Building a Literal Bridge Between Robotics and Neuroscience using Functional Near Infrared Spectroscopy (NIRS)

Megan Strait
Tufts University
200 Boston Avenue
Medford, MA 02155
megan.strait@tufts.edu

Matthias Scheutz
Tufts University
200 Boston Avenue
Medford, MA 02155
matthias.scheutz@tufts.edu

ABSTRACT

Functional near infrared spectroscopy (NIRS) is a promising new tool for research in human-robot interaction (HRI). The use of NIRS in HRI has already been demonstrated both as a means for investigating brain activity during human-robot interactions, as well as in the development of brain-robot interfaces that passively monitor a person's cognitive state to adapt robots' behaviors accordingly. In this paper, we survey the utility of NIRS and its range of applications in HRI. We discuss both some exemplary applications as well as several pressing challenges to the deployment of NIRS in more realistic settings.

Keywords

Functional near infrared spectroscopy (NIRS), human-robot interaction, brain-computer interfaces

1. INTRODUCTION

One of the main goals of research in human-robot interaction (HRI) is to develop effective interactions between people and robots. While current modes of interaction primarily rely on visual and verbal communication, there is a growing body of work employing brain-based sensors as an additional information channel to equip robots with an awareness of (and thus, ability to respond appropriately to) a person's cognitive state (for a review, see [6]). Given recent advances in brain-imaging technology, inexpensive sensors are becoming increasingly accessible to researchers and consumers alike. This production of affordable sensors (that are also small and wireless) is promising for HRI, as that allows for the wearing of such sensors in a variety of human-robot interaction settings without being intrusive. Neural data, in particular, is highly relevant for HRI research as it can complement traditional survey methods such as questionnaires and thus yield further understanding of users' genuine responses toward robots during an interaction (e.g., [41]). Moreover, a growing body of work suggests that it can be used for real-time detection of a person's cognitive or affective state (e.g., [16, 26, 45, 50]), which could be used to further inform a robot's user model and thus modify its behavior accordingly. Already, socially-aware robots that can capture *and respond to* changes in anxiety, attention, arousal, and other states have been found to be more effective in engaging people (e.g., [45]). For these reasons, research on neurophysiological signals has been attracting the attention of researchers in the HRI community over recent years [4, 6, 32, 44, 41].

Amongst work employing brain-based measures, electroencephalography (EEG) is the most commonly-used technology for measuring brain activity as it provides high temporal resolution (which allows for the monitoring of quickly-changing processes such as attention; see [13] for a review). However, it is also disadvantaged by limited spatial resolution, thus reducing its efficacy for measuring region-specific brain activity. Conversely, high spatial resolution can be achieved using functional magnetic resonance imaging (fMRI), but at a cost to both participant mobility and temporal resolution [6]. Hence, functional near infrared spectroscopy (NIRS or fNIRS) is a promising alternative, achieving some middle ground in both spatial resolution and mobility constraints between the EEG and fMRI technologies (e.g., [20]).

Within the field of human-robot interaction, two distinct usage patterns have emerged with respect to NIRS: (1) as a measurement tool for evaluating interactions between humans and robots (e.g., [23, 41, 44]), and (2) as a passive input channel to better inform a robot of a person's cognitive or affective state (e.g., [45]). In this paper, we discuss (1) the utility (i.e., cognitive states measurable by NIRS) and applications of NIRS in HRI, and (2) the limitations of NIRS with respect to naturalistic and unconstrained interactions of more realistic HRI settings. We start with an overview of the technology, followed by a survey of its applications to HRI, detailing selected studies as exemples. We then discuss the challenges of using NIRS in HRI and review general guidelines for future research directions.

2. NEAR INFRARED SPECTROSCOPY

Functional near infrared spectroscopy (NIRS, also referred to as fNIRS or fNIR) is a neuroimaging methodology similar to fMRI. By measuring changes in oxygenated and deoxygenated hemoglobin, NIRS non-invasively gauges regional cortical activation (e.g., [20]). However, as the hemodynamic response is a secondary response following neural activity, it lags behind the triggering neuronal events by several seconds (e.g., [10]). This response peaks about 5 seconds after the onset of the stimulus and then dips back down as homeostasis is reestablished.

