


Social Sensing for Psychology: Automated Interpersonal Behavior Assessment

Marianne Schmid Mast¹, Daniel Gatica-Perez², Denise Frauendorfer¹, Laurent Nguyen², and Tanzeem Choudhury³

¹Department of Organizational Behavior, University of Lausanne, Switzerland; ²Social Computing Group, Idiap Research Institute, Martigny, Switzerland; and ³Department of Information Science, Cornell University

Current Directions in Psychological Science
2015, Vol. 24(2) 154–160
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0963721414560811
cdps.sagepub.com


Abstract

In this article, we show how the use of state-of-the-art methods in computer science based on machine perception and learning allows the unobtrusive capture and automated analysis of interpersonal behavior in real time (*social sensing*). Given the high ecological validity of the behavioral sensing, the ease of behavioral-cue extraction for large groups over long observation periods in the field, the possibility of investigating completely new research questions, and the ability to provide people with immediate feedback on behavior, social sensing will fundamentally impact psychology.

Keywords

social interaction, interpersonal behavior, nonverbal behavior, ubiquitous computing, social sensing, honest signals

In this article, we discuss how knowledge, technology, and tools stemming from computer science can facilitate the assessment and analysis of interpersonal behavior via *social sensing* (Lazer et al., 2009; Pentland, 2008; Vinciarelli, Pantic, & Bourlard, 2009). Interpersonal behavior (or social interaction behavior) is defined as any verbal or nonverbal behavior directed toward or elicited by one or many real or imagined social interaction partners. Making sense of social interaction behavior is key to advancing our comprehension of human psychology because much of human behavior occurs in a social context (e.g., adults spend 32%–75% of their waking time with others; Mehl & Pennebaker, 2003).

Social sensing allows for verbal and nonverbal interaction behavior to be captured by mobile sensors or sensor-equipped environments and to be analyzed continuously and automatically in real time. The possibility for computing and sensing to be available everywhere, in conjunction with advances in computational data processing, makes it possible to extract behavioral cues without active input from participants or human coders—a concept often referred to as *ubiquitous computing*.

The use of novel sensors and technology developed by computer scientists for the psychological sciences was called *psychoinformatics* by Yarkoni (2012), and it encompasses mobile sensing via smartphones (Miller, 2012), data mining, computational analyses, and data collection via crowdsourcing (Lazer et al., 2009). In social

sensing, the emphasis extends beyond the generation and synthesis of large amounts of psychological data to the collection of social interaction data in the field. Social sensing means processing perceptual- and physical-sensor data (e.g., video, audio, location, and movement information) using smartphones, new types of wearable devices (e.g., smartwatches and bracelets), and instrumented environments (e.g., Microsoft Kinect sensors). It makes use of the computation power of mobile devices and cloud-computing resources (“clouds” being networks that provide data, software, and information to computers). The innovation and advancement of social sensing lies both in the processing speed, which makes it possible to study social interaction patterns of large entities (e.g., companies, families, classes) over long time periods (e.g., months or years), and in the immediate availability of feedback on verbal and nonverbal behavior.

How Social Sensing Works

Social sensing involves two phases (Pentland, 2008): the sensing of social interaction behavior via ubiquitous

Corresponding Author:

Marianne Schmid Mast, Department of Organizational Behavior, Faculty of Business and Economics, University of Lausanne-Dorigny, Internef, CH-1015 Lausanne, Switzerland
E-mail: marianne.schmidmast@unil.ch

