

Using Multimodal Wearable Technology to Detect Conflict among Couples

Adela C. Timmons, Theodora Chaspari, Sohyun C. Han, Laura Perrone, Shrikanth S. Narayanan, and Gayla Margolin, University of Southern California

By monitoring human behavior unobtrusively, mobile sensing technologies have the potential to improve our daily lives. Initial results from a field study demonstrate that such passive technologies can detect a complex psychological state in an uncontrolled, real-life environment.

Mobile, real-time, multimodal monitoring of people in real-life settings represents a new computing trend with important implications for both our physical and mental health. From physical activity to physiological arousal to language use, we can monitor a vast number of variables on an ongoing basis. By collecting data via multiple channels over long periods of time, we can obtain an unprecedented amount of information about ourselves and our lives. For example, we can test data streams to determine how day-to-day events impact our emotions, behaviors, and physical well-being across time as well as how these factors are interconnected.

Wearable technologies can also be interactive, so information obtained from such data (such as increasing levels of vocal pitch indicative of emotional arousal) could trigger messages that prompt interventions aimed

at behavioral change and improved psychological functioning (for example, a text message prompting a guided meditation exercise). Such systems could be used alone or integrated with standard clinical interventions to increase their effectiveness and maximize therapeutic gains.

Although the applications of such mobile sensing systems are exciting, they also present many challenges. Specifically, monitoring human behavior in uncontrolled settings is wrought with difficulty; researchers must synchronize signals across devices, integrate multiple platforms, securely and efficiently store large amounts of data, and process and analyze that data. They must also minimize the burden and intrusiveness of these monitoring systems so that they are not overly disruptive to daily life and people are willing to wear them.



See www.computer.org/computer-multimedia for multimedia content related to this article.

Although devices and procedures for collecting big data have proliferated in recent years, the infrastructure for processing and interpreting such data has lagged behind, creating a bottleneck between the collection of data and its application to real-world uses. For example, developing algorithms that can detect events of interest in real time is a major hurdle because data collected in everyday settings can be noisy, introducing errors to modeling schemes. Beyond these practical concerns, researchers must keep in mind privacy and ethical considerations and work to maximize the data's security and reduce the risk of psychological harm to system users.

The University of Southern California (USC) Couple Mobile Sensing Project (homedata.github.io) is an interdisciplinary collaboration between engineers and psychologists that uses ambulatory computing technologies to study interpersonal relationships, with the eventual goal of developing interventions to improve couple functioning. Here, we report the initial results from a field application involving the use of wearable technology to detect psychological states. In our study, young-adult couples wore biosensors measuring their electrodermal and electrocardiographic activity, physical activity, and body temperature and carried smartphones that collected audio recordings and GPS coordinates for one day. The couples also completed phone surveys each hour to report on their ongoing moods and if they experienced conflict. Classification experiments using binary decision trees resulted in unweighted accuracies as high as 86.8 percent when we combined the features from all sensor modalities. The procedures described in this article could be

extended to develop interactive, real-time interventions to decrease conflict and prompt alternative behavior, improving couples' relationship functioning and quality of life.

USING WEARABLE TECHNOLOGY TO DETECT PSYCHOLOGICAL STATES

Being able to identify, monitor, and alter our emotional states in real time is an important next step toward developing more effective interventions for improving quality of life. For example, engaging in daily positive activities has been linked to improved well-being.¹ In addition to emotional states, the quality of our relationships plays a central role in our mental and physical health. Interpersonal conflicts, such as arguments with coworkers, greatly impact our daily moods.² Romantic relationships in particular play a central role in individuals' quality of life; high levels of relationship conflict, in particular divorce, are linked to increased risk of psychological problems.³

behavioral change outside of the therapy room. However, in their home lives, couples often find themselves pulled into arguments that escalate, become entrenched over time, and are hard to exit.

One method to address this problem is to use wearable technology to monitor problematic relationship dynamics in real life. Mapping these patterns across time could provide data on what variables predict conflict and what factors are associated with conflict resolution. Results from these data could eventually be used to detect conflict and send behavior prompts to alter maladaptive relationship processes as they occur at home.

See the "Previous Research on Detecting Psychological States" sidebar for other work in this area.

Features of conflict

A large body of laboratory research has examined how distressed and nondistressed couples differ in terms of how they interact, respond physiologically, and speak to each other

WEARABLE TECHNOLOGY CAN BE USED TO MONITOR PROBLEMATIC RELATIONSHIP DYNAMICS IN REAL LIFE.