The change in hemoglobin due to neural activity is called the *hemodynamic response* (HDR) and is measureable with NIRS due to the differences in molar absorptivity of hemoglobin versus the surrounding tissue. Specifically, in

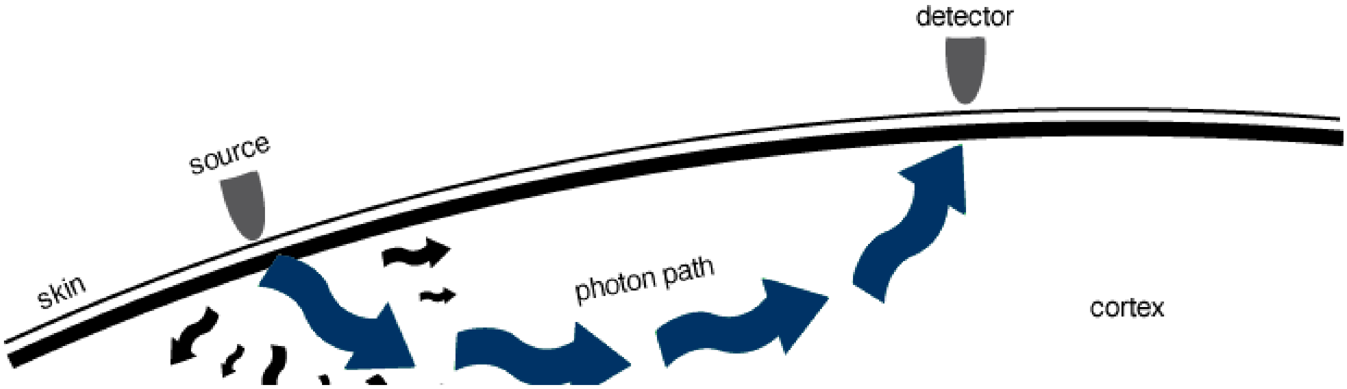

Figure 1: Photons follow a banana-shaped path through the skin, skull and cerebral cortex and are measured by a photo-receptor typically positioned between 1.5 to 4 cm from the optode.

the near infrared range of light (between 700-900nm), biological tissue is relatively transparent, whereas oxy- and deoxy-hemoglobin are strong absorbers of light. Estimates of the ratio of oxygenated to deoxygenated hemoglobin can thus be obtained through the coupling of infrared light emission and detection. In general, a light source emits photons which are carried through a fiber optic bundle to a sensor placed directly the skin. From there, the photons follow a banana-shaped path through the skin, skull and cerebral cortex to be caught by a photo-detector positioned several centimeters away from the source (see Figure 1). Since oxy- and deoxy-hemoglobin are both chromophores that absorb light in the near-infrared range, regional changes in hemoglobin concentration can be calculated based on wavelength-dependent changes in light attenuation ([6]).

Sensors (containing one or more light sources and detector) are positioned to target specific regions of the brain via multiple channels (light source-frequency pairings; see [24] for more details). These sensors are attached by various means including caps, headbands, and mechanical supports. While most sensors are tethered by cables that extend back to the NIRS instrument (see 2), there are also wireless, portable versions of NIRS devices in development [6]. For more detailed reviews of hemodynamics and NIRS instrumentation, see [12, 20, 24, 34].

3. APPLICATIONS AND UTILITY

There are a number of comprehensive reviews of the utility of NIRS in general (e.g., [5, 20, 34, 46]). However, as the literature on NIRS is dispersed across many publication outlets in the HRI, HCI, neuroimaging, and brain-computer interface (BCI) communities, we review here the cognitive and affective states measureable with NIRS, the applicability of NIRS in HRI, and the relevant considerations for its deployment.

3.1 Evaluation of Human-Robot Interaction

Within the field of human-robot interaction, two distinct usage patterns have emerged with respect to NIRS: (1) as a measurement tool for evaluating interactions between humans and robots (e.g., [41]), and (2) as a passive input channel that alters a robot's behavior based on a person's cognitive state (e.g., [26]. In both cases, a person performs a task (e.g., interacts with a robot) while wearing the NIRS probes. However, the principle difference between the two paradigms is what happens to that signal once it reaches the computer for processing. As an evaluation tool, NIRS enables offline analyses and inferences about cortical activity. Data recorded during a task are preserved and subsequently inspected or visualized by the experimenter post-hoc. A number of NIRS-based studies have been conducted in this regard to investigate people's cognitive reponses towards robots [23, 28, 36, 39, 41, 44]. In this work, NIRS is typically used as an additional measurement or as an objective measurement of people's genuine perceptions while they interact with or observe a particular robot or particular behavior of a robot. To-date, its usage in this manner has shown relevant findings regarding attention, engagement, empathy, and discomfort in human-robot interactions. These cognitive and affective states are of clear relevance to the field of HRI, as a robot designed for interactivity with a human agent need be approachable and interesting for there to be successful (and sustained) interaction.

