Psychophysiological Experimental Design for Use in Human-Robot Interaction Studies

Cindy L. Bethel, Jennifer L. Burke, Robin R. Murphy
Institute for Safety Security Rescue Technology
University of South Florida
Tampa, Florida, USA 33620
Email: (clbethel, jlburke4, murphy)@cse.usf.edu

Kristen Salomon
Department of Psychology
University of South Florida
Tampa, Florida, USA
Email: ksalomon@cas.usf.edu

Abstract— **This paper presents some foundations of psychophysiological concepts, related Human-Robot Interaction (HRI) studies, and outlines key experimental design issues associated with the use of psychophysiological measures in HRI studies. Psychophysiological measurements are one tool for evaluating participants' reactions to a robot with which they are interacting. Common physiological measures available for different experimental environments (e.g., laboratory setting vs. mobile field studies) are discussed. Issues and benefits are using psychophysiological measures are presented. Suggestions are given on what information should be obtained from participants before the psychophysiological measures are performed, to determine the impact on the actual measurements collected. There is focus on the placement of electrodes for each type of commonly used physiological measurement in HRI studies, timeframes for collecting usable data, and how to obtain a baseline measurement for comparison purposes to determine reactivity responses. There is a brief discussion of the effects of habituation, orientation, startle, and defensive responses by participants. Psychophysiological measures should be utilized as part of a multi-faceted approach to experimental design including self-assessments, participant interviews, and/or video-recorded data collection methods over the course of an experimental study. Two or more methods of measurement should be utilized for convergent validity. Although psychophysiological measures may not be appropriate for all HRI studies, they can provide a valuable evaluation tool of participants' responses when properly incorporated into a multi-faceted experimental design.**

I. Introduction

Human-robot interaction is an emerging field of research; however the development of effective methods to evaluation for these interactions is lacking. In general, methods of testing and evaluation have been adopted and modified from such fields as human-computer interaction, psychology, and social sciences [1]. The manner in which a human interacts with a robot is similar but not identical to interactions between a human and a computer or a human interacting with another human. As robots become more prevalent in day-to-day life, it will be increasingly important to have accurate methods of evaluating how humans feel about their interactions with robots and how they interpret the actions of the robots. Steinfeld *et al.* [2] describe the need for the development of common metrics as an open research issue in HRI. In order to obtain credibility in the research community, HRI studies need to be supported by quality experimental designs with multi-faceted methods of measurement to provide convergent validity.

There are three primary methods of evaluation of participants' responses to robots with which they are interacting in HRI studies: (1) *self-report measures*, (2) *behavioral measures*, and (3) *psychophysiological measures* [1]. The most common methods utilized in most HRI studies are self-report and behavioral measures; however psychophysiological measures and task performance metrics have been applied in some HRI studies. Each method has its advantages and disadvantages; however some of the disadvantages can be overcome by using more than one method of evaluation for redundancy and reliability [1].

The design of a quality research study for use in human-robot interaction applications that produces results that are verifiable, reliable, and reproducible is a major challenge. Psychophysiological measurements can complicate this process because the results are not always straightforward and confounds can lead to misinterpretation of data. There is a tendency to attribute more meaning to results because of the tangible nature of the recordings. Information needs to be obtained from participants prior to beginning a study to help in reducing these confounds (e.g., health information, state of mind). Multiple signals should be used in order to find correlations in the results to provide more validity and reliability.

The use of participants' self-reports is one of the most commonly used methods of evaluation in HRI studies. Participants provide a personal report of their motives and feelings about an object, situation, or interactions [3], [4]. A common problem is participants may not answer exactly how they are feeling but rather answer questions as they feel others would answer or in a way they think the researcher wants them to answer. Some participants may not always be cooperative in answering the questions presented. Participants' responses could be dependent on their mood and state of mind on the day of the study [3], [4]. Additionally, many of the psychological or psychometric scales developed and used in HRI studies have not been designed specifically for HRI; therefore the measures may require modification, which can impact the overall scale reliability [1]. The development of reliable scales designed specifically for HRI is an area that requires further research.