Current therapies aiming to improve relationship functioning typically alter couples' interaction patterns and communication processes. These interventions are usually administered during therapy sessions, with the expectation that any gains made in the session will translate into

during conflict. For example, heightened electrocardiographic (ECG) activity and electrodermal activity (EDA) during problem-solving discussions are linked to decreased marital satisfaction.⁴

Relatedly, covariation in physiological responses over time, or *synchrony*,

PREVIOUS RESEARCH ON DETECTING PSYCHOLOGICAL STATES

Numerous studies have attempted to use machine-learning techniques to automatically detect conflict in spoken conversations, but most of these human-activity recognition attempts have been in controlled settings.^{1,2} Applying these methods in real-life, uncontrolled environments is difficult for several reasons. First, many events of interest, such as conflict, have low base rates, meaning there is less information available for building the classification schemes. Second, signals recorded in such environments tend to be noisy, creating additional challenges for recovering and representing the inherent information. Third, fluctuations in variables (such as electrodermal activity) can reflect other processes (including exercise and anxiety) in addition to the event of interest, making it difficult to differentiate across events.

Some research has attempted to detect behaviors in daily life, but such projects usually focus on directly observable behaviors, such as whether people are talking,³ that are easier to identify than more subjective experiences, like whether conflict is occurring. Moreover, different people can exhibit a range of behavioral, emotional, and physiological reactions in response to conflict, making it difficult to create one system that works well for all individuals. Although detecting psychological states in uncontrolled environments is difficult, developing this ability could have many important applications.

References

1. S. Kim, F. Valente, and A. Vinciarelli, "Automatic Detection of Conflict in Spoken Conversations: Ratings and Analysis of Broadcast Political Debates," *Proc. IEEE Int'l Conf. Acoustic, Speech, and Signal Processing (ICASSP 12)*, 2012, pp. 5089–5092.
2. H. Hung and G. Englebienne, "Systematic Evaluation of Social Behavior Modeling with a Single Accelerometer," *Proc. ACM Int'l Joint Conf. Pervasive and Ubiquitous Computing (UbiComp 13)*, 2013, pp. 127–130.
3. Y. Lee et al., "SocioPhone: Everyday Face-To-Face Interaction Monitoring Platform Using Multi-phone Sensor Fusion," *Proc. Int'l Conf. Mobile Systems, Applications, and Services (MobiSys 13)*, 2013, pp. 375–388.

has been associated with both attachment style (that is, how secure versus anxious people feel in their interpersonal relationships) and marital satisfaction.^{5,6} Exactly what partners say to each other when they are angry is also important; studies have shown that even subtle aspects of language, such as the pronouns we use, are related to important relationship processes. That is, second-person singular pronoun use (such as "you" and "you'd") during problem-solving discussions is thought to reflect higher levels of blaming (for example, "You didn't do the dishes"), whereas first-person singular pronoun use (such as "I" and "I'll") is thought to reflect better communication skills (for example, "I felt frustrated when I saw the dishes in the sink").⁷ It is also possible that couples in more distressed relationships use more negative emotion words (such as "sad") and certainty words (for example, "You always criticize me").

Beyond what couples say to each other, the tone of speech might be an important indicator of relationship functioning. Thus, we can use acoustic measures such as vocal intensity (or loudness) and fundamental frequency (or pitch, F0) to obtain additional data on couples' communication patterns.⁸

Prototype model

Using mobile computing technology, our field study collected self-reports of mood and the quality of interactions (MQI) between partners, EDA, ECG activity, synchrony scores, language use, acoustic quality, and other relevant data (such as whether partners were together or communicating remotely) to detect conflict in young-adult dating couples in their

daily lives. We conducted classification experiments with binary decision trees to retroactively detect the number of hours of couple conflict.