Mental workload. The discrimination of mental workload states (i.e., low versus high workload) are the predominate focus of using NIRS in human-computer interaction (e.g., [17, 35]. These states are typically induced using arithmetic-based or memory-based tasks such as mental math or the n-back task [13] and are generally measured via the prefrontal cortex (an area easily accessible to NIRS via the forehead, as there is typically no hair to influence the measurements). However, while adaptivity of robot behavior based on a person's workload level has suggested promise for improving human-robot interactions [26, 37], the discernment of different states (e.g., low versus high workload) using NIRS is relatively unstable across investigations (e.g., [42, 7, 13, 18]). Specifically, classification accuracies range from barely above chance (e.g., [42]) to high-80s (e.g., [14]) with no clear relation to the device or signal processing methods used. Moreover, while some work suggests that NIRS outperforms EEG on detecting mental workload (e.g., [18]), others suggest the reverse (e.g., [7]).

Positive affect. Compared to the volume of workload-based investigations, there have been only a few NIRS-based studies targeting postiviely-charged states such as attention, en-

Figure 2: A subject fitted with a two-probe NIRS instrument. The sensors are positioned along the forehead to measure the prefrontal cortex and secured using a spandex cap. Data is transmitted via fiber-optic cables to the computer (back, left) that is connected to the oximeter (right). Image reproduced from [40] with permission.

gagement, and empathy [23, 28, 36, 39]. Specifically, Kawaguchi and colleagues [23, 36] investigated the effectiveness of Paro, the robotic seal, for robot-based therapy. Using NIRS, they observed activation in both sides of the Sylvian fissure while participants interacted with Paro, which indicated participants recognized and responded to Paro's emotional gesture expression. Similarly, Nozawa and Kondo (2010) used NIRS to evaluate human interactions with three different adaptive agents, finding that the degree of intrinsic motivation of an agent increased participants' brain activity in the the dorsolateral cortex – an area associated with both the control and sustanence of attention. More recently, we investigated the degree to which people 'empathized' with robots in moral utilitarian dilemmas and the relative effects of agency, moral value, and incentive on both participantsâĂŹ behavior and prefrontal hemodynamics in utilitarian decision-making. Consistent with a recent fMRI investigation of empathy towards robots (e.g., [33]), we initially observed that seeing a humanoid robot (the Aldebaran Nao) in danger elicited substantially more activity in the prefrontal cortex than other inanimate objects, but significantly less activity compared to seeing a human in danger [39]. However, in a set of follow-up studies, the agency effects on prefrontal activity (the differences in activity levels between seeing a robot versus human in danger) did not persist [43]. As participants only viewed images of the humanoid robot agents, however, it is possible that, by increasing the presence of the robot agents, such agency effects may become more apparent from the NIRS data (as was the case in [41]).

Negative affect. There are also a growing number of NIRS-based studies showing the measurability of negatively-charged emotional states such as negative mood [3, 2], low versus high arousal/valence [16], emotion regulation [47], and frustration [19]. While still a relatively unexplored application area, the results here are highly consistent across the various efforts to detect negative affect and moreover, across a diverse set of contexts (e.g., threat, working memory tasks, moral dilemma, human-robot interactions) in which that has been achieved. Regarding human-robot interactivity in particular, we have used NIRS to investigate discomfort in various interactions and contexts with a number of robots [41, 44]. Specifically, in a recent mixed-methods study, we employed both survey-based subjective and brain-based objective measures to investigate the effects of several situational factors on the perceived effectiveness of the human-robot interaction. The aim of our study was to measure the effects of robot communication strategies such as direct vs. indirect speech in advice-giving contexts. We manipulated three additional factors – robot appearance, interaction modality, and interaction distance – as they have been shown to modulate the effectiveness of human-robot interactions. Here we observed significant differences in prefrontal activity in response to two robots (the Xitome Designs' MDS and Willow Garage's PR2) based on whether the participant was co-located or remotely located with the given robot. Specifically, we found that when co-located with the very human-like MDS, significant brain activity was elicited and moreover, participants reported severely decreased preferences for further interaction with the MDS. Whereas, participants showed no significant reaction to co-located interactions with the PR2. Moreover, participants showed no significant reactions (neither in brain activity nor in subjective preferences) towards the MDS when there was substantial distance in the interaction (i.e., interaction via a Skype call or observation of a video recording). The survey-based and brain-based measures hence, in combination, suggest there may be emotion-regulatory mechanisms evoked when directly interacting with a co-located, humanoid robot. Whereas, in a removed context such as that of observing video of the two – much like viewing a movie – the fear or anxiety elicited by the MDS' eerie appearance may have been reduced or non-present.

3.2 Passive Adaptivity of Robot Behavior

As a passive input to adapt various robot behaviors, the NIRS measurements are analyzed in realtime and interpreted by a system trained to identify a state of the NIRS signal (e.g., high workload). When a particular state of the user is identified, a control signal is transmitted to the robot to change its behavior accordingly. While NIRS-based robot adaptivity has been suggested to improve the efficacy of human-robot interactions, the results and their reliability are still limited. To-date, there have been just a couple demonstrations of such systems [26, 37].