The next most common method of evaluation in HRI studies are behavioral measures. This method of evaluation involves some form of observation of the behaviors of participants. Johnson and Christensen [4] define observation as "the watching of behavioral patterns of people in certain situations to obtain information about the phenomenon of interest." Another issue associated with behavioral measures is the "Hawthorne effect", a situation where if participants know that they are being observed, it will impact their behaviors [3], [4]. Behavioral measures require two or more independent raters to evaluate the observations [1], [3], [4]. A problem associated with independent raters is the ability for observers to have agreement on the observed behaviors to assess reliably participants' behaviors that have been recorded. The ability of the raters to categorize consistently the behaviors can be another challenge to overcome in behavioral measures [3], [4]. The benefit of using behavioral measures is that researchers are able to record the actual behaviors of participants and do not have to rely on participants to report accurately their intended behaviors or preferences [3].

None of these measures alone are sufficient to interpret accurately the responses of participants to a robot with which they are interacting. In order for a study to have corroboration and consistency in its evaluations, at least two methods of measurement should be used [1], [3], [4]. Steinfeld *et al.* [2] discuss an approach of developing common metrics for human-robot interaction; however this approach is oriented more towards an engineering perspective and does not completely address the social interaction perspective and evaluation. Both perspectives appear to have value but further investigation is required.

The paper begins with foundations related to psychophysiological measures, including general information, terminology, acitivities and response types, commonly used psychophysiological measures in different environments, advantages of using psychophysiology, common issues, habituation, and response problems (Section II). In Section III, coverage is given to a limited number of psychophysiological evaluations in HRI studies. Psychophysiological experimental design issues and considerations in HRI studies are presented in Section IV.

II. Psychophysiology Foundations

Psychophysiology focuses on the interaction between the mind and body [5]. John Stern defined psychophysiology as "any research in which the independent variable (the subject's response) is a physiological measure and the independent variable (the factor manipulated by the experimenter) a behavioral one" [5]. Psychophysiological measures are useful evaluation tools for HRI studies if used appropriately. There is a tendency in the research community to attribute the results of psychophysiological measures to specific causes and/or emotions; however the readings may be attributed to multiple factors or confounds making it difficult to isolate the specific factors [1], [6], [7], [8], [9], [10]. Because the results are visible as tangible output, researchers have a tendency to make stronger assumptions than may be accurate [5].

A. Psychophysiological Activities and Response Tendencies

Psychophysiological measures can be analyzed in terms of three basic types of activities:(1) *spontaneous*, (2) *tonic*, and (3) *phasic*. A *spontaneous response* is a measurable response that occurs when there is no known stimulus presented [5]. *Tonic responses* are the baseline or resting level responses of activity for a particular physiological measure. This level occurs when participants being measured are not responding to a known or unknown stimulus. This measure is typically taken at the end of a resting period, typically in the last three to five minutes of a ten minute resting period. The *phasic response* occurs when participants have discrete responses to a specific or known stimulus (*an evoked response*). It is important during this type of measurement to account for internal and external stimuli that may impact participants' responses to the presented stimuli. This can be accomplished through self-reports or interviews to make sure other factors (e.g., state of mind, mood, health) are not contributing to the measured responses.

There are two types of psychophysiological response tendencies, (1) *stimulus-response specificity* and (2) *individual-response stereotypy* that commonly occur in psychophysiological studies; however they are not mutually exclusive. *Stimulus-response specificity* is when a stimulus or stressor produces a similar pattern of physiological responses among most subjects or participants studied. Typically, more than one type of response is involved but the pattern of responses would be consistent among most participants subjected to the same stimulus or stressor. *Individual-response stereotypy* occurs when a few individuals exhibit a pattern of responses different than expected to a specific stimulus or stressor. Also, individuals may have the same idiosyncratic response to different stressors, no matter what the stressors may be [5].