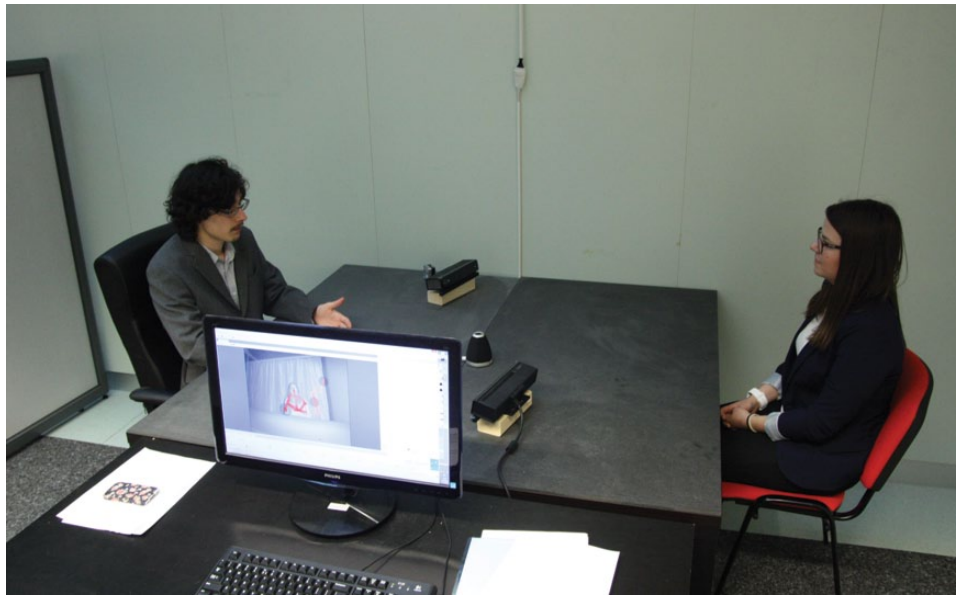


Fig. 1. Stationary social sensing of a dyadic interaction via Microsoft Kinect (for video and depth recording; shown to the front and left of each interaction partner) and Dev-Audio Microcone (for voice recording; shown at center of table).

computing devices and the extraction of verbal and non-verbal cues with computational models and algorithms.

Ubiquitous sensing

Ubiquitous computing frameworks are sensing devices that are incorporated into everyday human environments and capture signals that are informative of interaction behavior in a relatively unobtrusive manner. Social-sensing platforms can be mobile (e.g., smartphones) or stationary (e.g., *smart rooms*, or spaces equipped with cameras and microphones; see Fig. 1). Mobility allows for the registration of interpersonal behavior in the field and thus also captures spontaneous social interactions, in the spirit of experience sampling (a diary method that reminds participants to report their experience at specific moments during the day). In psychological research, a predecessor of mobile social sensing in the vocal domain is the electronically activated recorder (EAR; Mehl, Pennebaker, Crow, Dabbs, & Price, 2001), a modified portable audio device that registers thin slices of daily social interactions randomly throughout the day. Today, mobile sensing is typically performed using smartphones (Lane et al., 2010; Miller, 2012) that are equipped with multiple sensors (e.g., microphone, camera, accelerometer, GPS).

Wearable sensors (e.g., Google Glass or smartwatches) are rising in popularity and will become less invasive as technology advances. Their advantage is that they make it easier to extract behavioral cues. For instance, measuring gaze using the Google Glass camera is easier than

having to develop algorithms that extract gazing information from a video recording. However, wearable sensors can still be a threat to ecological validity because they can make a person more aware and controlling of his or her own interpersonal behavior (thus, e.g., evoking more socially desirable behavior) and may elicit reactions toward the person wearing the sensing device (which may, e.g., cause reluctance to interact with the person). Figure 2 shows some of the typically used ubiquitous sensing devices.

Automated extraction of behavioral cues

Through machine-perception and -learning technologies, computers are trained to extract the verbal and nonverbal behavior of the social interaction partners from the sensed data. For example, head nods in social interactions can be extracted via frontal video images from a sensor-equipped room (Nguyen, Odobez, & Gatica-Perez, 2012). The computational system starts by tracking the location of the person's face and estimates the horizontal and vertical visual motion of the face region. This motion information, extracted from a sequence of frames, is fed to a machine-learning module combined with human-annotated nod and non-nod labels (i.e., human coders rate whether the person in the video nodded or not). The system uses the data and labels to generate a model that can automatically classify nods in new video frames.