To assess our approach's usefulness, our study addressed four inter-related research questions that generated four tasks:

- › **Question 1:** Are theoretically driven features related to conflict episodes in daily life? **Task 1:** We conducted individual experiments for theoretically driven features, including self-reported MQI, EDA, ECG activity, synchrony scores, personal pronoun use, negative emotion words, certainty words, FO, and vocal intensity.
- › **Question 2:** Are unimodal feature groups related to conflict episodes in daily life? **Task 2:** We combined the features into unimodal groups to determine the classification accuracy of different categories of variables.
- › **Question 3:** Are multimodal feature combinations related to conflict episodes in daily life? **Task 3:** We combined the feature groups into multimodal indices to examine the performance of multiple sensor modalities.
- › **Question 4:** How do multimodal feature combinations compare with the couples' self-report data? **Task 4:** We statistically compared the classification accuracy of our multimodal indices to the couples' self-reported MQI to ascertain the potential of these methods to identify naturally occurring conflict episodes beyond what participants themselves reported, hour by hour.

Our objective here is to present preliminary data and demonstrate our classification system's potential utility for detecting complex psychological states in uncontrolled settings. Although this study collected data on dating couples, these methods could be used to study other types of relationships, such as friendships or relationships between parents and children.

RESEARCH METHODOLOGY

The participants in our study consisted of young-adult dating couples from the Couple Mobile Sensing Project, with a median age of 22.45 years and a standard deviation (SD) of 1.60 years. The couples were recruited from the greater Los Angeles area and had been in a relationship for an average of 25.2 months (SD = 20.7). Participants were ethnically and racially diverse, with 28.9 percent identifying as Hispanic, 31.6 percent Caucasian, 13.2 percent African-American, 5.3 percent Asian, and 21.1 percent multiracial.

Out of 34 couples who provided data, 19 reported experiencing at least one conflict episode and thus were included in the classification experiments.

All study procedures were approved by the USC Institutional Review Board.

Measures

All dating partners were outfitted with two ambulatory physiological monitors that collected EDA and ECG data for one day during waking hours. They also received a smartphone that alerted them to complete hourly self-reports on their general mood states and the quality of their interactions. The self-report options, which were designed to assess general emotional states relevant to couple interactions, included feeling stressed, happy,

sad, nervous, angry, and close to one's partner. Responses ranged from 0 (not at all) to 100 (extremely).

Additionally, each phone continuously collected GPS coordinates, as well as 3-minute audio recordings every 12 minutes from 10:00 a.m. until the couples went to bed.

Physiological indices. We collected physiological measures continuously for one day, starting at 10:00 a.m. and ending at bedtime. EDA, activity count, and body temperature were recorded with a Q-sensor, which was attached to the inside of the wrist using a band. ECG signals were collected with an Actiwave, which was worn on the chest under the clothing. ECG measures included the interbeat interval (IBI) and heart rate variability (HRV), and EDA features consisted of the skin conductance level (SCL) and the frequency of skin conductance responses (SCRs). Estimates of synchrony, or covariation in EDA signals between romantic partners, were obtained using joint-sparse representation techniques with appropriately designed EDA-specific dictionaries.⁶

We used computer algorithms to detect artifacts, which were then visually inspected and revised. All scores were averaged across each hour to obtain one estimate of each measure per hour-long period.

Language and acoustic feature extraction. A microphone embedded in each partner's smartphone recorded audio during the study period. The audio clips were 3 minutes long and collected once every 12 minutes, resulting in 6 minutes of audio per 12 minutes per pair (male and female within a couple). This resulted in a reasonable tradeoff between the size of the audio

data available for storage and processing and the amount of acquired information. For privacy considerations, participants were instructed to mute their microphones when in the presence of anyone not in the study.

We transcribed and processed audio recordings using Linguistic Inquiry

acoustic descriptors onto a vector of fixed dimensionality—independent of the audio clip duration—we further computed the mean, SD, maximum value, and first-order coefficient of the linear regression curve over each speech segment, resulting in eight features. All acoustic and language

OUR STUDY RESULTS SHOW THAT DATA COLLECTED VIA MOBILE COMPUTING ARE VALID INDICATORS OF INTERPERSONAL FUNCTIONING IN DAILY LIFE.

and Word Count (LIWC) software.⁹ For our theoretically driven features (task 1), we used preset dictionaries representing personal pronouns (such as “I” and “we”), certainty words (such as “always” and “must”), and negative emotion words (such as “tension” and “mad”). To test unimodal combinations of features, we used four preset LIWC categories, including linguistic factors (25 features including personal pronouns, word count, and verbs), psychological constructs (32 features such as words relating to emotions and thoughts), personal concern categories (seven features such as work, home, and money), and paralinguistic variables (three features such as assents and fillers).

