matchnmingle dataset: A novel multi-sensor resource for the analysis of social interactions and group dynamics in-the-wild during free-standing conversations and speed dates," *IEEE Trans. Affective Comput.*, vol. 12, no. 1, pp. 113–130, Jan.–Mar. 2018.

[17] W. Pouw, "Quantifying gesture-speech synchrony," in *Proc. 6th Gesture Speech Interact. Conf.*, 2019, pp. 75–80.

[18] P. Wagner, Z. Malisz, and S. Kopp, "Gesture and speech in interaction: An overview," *Speech Commun.*, vol. 57, pp. 209–232, 2014.

[19] C. Foster, B. Witcher, W. K. Campbell, and J. Green, "Arousal and attraction: Evidence for automatic and controlled processes," *J. Pers. Soc. Psychol.*, vol. 74, pp. 86–101, Jan. 1998.

[20] G. Lewandowski Jr and A. Aron, "Distinguishing arousal from novelty and challenge in initial romantic attraction between strangers," *Soc. Behav. Pers. Int. J.*, vol. 32, pp. 361–372, Jan. 2004.

[21] R. A. Dienstbier, "Attraction increases and decreases as a function of emotion-attribution and appropriate social cues," *Motivation Emotion*, vol. 3, no. 2, pp. 201–218, Jun. 1979. [Online]. Available: http://link.springer.com/10.1007/BF01650604

[22] T. L. Chartrand and J. A. Bargh, "The chameleon effect: The perception-behaviour link and social interaction," *J. Pers. Soc. Psychol.*, vol. 76, no. 6, pp. 893–910, 1999.

[23] W. Tschacher, G. M. Rees, and F. Ramseyer, "Nonverbal synchrony and affect in dyadic interactions," *Front. Psychol.*, vol. 5, pp. 1–13, Nov. 2014.

[24] S. D. Farley, "Nonverbal reactions to an attractive stranger: The role of mimicry in communicating preferred social distance," *J. Nonverbal Behav.*, vol. 38, no. 2, pp. 195–208, 2014.

[25] A. Kendon, "Movement coordination in social interaction: Some examples described," *Acta Psychol.*, vol. 32, no. 2, pp. 101–125, 1970.

[26] I. M. Vicaria and L. Dickens, "Meta-analyses of the intra- and interpersonal outcomes of interpersonal coordination," *J. Nonverbal Behav.*, vol. 40, no. 4, pp. 335–361, 2016.

[27] C. Cornejo, Z. Cuadros, R. Morales, and J. Paredes, "Interpersonal coordination: Methods, achievements, and challenges," *Front. Psychol.*, vol. 8, pp. 1–16, 2017.

[28] T. L. Chartrand and J. L. Lakin, "The antecedents and consequences of human behavioral mimicry," *Annu. Rev. Psychol.*, vol. 64, no. 1, pp. 285–308, 2013.

[29] F. J. Bernieri and R. Rosenthal, "Interpersonal coordination: Behavior matching and interactional synchrony," in *Fundamentals of Nonverbal Behavior*. Cambridge, U.K.: Cambridge Univ. Press, 1977, pp. 401–432.

[30] A. Paxton and R. Dale, "Interpersonal movement synchrony responds to high- and low-level conversational constraints," *Front. Psychol.*, vol. 8, pp. 1–16, 2017.

[31] E. G. Krumhuber, K. U. Likowski, and P. Weyers, "Facial mimicry of spontaneous and deliberate duchenne and non-duchenne smiles," *J. Nonverbal Behav.*, vol. 38, no. 1, pp. 1–11, 2014.

[32] H. Griffin, G. Varni, G. Volpe, G. Tome, M. Mancini, and N. Bianchi-Berthouze, "Gesture mimicry in expression of laughter," in *Proc. Int. Conf. Affect. Comput. Intell. Interact.*, 2015, pp. 677–683.

[33] T. L. Chartrand, W. W. Maddux, and J. L. Lakin, "Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry," in *The new unconscious*, R. R. Hassin, J. S. Uleman, and J. A. Bargh, Eds. Oxford, U.K.: Oxford Univ. Press, 2012, pp. 334–361.

[34] A. Fischer and U. Hess, "Mimicking emotions," *Curr. Opinion Psychol.*, vol. 17, pp. 151–155, 2017.

[35] B. Xiao, P. G. Georgiou, C. C. Lee, B. Baucom, and S. S. Narayanan, "Head motion synchrony and its correlation to affectivity in dyadic interactions," in *Proc. IEEE Int. Conf. Multimedia Expo*, 2013, pp. 2–7.