Fig. 2. Devices often used for social sensing (from left to right): Google Glass, Samsung Galaxy SIII smartphone, Dev-Audio Microcone, Microsoft Kinect, and Affectiva Q-Sensor (for physiological data).

Motion-sensing devices such as the Microsoft Kinect (Shotton et al., 2011) can record information on *depth* (i.e., the distance between the camera and objects seen in the images) in addition to video. This depth information enables computational algorithms to more accurately estimate various behavioral cues, such as eye gaze, body posture, head pose, and hand gestures.

Nonverbal vocal cues can be extracted from conversational data collected in the field (Choudhury & Basu, 2005; Wyatt, Choudhury, Bilmes, & Kitts, 2011). The computational capabilities of modern mobile devices allow the processing of speech in real time and the extraction of attributes of audio that can be used to infer the presence of human speech and paralinguistic aspects of speech (e.g., speaking rate, turn-taking behavior, pitch, energy level). This extraction can be achieved without recording the raw audio content or preserving the ability to reconstruct intelligible speech, thus respecting the privacy of the speaker. Nonverbal vocal cues can also be used to classify emotions (EmotionSense app; Rachuri et al., 2010) or detect stress from continuously captured interaction data (StressSense app; Lu et al., 2012).

Some extraction methods have proven to be accurate and are publicly available—for example, methods for extracting voice energy, pitch, and speaking rate (Basu, 2002); visual motion (Biel, Aran, & Gatica-Perez, 2011); head poses (Ba & Odobez, 2011); and facial expressions (Littlewort et al., 2011). Other cue extractions are more challenging to achieve—like, for instance, the automatic recognition of arm and body postures based on frontal video recordings (Marcos-Ramiro, Pizarro-Perez, Marron-Romera, Nguyen, & Gatica-Perez, 2013). Table 1 provides a non-exhaustive list of validated behavior-extraction methods.

Challenges for social sensing

To the extent that a computer algorithm uses a consistent decision-making framework (in the sense of a deterministic, repeatable output), the extraction is completely reliable. Because training of machine-learning algorithms often uses human-coded labels as *ground truth* (the criterion or gold standard used to assess the accuracy of the algorithm), the automated extraction is only as valid as is the ground truth. To avoid bias, extra care needs to go into training humans who code the ground truth.

The extraction process is intimately linked to the recording process, which means that the extraction algorithm works best if the data have been sensed in similar physical conditions (Lane et al., 2010). For example, if the extraction of conversational speech cues was developed using data collected only in indoor settings, the system may not perform well outdoors.

Integrated systems combine sensing and extracting into one device. Such devices are easy to use, as they do not require advanced computational skills, but they come with the challenge that what is extracted (e.g., duration of smiling) does not always correspond to what the researcher wants (e.g., frequency of smiling). On the other hand, if many researchers use the same integrated system, results among different studies become comparable, which is so far often not the case, because researchers often use different operationalizations of behavioral measures. To date, only a few integrated systems are available. One example is the Dev-Audio Microcone, a microphone array that can record and separate the voices of up to six speakers. The direct-speaker segmentation is a key advantage for the automated extraction of the vocal nonverbal cues (e.g., interruptions).

Privacy is a huge concern for ubiquitous sensing that is too broad to be discussed in detail here. Technical solutions for protecting privacy, such as content filtering for speech registration or privacy-sensitive speech-processing techniques (Wyatt et al., 2011), need to be developed (e.g., algorithms detecting whether a person is speaking while at the same time making verbal content impossible to be reconstructed). Moreover, computer science differs from psychology in its research traditions concerning data use and, thus, informed consent and privacy. In computer science, data (including video and audio) are readily shared with other researchers, and the same data are analyzed multiple times. In psychology, participants typically give consent for their data to be used only for the study in which the data have been collected, and, once the data are analyzed and published, they cannot be used for other original publications.