Voice-activity detection (VAD) was used to automatically chunk continuous audio streams into segments of speech or nonspeech. We used speaker clustering and gender identification to automatically assign a gender to each speech segment. We then extracted vocally encoded indices of arousal (F0 and intensity). To map the low-level

features were calculated separately by partner and averaged per hour.

Context and interaction indices. In addition to our vocal, language, self-reported, and physiological variables, we assessed numerous other factors that are potentially relevant for identifying conflict episodes. The contextual variables included whether participants consumed caffeine, alcohol, tobacco, or other drugs; whether they were driving; whether they exercised; body temperature; and physical activity level. The interactional variables involved the GPS-based distance between partners and information related to whether the dating partners were together, interacting face to face, or communicating via phone call or text messaging and if they were with other people.

The data for the contextual and interactional feature groups were collected via various mechanisms, including physiological sensors, smartphones, self-reports based on the hourly surveys, and interview data.

Conflict. We identified the hours in which conflicts occurred using the self-report phone surveys. For each hour, participants reported whether they “expressed annoyance or irritation” toward their dating partner using a dichotomous yes/no response option. Because determining what constitutes a conflict is subjective, we elected to use a discrete behavioral indicator (that is, whether the person said something out of irritation) as our ground-truth criterion for determining if conflict behavior occurred within a given hour. This resulted in 53 hours of conflict behavior and 182 hours of no conflict behavior for females and 39 hours of conflict behavior and 206 hours of no conflict behavior for males.

Conflict classification system

The goal of the classification task was to retroactively distinguish between instances of conflict behavior and no conflict behavior, as reported by the participants. The analysis windows constituted nonoverlapping hourly instances starting at 10:00 a.m. and ending at bedtime.

To classify conflict, we used a binary decision tree because of its efficiency and self-explanatory structure. We employed a leave-one-couple-out cross-validation setup for all classification experiments. For tasks 2 and 3, feature transformation was performed through a deep autoassociative neural network, also called an *autoencoder*, with three layers in a fully unsupervised way. The autoencoder's bottleneck features at the middle layer consisted of the input of a binary tree for the final decision (Y = conflict and N = no conflict). Unimodal classification followed a similar scheme, under which the autoencoder transformed only the within-modality features.