B. Common Psychophysiological Measures

There are numerous types of psychophysiological measures available to researchers; however they tend to be application and environment specific. If a participant is in a laboratory setting, in a fixed location connected directly to stationary equipment, the available methods of measurement are numerous. The most common measures used in a controlled laboratory setting are: cardio-vascular system (*heart rate variability (HRV), cardiac output, interbeat interval (IBI), blood pressure (BP)*); electrodermal activity (*skin conductance activity (SCA), skin conductance response (SCR)*); respiratory system (*respiratory sinus arrhythmia (RSA)*); muscular system (*electromyography (EMG)*); and brain activity (*electroencephalography (EEG) and imaging*) [5], [11], [12], [13]. Some of these measures such as EEG and imaging are not conducive to HRI studies because the participant must be in a specific testing location, with little or no movement, and directly connected to the testing and data collection equipment for accurate recording of data.

The most common measures used in Human-Robot Interaction studies include: HRV, IBI, BP, SCR, RSA, and EMG [7], [10], [8], [14], [9]. These psychophysiological measures are available in ambulatory recording units which allows participants to be placed in a field location or even allows them to be mobile. However, in most cases the measures must be adjusted for movement artifacts or signal noise.

C. Advantages of Using Psychophysiology in HRI Studies

There are advantages of using psychophysiological measures in human-robot interaction applications and experiments. The primary advantage is that participants cannot consciously manipulate the activities of their autonomic nervous system; therefore the readings reflect participants' state during the time of evaluation [1], [6], [7], [8], [9], [10]. Additionally, psychophysiological measures offer a non-invasive method that can be used to determine the stress levels and reactions of participants interacting with technology [6], [7], [8], [9], [10].

D. Issues with Using Psychophysiology in HRI Studies

The use of psychophysiological measures can pose significant challenges. The ability to gather reliable data from participants in real-world human-robot interaction scenarios can be difficult [1]. Proper preparation of the area where electrodes are placed, location of electrode placement, and making sure appropriate amounts of conducting gel or paste are used are factors which impact the quality of recordings.

It is important and sometimes complicated to determine baseline values; and the law of initial values can make this issue even more problematic [5], [11], [12], [13]. The "Law of Initial Values" indicates that the initial state of a participant determines the level of possible changes in that state that can occur [5]. If participants are recorded at a high initial state, then further increases in physiological response levels are limited, similarly if participants start at a lower initial state, further decreases in response levels will be limited.

E. Habituation and Response Factors

Habituation is a factor that reduces participants' responses due to repetitive presentation of the same or similar stimuli in psychophysiological studies. There are two primary types of habituation: (1) *short-term* - occurs during a single evaluation session and (2) *long-term* - occurs over multiple settings over a period of days or weeks. Habituation occurs more rapidly when stimuli is presented frequently. One method to reduce the effects of habituation is to ask participants to complete a rating questionnaire between the presentation of each stimulus to induce a behavioral response. Habituation has its strongest effects towards the end of any study and needs to be considered in the evaluation of data during psychophysiological studies [5].

Three response factors may need to be considered in any psychophysiological study: (1) *orienting response*, (2) *startle response*, and (3) *defensive response*. The *orienting response* relates to how a participant responds to novel stimuli. It causes the participant to orient towards the novel stimuli to identify what it is and its location. Once the participant determines the stimuli is not a threat or concerning, the effects of the orienting response are inhibited. Therefore, depending on the type of test the first few seconds following the presentation of novel stimuli should in some cases be disregarded when evaluating the data depending on the application. There are some cases where researchers may want to evaluate or measure the orienting response toward a robot presented to participants as part of their study. The *startle response* occurs due to a sudden onset of an intense type of stimuli (e.g., door slam or lightning strike). Data collected after a startle response would be handled similar to an orienting response by disregarding the data for the first few seconds following the presentation of the stimuli; however this would be dependent on the focus of the research study. The *defensive response* occurs as a result of intense, threatening, dangerous, or painful stimuli. This type of response prepares the body for "fight or flight" activation in participants. The inclusion of this data would depend on the type of study conducted [5].