Table 1. Validated Behavior-Extraction Methods

Extracted cues	Sensing devices	Key words for describing the algorithms	Examples of nonverbal behaviors extracted	References
Arm and body postures	Video; Kinect	Hand likelihood map extraction (optical flow, face detection, edges, skin segmentation); hand tracking; 3D torso-pose extraction	Self-touch, gestures	Marcos-Ramiro, Pizarro-Perez, Marron-Romera, Nguyen, and Gatica-Perez, 2013
Head pose (as proxy for gaze)	Video	Dynamic Bayesian Network; contextual prior models; observations: speaking proportion, visual head pose based on exemplars	Gazing	Ba and Odobez, 2011
Eye gaze	Video; Kinect	3D head-pose tracking from depth data; eye gaze direction tracking from pose-corrected eye appearance	Gazing	Funes and Odobez, 2012
Face location and motion		Frame-based classification between nodding and not-nodding; Fourier representation of face-region optical flow	Nodding	Nguyen, Odobez, and Gatica-Perez, 2012
Facial-feature localization and FACS-based classification of facial expressions	Video	Face detection; facial-feature detection (eye corners and center, tip of the nose, mouth corners and center); feature extraction (using Gabor filters); action-unit recognition; expression intensity and dynamics	Smiling and FACS codes	Littlewort et al., 2011
Face detection and geometric analysis	Video	Upper-body detection, head detection inside upper-body areas, head-pose estimation, looking-at-each-other scoring between pairs of heads	Eye contact	Marin, Zisserman, Eichner, and Ferrari, 2014
Full-body pose	Video; Kinect	Body-part representation; depth image feature extraction; randomized decision-forest body part classification	Arms on hips, arms crossed	Shotton et al., 2011
Speech qualities	Close talk microphone, microphone array, smartphone mic	Speech qualities: voiced/non-voiced classification using hidden Markov models, pitch tracking; speaker segmentation: filter-sum beamforming using an array of microphones	Voice energy, pitch, speaking rate	Basu, 2002; Boersma and Weenink, 2013; Kiran et al., 2010; Lu et al., 2012

Note: FACS = Facial Action Coding System (Ekman & Friesen, 1978).

Making Sense of Social Sensing for Psychology

Social sensing has the power to put interpersonal behavior on center stage in psychology (Baumeister, Vohs, & Funder, 2007), not only because it will allow researchers to tackle more research questions linked to interpersonal behavior (given that it makes cue extraction manageable), but mostly because completely new research questions, impossible to study to date, can be addressed. For

instance, how daily social interaction behavior within large entities, such as companies, organizations, or work teams, is related to job satisfaction or to the entities' productivity, and how this link can change over time and with different events (e.g., a merger, the appointment of a new CEO) are questions that can be enlightened only by using social sensing. In an initial study, Olguin and colleagues (2009) assessed the social interaction behavior of 22 employees of a company who wore a social-sensing device (capturing proximity and speech, and

thus estimating time spent in face-to-face communication) for over 1 month. The researchers showed that the more an employee communicated (a composite of face-to-face communication and e-mail communication), the less satisfied he or she was with the job.

Automated performance prediction and automated social inferences

Automatically sensed and extracted interpersonal behavior can be related to real-world outcomes (e.g., task performance) and can be used to make social inferences. The automated extraction of behavioral cues that are diagnostic of a person's performance on a specific task can be used for training or for selection. As an example, consumer satisfaction with a company's call-center service was related to a series of automatically sensed and analyzed vocal-behavior cues of the call-center respondents. Once the vocal-behavior pattern associated with the greatest level of client satisfaction is identified, the social-sensing system can be used to select or train call-center employees (Pentland, 2008). Using social sensing, we identified the nonverbal-behavior pattern during a job interview that best predicted whether the job applicant would later be successful on a sales-type job (Frauendorfer, Schmid Mast, Nguyen, & Gatica-Perez, 2014). Recruiters could use this information to screen new applicants during job interviews by automatically sensing their nonverbal behavior and comparing it to the nonverbal-behavior patterns of applicants who were successful on the job.