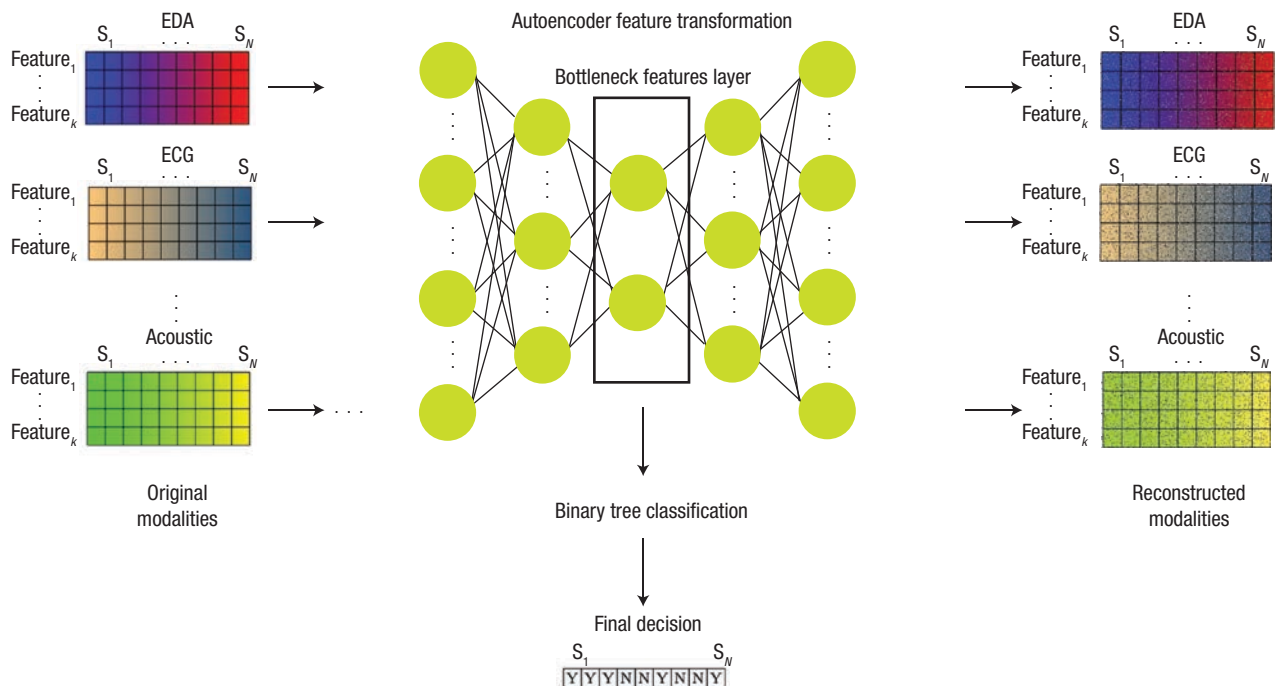


FIGURE 1. Schematic representation of the final classification system. Multimodal classification (task 3) between conflict and non-conflict samples ($S_1 \dots S_N$) used combinations of features from each modality: self-reported mood and quality of interactions (MQI), electrodermal activity (EDA), electrocardiogram (ECG) activity, EDA synchrony measures, language, and acoustic information.

Figure 1 presents a schematic representation of the classification system as it applies to our dataset.

Further details regarding the system, a list of the entire feature set, complete results from all our experiments, and confusion matrices (that is, tables showing the performance of the classification model) are available online at homedata.github.io/statistical-methodology.html.

RESULTS

Our study results showed that several of our theoretically driven features (such as self-reported levels of anger, HRV, negative emotion words used, and mean audio intensity) were associated with conflict at levels significantly higher than chance, with an unweighted accuracy (UA) reaching up to 69.2 percent for anger and 62.3 percent for expressed negative emotion (task 1). This initial set of results is in line with laboratory research linking physiology and language use to couples' relationship functioning.^{4–8} When testing unimodal feature groups (task 2), the levels of accuracy reached

up to 66.1 and 72.1 percent for the female and male partners, respectively. Combinations of modalities based on EDA, ECG activity, synchrony scores, language and context and interaction resulted in UAs up to 79.6 percent (sensitivity = 73.5 percent and specificity = 85.7 percent) for females and 86.8 percent (sensitivity = 82.1 percent and specificity = 91.5 percent) for males. Using all features except self-reports, the UA reached up to 79.3 percent. These findings generally indicate that it is possible to detect a complex, psychological state with reasonable accuracy using multimodal data obtained in uncontrolled, real-life settings.

Because we aim to eventually detect conflict using passive technologies only—that is, without requiring couples to complete self-report surveys—we compared the UAs based only on self-reported MQI to combinations incorporating passive technologies (task 4). These results showed several setups where multimodal feature groups with and without self-reported MQI data significantly exceeded the

UA achieved from MQI alone. This indicates that the passive technologies added predictability to our modeling schemes.

Figure 2 shows receiver operating characteristic (ROC) curves for several feature combinations. The results showed that the area under the curve (AUC) for our multimodal indices reached up to 0.79 for females and 0.76 for males.

DISCUSSION

The results we report here provide a proof of concept that the data collected via mobile computing methods are valid indicators of interpersonal functioning in daily life. Consistent with laboratory-based research, we found statistically significant above-chance associations between conflict behavior and several theoretically driven, individually tested data features. We also obtained significant associations between conflict and both unimodal and multimodal feature groups with and without self-reported MQI included. In fact, our best-performing combinations of data features in