III. Human-Robot Interaction Studies Using Psychophysiological Measures

Five studies have been conducted using psychophysiology using actual robots in the human-robot interaction community. Rani *et al.* [7] performed a study that included the development of a robotic system that monitored a participant's anxiety level and would respond appropriately to assist the participant. They used self-report and performed measurements for HRV, IBI, skin conductance response (SCR), and electromyography (EMG) of the corrugator supercilii and masseter muscles. The authors found that cardiac activity, SCR, and EMG were all good indicators of anxiety and correlated with the participant's self-report. One limitation of the study was that the authors performed experiments on one participant.

Lui *et al.* [10] performed a study in which a robot would modify its behavior based on the psychophysiological responses of the participants with which it was interacting. In this study 14 participants performed two different versions of robot-based basketball (RBB), counterbalanced. In one version, the game difficulty was based on performance and in the other condition game difficulty was based on the psychophysiological readings for the participant. They used self-report of anxiety and measured cardiovascular activity (IBI, relative pulse volume, pulse transit time, and pre-ejection period), SCR (tonic and phasic), and EMG activity (from the corrugator supercilii, zygomaticus, and upper trapezius muscles). They reported that 11 out of 14 participants had lower anxiety levels playing the psychophysiological measures version of RBB and 9 out of 14 had improved performance with the anxiety-based version.

A group of researchers from Japan, Itoh *et al.* [8], developed their own bioinstrumentation system to measure human stress when interacting with a fixed humanoid robot that had only an upper body. Their wearable system measured ECG, heart rate, IBI, respiration, EDA (changes in skin resistance), pulse wave transit time, blood pressure, and upper body movements. If participants' stress level increased the robot would modify its actions to decrease participants' stress levels by shaking their hand. They found that blood pressure and pulse wave transit time were disrupted due to movement artifacts and the data was not useful. The physiological responses did indicate a reduction in participants' stress after the robot shook their hands [8].

Kulić and Croft have reported two studies using psychophysiological measures in HRI [14], [9]. These two studies they performed utilized a robot manipulator arm. Participants were evaluated for their anxiety levels while experiencing various movements of the robotic arm. The robot performed two sets of movements: (1) pick and place, and (2) reach and retract. There were also two scenarios for each movement type; a set of planned standard motions and a set of planned safe motions. In both studies they measured heart rate, SCR, and EMG of the corrugator supercilii muscle. The authors determined that participants' arousal responses could be most reliably detected with SCR, but heart rate also had a contributory impact, although less reliable. Psychophysiological responses were compared with participants' self-reports. On average, based on both psychophysiological and self-report data, 94% of the time arousal levels were detected. On average, 75% of the time the correct valence was detected in participants. EMG of the corrugator supercilii muscle was not a reliable predictor of participants' valence (positive or negative) and arousal level in the interactions between the robot and participants. In most participants no changes were noted. The results indicate that participants have lower arousal responses with the planned safe motions of the robotic manipulator arm and feel calmer when the robot motions are slower.

IV. Psychophysiological Experimental Design in HRI

Psychophysiological measures can be a useful assessment tool in HRI studies; however there are several factors that need to be considered when designing these studies. In order to determine possible confounding factors and participants' status prior to obtaining measurements, it is important to perform pre-testing assessments. Studies should be designed such that habituation effects are reduced, typically through some type of self-assessment between the presentation of each stimulus. As part of the design process, researchers must determine the most appropriate measures to use. More than one psychophysiological signal should be utilized when designing experimental studies to corroborate data and for more accurate interpretation of participants' responses [1], [6], [7], [8], [9], [10]. Once the types of measures are decided, it is necessary to determine the appropriate placements of electrodes for each type of psychophysiological signal. For convergent validity, it is important to utilize other methods of assessment such as self-reports, interviews, and/or video-recording methods. As part of the study, it is necessary to obtain appropriate informed consent from participants.