Such *automated performance predictions* are data-driven. The meaning of the extracted behavior is generated by its relation to some external criterion (e.g., client satisfaction, job performance). Generalizability of the identified relations is restricted, however, and might be confined to the specific situation in which the data were gathered (e.g., the specific behavioral pattern that predicts job success in sales might not do so in accounting).

We make sense of our social interaction partners by perceiving their expressive verbal and nonverbal behavior and appearance, and, based on this perception, we draw inferences about their traits and states. Social-sensing machines, similar to a human perceiver, not only capture interpersonal behavior but also can be taught to draw social inferences. If we know which verbal- and nonverbal-behavior cues are related to being perceived as dominant, for example, social sensing can be used to inform us about dominance hierarchies in a group (*automated social inferences*). Such information can be fed back to the user via an application. In the future, such applications may inform users about how persuasive they are perceived to be while giving a speech, how trustworthy they are perceived to be in a negotiation task, or how attractive they are perceived to be in a dating context.

Effects of automated instant behavior feedback

The automation and speed makes it possible to provide people with instant feedback about their interpersonal behavior, about how they perform, or about how they are perceived by others. How such *automated instant feedback* affects people's behavior during their social interaction remains to be investigated. Also, immediate feedback can be used for training and to increase self-awareness, which can lead to personal development. Moreover, behavioral feedback can affect how we experience ourselves, as we sometimes make inferences about our own attitudes based on observations of our interpersonal behavior (Bem, 1972). To illustrate, when people realize that they talk more often to a specific person in their work group, they might conclude that they like this person.

The study of the effect of automated instant feedback is still in its infancy. Research has shown, for instance, that group collaboration can be enhanced using real-time visualized feedback about group dynamics (e.g., dominance structures) that are automatically sensed (e.g., via speaking-time distribution among group members; Kim, Chang, Holland, & Pentland, 2008). Another example is a visual feedback system developed for physician-patient interactions. Based on a physician's automatically assessed nonverbal cues in the audio and video stream, the physician's amount of expressed caring and dominance during the medical encounter is visually displayed in real time (Patel et al., 2013). Whether this feedback affects the way the physician communicates with the patient or consultation outcomes (e.g., adherence) remains to be tested. In a third example, Hoque, Courgeon, Martin, Mutlu, and Picard (2013) developed MACH (My Automated Conversation coach), a desktop application featuring a virtual coach for social-skills training. MACH has been used for training job applicants. It asks interview questions, uses ubiquitous computing for analyzing the applicants' behavioral cues while answering the questions, and renders individualized behavioral feedback (e.g., about an applicant's frequency of smiling, frequency of nodding, and speaking time during the interview). Training job applicants with MACH has been shown to improve interview performance compared to a control group (Hoque et al., 2013).

Using Social Sensing as a Psychologist

The development of algorithms to extract relevant behavioral cues still involves expert domain knowledge and expert computational skills. The latter can be a barrier to wide adoption among psychologists who are not trained as computer scientists. Therefore, if researchers

in psychology want to use these newly developed algorithms for extracting verbal- and nonverbal-behavior data, this may not be a straightforward plug-and-play process (except for in the few available integrated systems; e.g., the Microcone) and thus may be less usable without consulting or collaborating with computer scientists. Rather than considering this a limiting factor, we believe that it provides an opportunity for the emergence of a new, interdisciplinary field in which behavioral scientists collaborate with computer scientists and train the next generation of psychologists.