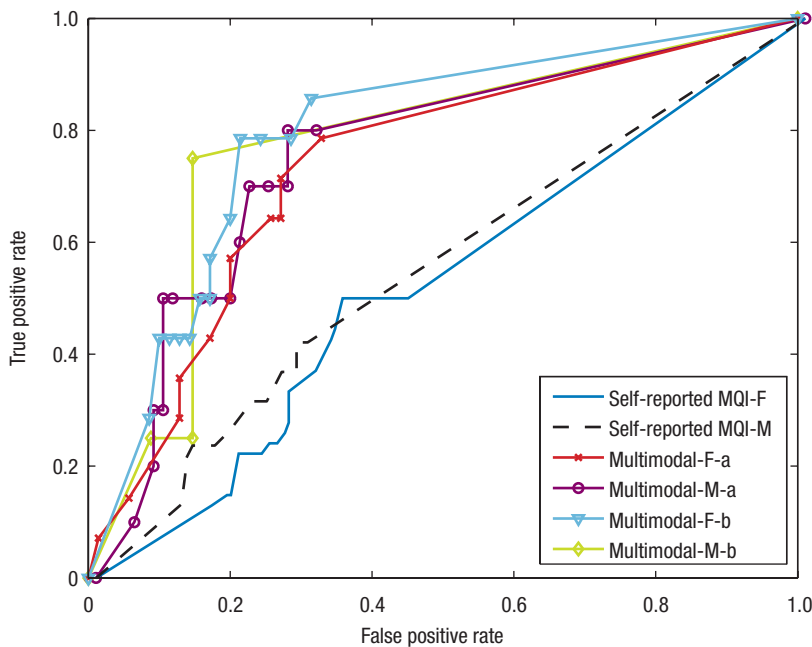


FIGURE 2. Receiver operating characteristic (ROC) curves for self-reported mood and quality of interaction (such as stressed, happy, sad, nervous, angry, and close) and multimodal feature groups. Multimodal-F-a = female EDA, psychological, personal, interaction, context; Multimodal-F-b = female self-reported MQI, EDA, psychological, personal, acoustic, interaction, context; Multimodal-M-a = male ECG activity, psychological, acoustic, interaction, context; Multimodal-M-b = male self-reported MQI, EDA synchrony, paralinguistic, personal, acoustic, interaction, context.

several cases reached or exceeded the UA levels obtained via self-reported MQI alone.

To our knowledge, the prototype model developed for this study is the first to use machine-learning classification to identify episodes of conflict behavior in daily life using multimodal, passive computing technologies. Our study extends the literature by presenting an initial case study indicating that it is possible to detect complex psychological states using data collected in an uncontrolled environment.

Implications

Couples communicate using complex, cross-person interactional sequences where emotional, physiological, and behavioral states are shared via vocal cues and body language. Multimodal feature detection can provide a comprehensive assessment of these interactional sequences by monitoring the way couples react physiologically, what

they say to each other, and how they say it. Couples in distressed relationships can become locked into maladaptive patterns that escalate quickly and are hard to exit once triggered. Detecting and monitoring these sequences as they occur in real time could make it possible to interrupt, alter, or even prevent conflict behaviors.

Thus, although preliminary, our data are an important first step toward using mobile computing methods to improve relationship functioning. The proposed algorithms could be used to identify events or experiences that precede conflict and send prompts that would decrease the likelihood that such events will spill over to affect relationship functioning. Such interventions would move beyond the realm of human-activity recognition to also include the principles of personal informatics, which help people to engage in self-reflection and self-monitoring to increase self-knowledge and improve functioning.

For example, a husband who is criticized by his boss at work might experience a spike in stress levels, which could be reflected in his tone of voice, the content of his speech, and his physiological arousal. Based on this individual's pattern of arousal, our system would predict that he is at increased risk for having an argument with his spouse upon returning home that evening. A text message could be sent to prompt him to engage in a meditation exercise, guided by a computer program, that decreases his stress level. When this individual returns home, he might find that his children are arguing and that his wife is in an irritable mood. Although such situations often spark conflict between spouses, the husband might feel emotionally restored following the meditation exercise and thus be able to provide support to his wife and avoid feeling irritable himself, thereby preventing conflict.

A second option is to design prompts that are sent after a conflict episode to help individuals calm down, recover, or initiate positive contact with their partners. For example, a couple living together for the first time might get in an argument about household chores. After the argument is over, a text message could prompt each partner to independently engage in a progressive muscle relaxation exercise to calm down. Once they are in a relaxed state, the program could send a series of prompts that encourage self-reflection and increase insight about the argument—for example, what can I do to communicate more positively with my partner? What do I wish I had done differently?

In addition to detecting conflict episodes, amplifying positive moods or the frequency of positive interactions could be valuable. Potential

behavioral prompts could include exercises that build upon the positive aspects of a relationship, such as complimenting or doing something nice for one's partner. Employing these methods in people's daily lives could increase the efficacy of standard therapy techniques and improve both individual and relationship functioning. Because the quality of our relationships with others plays a central role in our emotional functioning, mobile technologies thus provide an exciting approach to promoting well-being.

Limitations

Although the results from our classification experiments suggest that these methods hold promise, our findings should be interpreted in light of several limitations. Our system's classification accuracy, while moderately good given the task's inherent complexity, will need to be improved before our method can be employed widely. In our best-performing models, we missed 18 percent of conflict episodes and falsely identified 9 percent of cases as conflict. Classification systems that miss large numbers of conflict episodes will be limited in their ability to influence people's behaviors. At the same time, falsely identifying conflict would force people to respond to unnecessary behavior prompts, which could annoy them or cause them to discontinue use.