[36] B. Xiao, P. Georgiou, B. Baucom, and S. S. Narayanan, "Head motion modeling for human behavior analysis in dyadic interaction," *IEEE Trans. Multimedia*, vol. 17, no. 7, pp. 1107–1119, Jul. 2015.

[37] T. L. Chartrand and A. N. Dalton, *Mimicry: Its Ubiquity, Importance, and Functionality*. New York, NY, USA: Oxford, 2009.

[38] C. L. Crown, "Coordinated interpersonal timing of vision and voice as a function of interpersonal attraction," *J. Lang. Soc. Psychol.*, vol. 10, no. 1, pp. 29–46, 1991.

[39] Y. Y. Kim, "Achieving synchrony: A foundational dimension of intercultural communication competence," *Int. J. Intercultural Relations*, vol. 48, pp. 27–37, 2015.

[40] A. Paxton and R. Dale, "Argument disrupts interpersonal synchrony," *Quart. J. Exp. Psychol.*, vol. 66, no. 11, pp. 2092–2102, 2013.

[41] D. Lakens, "Movement synchrony and perceived entitativity," *J. Exp. Soc. Psychol.*, vol. 46, no. 5, pp. 701–708, 2010.

[42] S. L. Koole and W. Tschacher, "Synchrony in psychotherapy: A review and an integrative framework for the therapeutic alliance," *Front. Psychol.*, vol. 7, pp. 1–17, 2016.

[43] F. Ramseyer and W. Tschacher, "Synchrony in dyadic psychotherapy sessions," in *Simultaneity: Temporal Structures and Observer perspectives*, 2008, pp. 329–347. [Online]. Available: papers3://publication/uuid/EFF6F812–94D3-4951-8DCD-6CD74E408EA0

[44] J. L. Lakin and T. L. Chartrand, "Using nonconscious behavioral mimicry to create affiliation and rapport," *Psychol. Sci.*, vol. 14, no. 4, pp. 334–339, 2003.

[45] M. Stel and R. Vonk, "Mimicry in social interaction: Benefits for mimickers, mimickees, and their interaction," *Brit. J. Psychol.*, vol. 101, pp. 311–323, 2010.

[46] H. Hung and D. Gatica-Perez, "Estimating cohesion in small groups using audio-visual nonverbal behavior," *IEEE Trans. Multimedia*, vol. 12, no. 6, pp. 563–575, Oct. 2010.

[47] M. C. Nanninga, Y. Zhang, N. Lehmann-Willenbrock, Z. Szlávik, and H. Hung, "Estimating verbal expressions of task and social cohesion in meetings by quantifying paralinguistic mimicry," in *Proc. 19th ACM Int. Conf. Multimodal Interact.*, 2017, pp. 206–215.

[48] J. C. Karremans and T. Verwijmeren, "Mimicking attractive opposite-sex others: The role of romantic relationship status," *Pers. Soc. Psychol. Bull.*, vol. 34, no. 7, pp. 939–950, 2008.

[49] M. E. Ireland, R. B. Slatcher, P. W. Eastwick, L. E. Scissors, E. J. Finkel, and J. W. Pennebaker, "Language style matching predicts relationship initiation and stability," *Psychol. Sci.*, vol. 22, no. 1, pp. 39–44, 2011.

[50] N. Guéguen,, "Mimicry and seduction: An evaluation in a courtship context," *Soc. Influence*, vol. 4, no. 4, pp. 249–255, 2009.

[51] E. Prochazkova, E. E. Sjak-Shie, F. Behrens, D. Lindh, M. E. Kret, and Affiliations, "The choreography of human attraction: Physiological synchrony in a blind date setting," 2019.

[52] W. Tschacher and D. Meier, "Physiological synchrony in psychotherapy sessions," *Psychotherapy Res.*, vol. 30, no. 5, pp. 558–573, 2020.

[53] A. Chang, H. E. Kragness, W. Tsou, D. J. Bosnyak, A. Thiede, and L. J. Trainor, "Body sway predicts romantic interest in speed dating," *Soc. Cogn. Affect. Neurosci.*, vol. 16, no. 1–2, pp. 185–192, Jan. 2021. [Online]. Available: https://academic.oup.com/scan/article/16/1–2/185/5873242

[54] S. Feese, B. Arnrich, G. Troster, B. Meyer, and K. Jonas, "Quantifying behavioral mimicry by automatic detection of nonverbal cues from body motion," in *Proc. ASE/IEEE Int. Conf. Privacy Secur. Risk Trust ASE/IEEE Int. Conf. Soc. Comput.*, 2012, pp. 520–525.

[55] F. Ramseyer, "Synchronized movement in social interaction," in *Proc. ACM Int. Conf. Proc. Ser.*, 2013, pp. 1–5.