Recommended Reading

- Bishop, C. (2012). *Pattern recognition and machine learning*. New York, NY: Springer. An extensive treatise on pattern recognition and machine learning.
- Lazer, D., Pentland, A., Adamic, L., Aral, S., L., B. A., Brewer, D., . . . Van Alstyne, M. (2009). (See References). An account of how social interaction information is readily available from our everyday behavior (e.g., interacting with computers, being active in computer networks, or being sensed by security cameras).
- Miller, G. (2012). (See References). A comprehensive overview of how smartphones can be used for research in psychology.
- Pentland, A. (2008). (See References). An accessible introduction and exemplification of how computer science can capture and extract interpersonal behavior (social sensing).
- Yarkoni, T. (2012). (See References). An overview of how tools and techniques from computer and information sciences can be used with psychological data.

Acknowledgments

We thank Nora Murphy, Ioana Latu, Dario Bombari, Elena Canadas, and Valérie Carrard for their helpful comments on an earlier version of this manuscript.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This research was supported by Swiss National Science Foundation Grants CRSII2_127542 and CRSII2_147611 to M. Schmid Mast, D. Gatica-Perez, and T. Choudhury.

References

- Ba, S., & Odobez, J.-M. (2011). Multiperson visual focus of attention from head pose and meeting contextual cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *3*, 101–116.
- Basu, S. (2002). *Conversational scene analysis*. Cambridge, MA: MIT Department of EECS. Retrieved from <http://alummi.media.mit.edu/~sbasu/papers.html>
- Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-reports and finger movements: Whatever happened to actual behavior? *Perspectives on Psychological Science*, *2*, 396–403. doi:10.1111/j.1745-6916.2007.00051.x
- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York, NY: Academic Press.
- Biel, J.-I., Aran, O., & Gatica-Perez, D. (2011, July). *You are known by how you vlog: Personality impressions and non-verbal behavior in YouTube*. Paper presented at the AAAI Conference on Weblogs and Social Media, Barcelona, Spain.
- Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [Computer program]. Retrieved Version 5.3.51. Available from <http://www.praat.org/>
- Choudhury, T., & Basu, S. (2005). Modeling conversational dynamics as a mixed memory Markov process. In M. I. Jordan, Y. LeCun, & S. A. Solla (Eds.), *Advances in neural information processing systems* (pp. 281–288). Vancouver, British Columbia, Canada: MIT Press.
- Ekman, P., & Friesen, W. (1978). *Facial Action Coding System: A technique for the measurement of facial movement*. Palo Alto, CA: Consulting Psychologists Press.
- Frauendorfer, D., Schmid Mast, M., Nguyen, L., & Gatica-Perez, D. (2014). *A step towards automated applicant selection: Predicting job performance based on applicant nonverbal interview behavior*. Manuscript in preparation.
- Funes, K., & Odobez, J.-M. (2012, June). *Gaze estimation from multimodal kinect data*. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI.
- Hoque, M., Courgeon, M., Martin, J.-C., Mutlu, B., & Picard, R. W. (2013, September). *MACH: My automated conversation coach*. Paper presented at the Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Zurich, Switzerland.
- Kim, T., Chang, A., Holland, L., & Pentland, A. S. (2008, November). *Meeting mediator: Enhancing group collaboration using sociometric feedback*. Paper presented at the 2008 ACM Conference on Computer-Supported Cooperative Work, San Diego, CA.
- Kiran, K., Rachuri, M., Musolesi, M., Mascolo, C., Rentfrow, P. J., Longworth, C., & Aucinas, A. (2010, September). *EmotionSense: A mobile phones based adaptive platform for experimental social psychology research*. Paper presented at the 12th ACM International Conference on Ubiquitous Computing (UbiComp), Copenhagen, Denmark.
- Lane, N. D., Miluzzo, E., Hong, L., Peebles, D., Choudhury, T., & Campbell, A. T. (2010). A survey of mobile phone sensing. *Communications Magazine, IEEE*, *48*(9), 140–150. doi:10.1109/MCOM.2010.5560598
- Lazer, D., Pentland, A., Adamic, L., Aral, S., L., B. A., Brewer, D., . . . Van Alstyne, M. (2009). Life in the network: The coming age of computational social science. *Science*, *323*, 721–723. doi:10.1126/science.1167742
- Littlewort, G., Whitehill, J., Wu, T., Fasel, I., Frank, M., Movellan, J., & Bartlett, M. (2011, March). *The computer expression*