A. Pre-testing Information

Prior to performing any psychophysiological measures certain information should be obtained from participants to help reveal any possible confounding factors. Caffeine, nicotine, medications, and even food can have an impact on the cardiovascular system and needs to be considered when evaluating participant data. A pre-test questionnaire should be given to participants to provide data regarding the time of their last meal; whether they drink caffeine and if so the last time they had caffeine; whether they smoke and if so the last time they smoked; and what medications they are taking at the time of testing that could impact the results.

Determining how participants feel at the time of the assessment, how they describe themselves, and what emotions they are experiencing at the time of the study can provide useful information when evaluating the physiological signals obtained during the testing. There are two self-report assessments that can provide this information: (1) *the State Trait Anxiety Inventory (STAI)* which evaluates how a person describes themselves and how they generally deal with life and situations that may induce anxiety [15] and (2) *the Positive and Negative Affect Schedule (PANAS)* which describes how participants are feeling and what emotions they are experiencing at the time of the study [16]. This provides the researcher some insight as to other factors that might impact the results from participants.

Prior to starting experiments and before attaching any physiological measurement equipment, participants should receive some basic instruction regarding the study and be presented with the appropriate informed consent form(s) for their signature. A separate form is needed for any audio/video-recording that may be performed during the course of the experiments [5].

B. Habituation Reduction During HRI Studies

Habituation can impact psychophysiological measures when repeated stimuli is presented to participants frequently. This factor can be reduced by having participants complete assessments between the presentation of stimuli. The self-assessment manikin (SAM) [17] evaluates participants' valence, arousal, and dominance feelings in response to the presented stimulus and can be completed either through paper-and-pencil or computerized forms and can be used to reduce the effects of habituation.

C. Post-Testing Assessments

Following an assessment, a post-study debriefing should be conducted with participants to determine their reactions to the study and to inform them of the purpose of the study. A post-study self-assessment or personal interview can be performed with participants to determine how they felt about the study and if there were any situations that may have influenced their responses during the study. This information should be collected to aid in the design of future studies.

D. Psychophysiological Testing and Electrode Placement

The determination of the most appropriate psychophysiological measures to use in HRI studies is an open research question [18]. From prior studies, it appears that HRV and SCR have been fairly reliable in detecting anxiety in participants [8], [7]. When measuring HRV and SCR it is important to obtain tonic or baseline readings for participants. These baseline readings are subtracted from the phasic response measurements to determine the variability levels experienced by participants. There are mixed results related to EMG signals; however EMG was included in four of the five studies presented in Section III of this paper [7], [10], [14], [9]. RSA was only used in one of the studies presented and the results obtained were not presented or useful [8].

The proper location of electrodes and the application of the appropriate amount of conducting gel or paste impacts the quality of the data collected. There are disposable pre-filled electrodes available for measuring cardiac physiological signals; however for EMG and EDA signal measurements, reusable electrodes with conducting paste must be used. For measuring cardiac activity, a modified lead II configuration is typically used (See Figure 1).

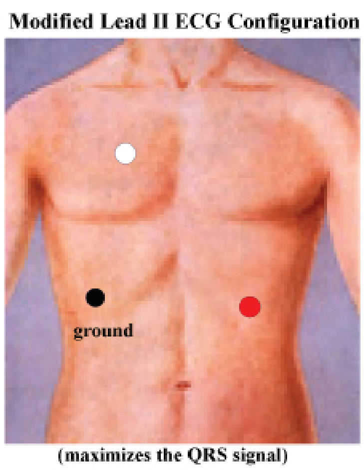

Fig. 1. Modified Lead II Configuration for ECG Measures

Dawson *et al.* [12] recommends the placement of electrodes on the volar surface of the distal phalanges for obtaining the most reliable electrodermal activity signal measurements (See Figure 2). There are other electrode placements for EDA such as on the palmar surface; however these placements do not provide the same level of reliability.