[56] N. Yang, Z. Wang, and W. Hu, "Synchrony expression analysis of human-human interaction using wearable sensors," in *Proc. IEEE Int. Conf. Control Autom.*, 2016, pp. 611–615.

[57] J. Paulick *et al.*, "Nonverbal synchrony: A new approach to better understand psychotherapeutic processes and drop-out," *J. Psychotherapy Integration*, vol. 28, no. 3, pp. 367–384, 2018.

[58] D. Schoenherr *et al.*, "Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony," *Behav. Res. Methods*, vol. 51, no. 1, pp. 361–383, 2019.

[59] N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Phys. Rep.*, vol. 438, no. 5–6, pp. 237–329, 2007.

[60] M. I. Coco, R. Dale, and N. Nguyen, "Cross-recurrence quantification analysis of categorical and continuous time series : An R package," *Front. Psychol.*, vol. 5, 2014, Art. no. 510.

[61] W.-S. Chu, J. Zeng, F. De La Torre, J. F. Cohn, and D. S. Messinger, "Unsupervised synchrony discovery in human interaction," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 3146–3154.

[62] S. Bilakhia, S. Petridis, A. Nijholt, and M. Pantic, "The MAHNOB mimicry database: A database of naturalistic human interactions," *Pattern Recognit. Lett.*, vol. 66, pp. 52–61, 2015.

[63] X. Sun, J. Lichtenauer, M. Valstar, A. Nijholt, and M. Pantic, "A multimodal database for mimicry analysis," *Lecture Notes Comput. Sci. Including Subseries Lecture Notes Artif. Intell. Lecture Notes Bioinf.*, vol. 6974, pp. 367–376, 2011.

[64] R. Levitan, "Acoustic-prosodic entrainment in human-human and human-computer dialogue," Ph.D. dissertation, Columbia Univ., New York, NY, USA, 2014.

[65] J. Edlund, M. Heldner, and J. Hirschberg, "Pause and gap length in face-to-face interaction," in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, 2009, pp. 2779–2782.

[66] J. Michalsky and H. Schoormann, "Pitch convergence as an effect of perceived attractiveness and likability," in *Proc. Annu. Conf. Int. Speech Commun. Assoc.*, 2017, pp. 2253–2256.

[67] J. Michalsky, H. Schoormann, and O. Niebuhr, "Conversational quality is affected by and reflected in prosodic entrainment," in *Proc. Int. Conf. Speech Prosody*, pp. 389–392, 2018.

[68] J. Trouvain and K. P. Truong, "Convergence of laughter in conversational speech: Effects of quantity, temporal alignment and imitation," in *Proc. Int. Symp. Imitation Convergence Speech*, 2012, Art. no. 231287.

[69] L. Mol, E. Krahmer, A. Maes, and M. Swerts, "Adaptation in gesture: Converging hands or converging minds?" *J. Memory Lang.*, vol. 66, no. 1, pp. 249–264, 2012.

[70] T. Ogata *et al.*, "Interpersonal coevolution of body movements in daily face-to-face communication," *IEICE Trans. Inf. Syst.*, vol. E100D, no. 10, pp. 2547–2555, 2017.

[71] N. Higo *et al.*, "Interpersonal similarity between body movements in face-to-face communication in daily life," *PLoS One*, vol. 9, no. 7, pp. 1–10, 2014.

[72] S. Weidman, M. Breen, and K. C. Haydon, "Prosodic speech entrainment in romantic relationships," in *Proc. Int. Conf. Speech Prosody*, 2016, pp. 508–512.

[73] R. Levitan, S. Benus, A. Gravano, and J. Hirschberg, "Entrainment and turn-taking in human-human dialogue," in *Proc. Conf. Assoc. Advancement Artif. Intell. Spring Symp. Tech. Rep.*, pp. 44–51, 2015.

[74] R. Levitan, A. Gravano, L. Willson, S. Benus, J. Hirschberg, and A. Nenkova, "Acoustic-prosodic entrainment and social behavior," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics: Hum. Lang. Technol*, 2012, pp. 11–19.

[75] J. Phillips-Silver, A. Ktipis, and G. A. Bryant, "The ecology of entrainment: Foundations of coordinated rhythmic movement," *Music Percept.*, vol. 28, no. 1, pp. 3–15, 2010.

[76] R. G. Moulder, S. M. Boker, F. Ramseyer, and W. Tschacher, "Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses," *Psychol. Methods*, vol. 23, no. 4, pp. 757–773, 2018.

[77] L. Cabrera-Quiros, A. Demetriou, E. Gedik, L. V. d. Meij, and H. Hung, "The MatchNMingle dataset: A novel multi-sensor resource for the analysis of social interactions and group dynamics in-the-wild during free-standing conversations and speed dates," *IEEE Trans. Affective Comput.*, pp. 1–17, Jan.–Mar. 2018.