In our current model, classification accuracy is inhibited by several factors.

First, we relied on self-reports of conflict. Future projects could use audio recordings as an alternative, perhaps more accurate, way to identify periods of conflict.

Second, conflict and how it is experienced and expressed is highly variable across couples, with people

showing different characteristic patterns in physiology or vocal tone. For example, some couples yell loudly during conflict, whereas others withdraw and become silent. One method for addressing this issue could be to train the models on individual couples during an initial trial period. By tailoring our modeling schemes, we might be able to capture response patterns specific to each person and thereby improve our classification scores.

Third, we collapsed our data into hour-long time intervals, which likely caused us to lose important information about when conflict actually started and stopped. Many conflict episodes do not last for an entire hour, and physiological responding within an hour-long period could reflect various activities besides conflict. Using a smaller time interval would likely increase accuracy.

Fourth, outside of the synchrony scores, we did not take into account the joint effects of male and female

efficiency and intuitive nature, but different classifiers could provide additional benefits. With future iterations of these procedures, classification accuracies will likely improve, increasing the usefulness of these methods for influencing behavior in the real world. Still, the occasional false positive, which would prompt couples to engage in relaxation or other self-reflective activities, would most likely not be harmful, as long as the frequency of such events is minimal.

In addition to increasing our classification accuracy, several practical considerations for developing these methods deserve note. In the current study, we retroactively detected conflict, but interventions aimed at improving couple functioning would need to detect these episodes as they occur. In particular, such monitoring with audio data might be difficult because software must be capable of transcribing language, processing word counts, and extracting acoustic patterns in real

OUR MODELING SCHEME COULD BE USED TO CREATE INTERVENTIONS TO IMPROVE COUPLE FUNCTIONING AND QUALITY OF LIFE MORE GENERALLY.

responses. Considering these together (such as male and female vocal pitch increasing at the same time) could improve our results.

Future directions

Future research should examine the classification accuracy of multiple classifiers. We chose a decision tree for classification because of its running-time

time. We also manually transcribed the audio recordings, which could have resulted in less error than automated transcription techniques. We expect that methods for ongoing monitoring of physiology and behavior will continue to be developed and become more sophisticated over time.

There are also numerous ways that these methods could, albeit

ABOUT THE AUTHORS

ADELA C. TIMMONS is a doctoral candidate in clinical and quantitative psychology at the University of Southern California (USC). Her research interests include using ambulatory assessment methods to study interpersonal conflict and the impact of adversity on family functioning. Timmons received an MA in clinical and quantitative psychology from USC. Contact her at adelatim@usc.edu.

THEODORA CHASPARI is a doctoral candidate in the Signal Analysis and Interpretation Laboratory (SAIL) at USC. Her research interests include signal processing for wearable interfaces. Chaspari received an MS in electrical engineering from USC. She is a Student Fellow of IEEE. Contact her at chaspari@usc.edu.

SOHYUN C. HAN is a doctoral candidate in clinical psychology at USC. Her research interests include psychophysiology and family/couple conflict. Han received an MA in clinical psychology from USC. Contact her at sohyunha@usc.edu.

LAURA PERRONE is a lab manager for the Family Studies Project at USC. Her research interests include risk and resilience in families. Perrone received a BA in psychology from Pomona College. Contact her at lperrone@usc.edu.

SHRIKANTH S. NARAYANAN is a Niki and C.L. Max Nikias Chair in Engineering at USC, where he directs SAIL. Narayanan received a PhD in electrical engineering from the University of California, Los Angeles. He is a Fellow of IEEE and the Acoustical Society of America, International Speech Communication Association, National Academy of Inventors, and American Association for the Advancement of Science. Contact him at shri@sipi.usc.edu.

GAYLA MARGOLIN is a professor of psychology at USC, where she directs the Family Studies Project and conducts research on the intergenerational transmission of aggression and behavioral and physiological reactions in couples and families. Margolin received a PhD in clinical psychology from the University of Oregon. Contact her at margolin@usc.edu.

inadvertently, negatively impact the quality of life of the people using them. Answering surveys at regular intervals could be disruptive and annoying for participants, and wearing physiological sensors could be uncomfortable and even embarrassing. Future work should aim to make use of passive sensing technologies rather than rely on self-reported data. Many of our multimodal indices using only passive technologies performed well, suggesting it is possible to develop monitoring technologies that do not require active participation. In a similar vein, developing sensors that are smaller,

more stylish, and less obtrusive will be important for increasing couples' willingness to use the technology. Integrating these methods with trendy devices such as smartwatches or smart clothes will be an important component in encouraging the use of these methods.