There are several different electrode placements for measuring EMG signals depending on the experiments being conducted. To obtain measurements related to frowning, electrodes should be placed on the corrugator supercilii muscle at the brow line. Electrodes are placed in the cheek area, on the zygomaticus major muscle, to determine if participants are smiling. Electrodes can be placed on the masseter muscle located near the jawline close to the ear to measure jaw clenching behavior associated with stress and anxiety. There are many electrode placements used in psychophysiological studies depending on what signals are of interest to the researcher; however these placements are the most common in previous HRI studies (See Figure 3).

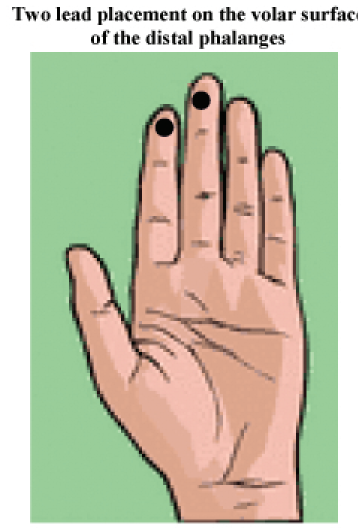

Fig. 2. Recommended electrode placement for EDA [12]

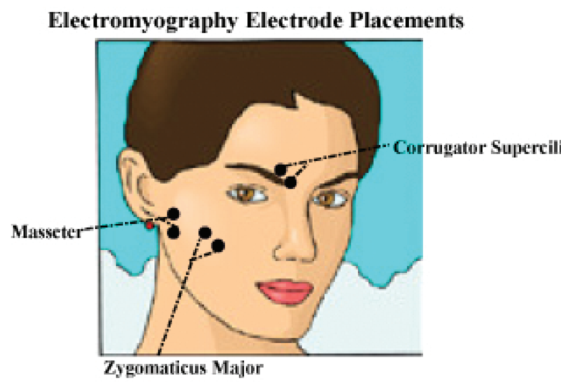

Fig. 3. Common electrode placements for EMG signals

V. Conclusion

Psychophysiological measures can be useful tools in the design of HRI studies to assess humans' responses to robots with which they are interacting. These measures have various issues and problems; however they do provide a means of evaluating participants' responses in a non-invasive manner that cannot be consciously manipulated by the participant [1], [6], [7], [8], [9], [10]. When designing a HRI study using psychophysiological measures it is important to evaluate what is to be measured, and what is the most appropriate tests to assess participants' responses. The selection of the most appropriate measure for use in a particular HRI study is an open question requiring further research [18]. It is important to select more than one psychophysiological measure when designing a study in order to determine if there are correlations and corroboration between these signals for higher reliability and validity in the study [1], [6], [7], [8], [9], [10]. One significant problem associated with psychophysiological studies is the tendency to attribute specific causation to the responses which may not be accurate due to confounds or other contributory factors [5].

In the process of designing a HRI study it is important to develop appropriate pre-testing assessments to obtain valuable information that can impact the validity and meaning of the responses recorded. It is also important to conduct some type of assessment or evaluation during the course of the experiments to reduce the possibility and/or impact of habituation [5]. Following the testing period, researchers should design some form of assessment or interview process to debrief participants regarding the purpose of the study and to determine any issues that arose as part of the experiments or that may impact the results of the study. Informed consent should be obtained from each participant in the study prior to performing any experiments, with separate consent forms for any audio/visual-recording.

The use of psychophysiological measures are just one facet of a comprehensive HRI study. For convergent validity, it is important to incorporate not only multiple pscyhophysiological measures but also one or more other methods of assessment such as: self-report assessments, interviews, and/or video-recording. One method of evaluation and measurement is not going to be sufficient for a complete evaluation of a human's response to a robot in human-robot interaction studies [1]. Instead, research should focus on developing a diverse set of complimentary measures that capture the full range of human-robot interactions.

ACKNOWLEDGMENT

This work is supported in part under a National Science Foundation Graduate Research Fellowship - Award Number DGE - 0135733. Additionally, this work was sponsored, in part, by the US Army Research Laboratory under Cooperative Agreement W911NF-06-2-0041. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ARL or the US Government. The US Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation hereon.