- recognition toolbox (CERT)*. Paper presented at the IEEE International Conference on Automatic Face and Gesture Recognition, Santa Barbara, CA.
- Lu, H., Rabbi, M., Chittaranjan, G. T., Frauendorfer, D., Schmid Mast, M., Campbell, A. T., . . . Choudhury, T. (2012, September). *StressSense: Detecting stress in unconstrained acoustic environments using smartphones*. Paper presented at the UbiComp, Pittsburgh, PA.
- Marcos-Ramiro, A., Pizarro-Perez, D., Marron-Romera, M., Nguyen, L. S., & Gatica-Perez, D. (2013, April). *Body communication cue extraction for conversational analysis*. Paper presented at the IEEE International Conference on Automatic Face and Gesture Recognition, Shanghai, China.
- Marin, J., M. J., Zisserman, A., Eichner, A., & Ferrari, V. (2014). Detecting people looking at each other in videos. *International Journal of Computer Vision*, *106*, 282–296.
- Mehl, M. R., & Pennebaker, J. W. (2003). The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations. *Journal of Personality and Social Psychology*, *84*, 857–870. doi:10.1037/0022-3514.84.4.857
- Mehl, M. R., Pennebaker, J. W., Crow, D. M., Dabbs, J., & Price, J. H. (2001). The Electronically Activated Recorder (EAR): A device for sampling naturalistic daily activities and conversations. *Behavior Research Methods, Instruments, & Computers*, *33*, 517–523. doi:10.3758/BF03195410
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science*, *7*, 221–237. doi:10.1177/1745691612441215
- Nguyen, L. S., Odobez, J.-M., & Gatica-Perez, D. (2012, October). *Using self-context for multimodal detection of head nods in face-to-face interactions*. Paper presented at the International Conference on Multimodal Interactions, Santa Monica, CA.
- Olguin, D. O., Waber, B. N., Taemie, K., Mohan, A., Ara, K., & Pentland, A. (2009). Sensible organizations: Technology and methodology for automatically measuring organizational behavior. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, *39*(1), 43–55. doi:10.1109/TSMCB.2008.2006638
- Patel, R. A., Hartzler, A., Czerwinski, M. P., Pratt, W., Back, A. L., & Roseway, A. (2013, April). *Leveraging visual feedback from social signal processing to enhance clinicians' nonverbal skills*. Paper presented at the CHI '13 Extended Abstracts on Human Factors in Computing Systems, Paris, France.
- Pentland, A. (2008). *Honest signals*. Cambridge, MA: MIT Press.
- Rachuri, K. K., Musolesi, M., Mascolo, C., Rentfrow, P. J., Longworth, C., & Aucinas, A. (2010, September). *EmotionSense: A mobile phones based adaptive platform for experimental social psychology research*. Paper presented at the Proceedings of the 12th ACM International Conference on Ubiquitous computing, Copenhagen, Denmark.
- Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R. K., . . . Blake, A. (2011, June). *Real-time human pose recognition in parts from a single depth image*. Paper presented at the IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, CO.
- Vinciarelli, A., Pantic, M., & Bourlard, H. (2009). Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, *27*, 1743–1759. doi:http://dx.doi.org/10.1016/j.imavis.2008.11.007
- Wyatt, D., Choudhury, T., Bilmes, J., & Kitts, J. A. (2011). Inferring colocation and conversation networks from privacy-sensitive audio with implications for computational social science. *ACM Transactions on Intelligent Systems and Technology*, *2*(1), 1–41. doi:10.1145/1889681.1889688
- Yarkoni, T. (2012). Psychoinformatics: New horizons at the interface of the psychological and computing sciences. *Current Directions in Psychological Science*, *21*, 391–397. doi:10.1177/0963721412457362