Once highly accurate classification systems have been developed, future research will need to concentrate on designing interventions and testing them in real-world settings. While our intention is that these methodologies will improve the quality of life of the individuals using them, it is possible

that there will be unforeseen negative consequences or other barriers to utilization, so several iterations of interventions will need to be developed and tested to ensure their efficacy.

Beyond these recommendations, it is important that researchers take steps to keep data secure and minimize the risk of harmful privacy breaches. Care should also be taken to assess if couples are appropriate for the intervention and if there are any risks to safety—for example, individuals with suicidal ideation or violent couples would require interventions more appropriate to their needs.

Although preliminary, the proposed modeling scheme could eventually be used to create interventions that provide feedback and behavioral prompts to improve couple functioning and quality of life more generally. Ultimately, our method could be expanded to other types of relationships (such as parent-child) and to other types of behaviors (such as increasing positive interactions). Future iterations of this model will aim to improve upon the classification accuracy and incorporate ongoing, interactive monitoring to detect and predict conflict as it occurs in real time. ■

ACKNOWLEDGMENTS

This research is based on work supported by National Science Foundation grant BCS-1627272 (Margolin, PI), NIH-NICHD grant R21HD072170-A1 (Margolin, PI), SC CTSI (NIH/NCATS) through grant UL1TR000130 (Margolin, PI), NSF GRFP grant DGE-0937362 (Timmons, PI), an APA Dissertation Research Award (Timmons, PI), and NSF GRFP grant DGE-0937362 (Han, PI). This article's content

is the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health (NIH). We thank Naveen Kumar for his help in acoustic processing.

REFERENCES

1. S. Lyubomirsky and K. Layous, "How Do Simple Positive Activities Increase Well-Being?," *Current Directions in Psychological Science*, vol. 22, no. 1, 2013, pp. 57–62.
2. N. Bolger et al., "Effects of Daily Stress on Negative Mood," *J. Personality and Social Psychology*, vol. 57, no. 5, 1989, pp. 808–818.
3. P.R. Amato, "The Consequences of Divorce for Adults and Children," *J. Marriage and Family*, vol. 62, no. 4, 2000, pp. 1269–1287.
4. R.W. Levenson and J.M. Gottman, "Physiological and Affective Predictors of Change in Relationship Satisfaction," *J. Personality and Social Psychology*, vol. 49, no. 1, 1985, pp. 85–94.
5. A.C. Timmons, G. Margolin, and D.E. Saxbe, "Physiological Linkage in Couples and Its Implications for Individual and Interpersonal Functioning," *J. Family Psychology*, vol. 29, no. 5, 2015, pp. 720–731.
6. T. Chaspari et al., "Quantifying EDA Synchrony through Joint Sparse Representation: A Case-Study of Couples' Interactions," *Proc. Int'l Conf. Acoustics, Speech, and Signal Processing (ICASSP 15)*, 2015, pp. 817–821.
7. R.A. Simmons, P.C. Gordon, and D.L. Chambless, "Pronouns in Marital Interaction: What Do 'You' and 'I' Say about Marital Health?," *Psychological Science*, vol. 16, no. 12, 2005, pp. 932–936.
8. B.R. Baucom et al., "Correlates and Characteristics of Adolescents' Encoded Emotional Arousal during Family Conflict," *Emotion*, vol. 12, no. 6, 2012, pp. 1281–1291.
9. J.W. Pennebaker et al., "The Development and Psychometric Properties of LIWC2007," 2007; www.liwc.net/LIWC2007LanguageManual.pdf.

myCS Read your subscriptions through the myCS publications portal at <http://mycs.computer.org>

IEEE computer society

Read all your IEEE magazines and journals your WAY on

myCS

Introducing myCS, the digital magazine portal from IEEE Computer Society. Go beyond static, hard-to-read PDFs with an easily accessible, customizable, and adaptive experience.

There's No Additional Cost!

Now there's even more to love about your membership...



▶ LEARN MORE AT: mycs.computer.org