REFERENCES

[1] C. D. Kidd and C. Breazeal, "Human-robot interaction experiments: Lessons learned," in *Proceeding of AISB'05 Symposium Robot Companions: Hard Problems and Open Challenges in Robot-Human Interaction*, Hatfield, Hertfordshire, UK, 2005, pp. 141 – 142.

[2] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common metrics for human-robot interaction," in *1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. Salt Lake City, Utah, USA: ACM Press, 2006.

[3] D. G. Elmes, B. H. Kantowitz, and H. L. Roediger III, *Research Methods in Psychology*, 8th ed. Belmont, CA: Thomson-Wadsworth, 2006.

[4] B. Johnson and L. Christensen, *Educational Research Quantitative, Qualitative, and Mixed Approaches*, 2nd ed. Boston, MA: Pearson Education, Inc., 2004.

[5] R. M. Stern, W. J. Ray, and K. S. Quigley, *Psychophysiological Recording*, 2nd ed. New York, NY, USA: Oxford University Press, 2001.

[6] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: analysis of affective physiological state," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 10, pp. 1175 – 1191, 2001.

[7] P. Rani, N. Sarkar, C. A. Smith, and L. D. Kirby, "Anxiety detecting robotic system - towards implicit human-robot collaboration," *Robotica*, vol. 22, no. 1, pp. 85 – 95, 2004.

[8] K. Itoh, H. Miwa, Y. Nukariya, M. Zecca, H. Takanobu, S. Roccella, M. C. Carrozza, P. Dario, and T. Atsuo, "Development of a bioinstrumentation system in the interaction between a human and a robot," in *International Conference of Intelligent Robots and Systems*, Beijing, China, 2006, pp. 2620 – 2625.

[9] D. Kulić and E. A. Croft, "Physiological and subjective responses to articulated robot motion," *Robotica*, vol. Forthcoming, no. -1, p. 15, 2006.

[10] C. Liu, P. Rani, and N. Sarkar, "Affective state recognition and adaptation in human-robot interaction: A design approach," in *International Conference on Intelligent Robots and Systems*, Beijing, China, 2006, pp. 3099 – 3106.

[11] K. A. Brownley, B. E. Hurwitz, and N. Schneiderman, "Cardiovascular psychophysiology," in *Handbook of Psychophysiology*, 2nd ed., J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, Eds. USA: Cambridge University Press, 2000, pp. 224 – 264.

[12] M. E. Dawson, A. M. Schell, and D. L. Filion, "The electrodermal system," in *Handbook of Psychophysiology*, 2nd ed., J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, Eds. USA: Cambridge University Press, 2000, pp. 200 – 223.

[13] L. G. Tassinary and J. T. Cacioppo, "The skeletomotor system surface electromyography," in *Handbook of Psychophysiology*, 2nd ed., J. T. Cacioppo, L. G. Tassinary, and G. G. Berntson, Eds. USA: Cambridge University Press, 2000.

[14] D. Kulić and E. A. Croft, "Anxiety detection during human-robot interaction," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, 2005, pp. 616 – 621.

[15] C. Spielberger, R. Gorsuch, and R. Lushene, *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologist Press, 1970.

[16] D. Watson, L. A. Clark, and A. Tellegen, "Development and validation of brief measures of positive and negative affect: the panas scales," *Journal of Personality and Social Psychology*, vol. 54, no. 6, pp. 1063 – 1070, 1988.

[17] M. M. Bradley and P. J. Lang, "Measuring emotion: The self-assessment manikin and the semantic differential," *Journal of Behavioral Therapy and Experimental Psychiatry*, vol. 25, pp. 49 – 59, 1994.

[18] C. L. Bethel, R. R. Murphy, and J. L. Burke, "Survey of psychophysiology measurements applied to human-robot interaction," in *16th IEEE International Symposium on Robot & Human Interactive Communication*, Jeju Island, South Korea, 2007, p. (submitted).