[78] L. Tickle-Degnen and R. Rosenthal, "The nature of rapport and its nonverbal correlates," *Psychol. Inquiry Psychol. Inq.*, vol. 1, pp. 285–293, Oct. 1990.

[79] S. Luo and G. Zhang, "What leads to romantic attraction: Similarity, reciprocity, security, or beauty? Evidence from a speed-dating study," *J. Pers.*, vol. 77, no. 4, pp. 933–963, 2009.

[80] E. Gedik, "Detecting conversing groups using social dynamics from wearable acceleration : Group size awareness," vol. 0, no. 0, pp. 1–24, 2018.

[81] Y. Hart *et al.*, "Automated video analysis of non-verbal communication in a medical setting," *Front. Psychol.*, vol. 7, pp. 1–8, 2016.

[82] F. Ramseyer and W. Tschacher, "Nonverbal synchrony in psychotherapy: Coordinated body movement reflects relationship quality and outcome," *J. Consulting Clin. Psychol.*, vol. 79, no. 3, 2011, Art. no. 284.

[83] W. Tschacher, G. M. Rees, and F. Ramseyer, "Nonverbal synchrony and affect in dyadic interactions," *Front. Psychol.*, vol. 5, 2014, Art. no. 1323.

[84] A. S. Won, J. N. Bailenson, and J. H. Janssen, "Automatic detection of nonverbal behavior predicts learning in dyadic interactions," *IEEE Trans. Affective Comput.*, vol. 5, no. 2, pp. 112–125, Apr.–Jun. 2014.

[85] A. S. Won, J. N. Bailenson, S. C. Stathatos, and W. Dai, "Automatically detected nonverbal behavior predicts creativity in collaborating dyads," *J. Nonverbal Behav.*, vol. 38, no. 3, pp. 389–408, Sep. 2014.

[86] C. Martella *et al.*, "How was it ? exploiting smartphone sensing to measure implicit audience responses to live performances categories and subject descriptors," in *Proc. 23rd ACM Int. Conf. Multimedia*, 2015, pp. 201–210.

[87] C. Nadeau and Y. Bengio, "Inference for the generalization error," in *Advances in Neural Information Processing Systems*. Cambridge, MA, USA: MIT Press, 2000. [Online]. Available: https://papers.nips.cc/paper/1999/hash/7d12b66d3df6af8d429c1a357d8b9e1a-Abstract.html

[88] R. R. Bouckaert and E. Frank, "Evaluating the replicability of significance tests for comparing learning algorithms," in *Advances in Knowledge Discovery and Data Mining (Lecture Notes in Computer Science Series)*, T. Kanade, J. Kittler, J. M. Kleinberg, F. Mattern, J. C. Mitchell, O. Nierstrasz, C. Pandu Rangan, B. Steffen, M. Sudan, D. Terzopoulos, D. Tygar, M. Y. Vardi, G. Weikum, H. Dai, R. Srikant, and C. Zhang, Eds. Berlin, Heidelberg: Springer, 2004, pp. 3–12. [Online]. Available: http://link.springer.com/10.1007/978–3-540-24775-3_3

Jose Vargas-Quiros is working towards the PhD degree with the Socially Perceptive Computing Lab, TU Delft, Delft, The Netherlands, since 2018. He is interested in multimodal action recognition and conversation quality assessment in-the-wild, the study of interpersonal adaptation and synchrony, and efficient annotation of in-the-wild data.

Öykü Kapcak received the MSc degree from the Delft University of Technology, Delft, The Netherlands, From 2017-2019. She currently works in data analysis with the finance industry. Her masters thesis work revolved around the estimation of attraction using bodily coordinated features extracted from wearable sensors.

Hayley Hung (Member, IEEE) received the PhD degree in computer vision from the Queen Mary University of London, London, U.K, in 2007. She is an associate professor with the Socially Perceptive Computing Lab, TU Delft, The Netherlands, where she works since 2013. Between 2010-2013 she held a Marie Curie fellowship at the Intelligent Systems Lab at the University of Amsterdam. Between 2007-2010 she was a post-doctoral researcher at IDIAP Research Institute in Switzerland. Her research interests include social computing, social signal processing, computer vision, and machine learning.

Laura Cabrera-Quiros is an assistant professor with the Electronics Engineering department, the Costa Rican Institute of Technology. Her research focuses on the use of machine learning and non-invasive technologies (e.g., wearable and embedded devices, cameras, physiological sensors) to understand human behavior, monitor health, and improve people's quality of life.

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