3.2.1 Exemplars

Matsuyma and colleagues created a simple, proof-of-concept system based on the detection of workload-related activity. Their study was a preliminary (and first) attempt at using passive monitoring of a person's cognitive state (using NIRS) to adapt a robot's behavior. There they measured participants' prefrontal cortex while they solved arithmetic problems. When an increase in oxygenated hemoglobin corresponding to the participant actively working on a arithmetic problem was detected, the robot received a motor command to raise its arm. However, this study exposed a particular shortcoming of NIRS that is an obstacle for its effectiveness in more realistic scenarios, namely that of onset detection latency [6]. Specifically, using their approach to workload monitoring, the time between a participant beginning the arithmetic problem and the transmission of the motor control signal ranged from just few seconds to over fifteen seconds [6, 26]. As task-related hemodynamic changes in oxygenated hemoglobin occur over several seconds [8], this delay was (and is) somewhat unavoidable due to the inherent hemodynamics; however, recent work has demonstrated vast reductions in temporal delays to onset detection [10] which suggests further improvement may be possible.

Similar to [26], we previously participated in the development of a passive NIRS-BCI aimed at adapting a robot's behavior based on a person's detected multitasking state [37]. A two-probe NIRS instrument (with four sources per probe) was used to image participants' prefrontal cortex, while they worked with two simulated robots on a human-robot team task. Here we designed a naive SVM classification model based on gross temporal dynamics and trained using data collected while participants performed a variant of the n-back task. In the human-robot team task, we hypothesized that adapting the level of a robot's autonomy would lead to better task performance and better perceptions of teamwork. Thus, while participants performed the team task, classifications of their mental workload dynamically adapted the autonomy of one of the robots according to the participant's multitasking state. An initial evaluation [37] showed successful task completion was significantly moderated by adaptivity: the dynamic adaptivity of the robot's autonomy improved task performance (82% of participants successfully completed the team task versus a baseline performance rate of 45%). This system was thus a substantial extension of [26], as it was the first NIRS-BCI to demonstrate effective improvements on a *realistic* task. However, in a recent series of reinvestigations [42] of this system's classification performance, the average classification accuracy on an alternative dataset (of mental arithmetic) was only 54.5% (SD=14.3%) suggesting limited generalizability of the system's signal processing. Additionally, this NIRS-based classification model was found effective (statistically better than chance) for only 10/40 participants in the alternative dataset [42], which suggested limited utility for a more realistic population sample (i.e., when N=40 vs. N=3 in the initial evaluation).

3.2.2 Utility

Despite the limited success of passive NIRS-based robot adaptivity thus far, there is evidence supporting its utility. For instance, although our more recent findings [42] suggest low replicability and extensibility of [37], the subjective reports in the original experiment indicate the dynamic autonomy improved perceptions of the robot teammates. Specifically, the adaptively autonomous robot was found to be more helpful and cooperative than its non-adaptive counterpart. Thus, there may be still some utility to the NIRS-based adaptivity (despite the confounds), though perhaps not measureable at the scale of task completion rates.

Additionally, NIRS may be particularly useful for evaluation and/or passive monitoring of a person's aversion or discomfort with a robot. Via a qualitative evaluation ([44]) of participants' prefrontal activity during human-robot interactions, we found persistent trends across participants based on situational factors of the interaction and similar changes in activity as a function of the duration of exposure to a robot (see Figure 3). Specifically, we analyzed prefrontal cortical activity over the course of two-minute (semi) free-form interactions. Additionally, we evaluated the changes in activity over repeat interactions and whether the effects observed were large enough to persist at an individual subject-level. The findings were consistent with our prior results [41], suggesting, in combination with subjective measures of preference, that PFC hemodynamics reflect a person's aversion to a robot and are moderated by the settings of the interaction and the human-likeness of the robot interlocutor. But our results also suggest that *participants' responses fluctuate over the course of a task, and may diminish with prolonged exposure to or interaction with a given robot.* This indicates that passive monitoring of a person's cognitive state would be particularly useful, as it would allow the robot to react or modify its behavior in order to reduce a person's aversion. However, given the high variability and noise in the NIRS data, it is unclear whether these effects are reliably detectable at an individual level. Thus the use of NIRS as a feasible realtime measurement for adapting robot behavior based on a person's affective state may be still premature without more controlled investigations to better understand individual variations in the signal.

3.3 Considerations

In the previous section, we discussed more prototypical examples of NIRS in HRI, which were intended to serve as proof-of-concept studies that NIRS can be utilized both as a component of an interface device and as an evaluation tool in HRI. However, many applications have encountered significant challenges that are not yet, in most cases, sufficiently addressed. The goal of this section is to discuss the most pressing considerations for the wide use of NIRS in HRI.

In the standard, *offline* approaches to signal processing of NIRS data, the signals are short (3 to 60 seconds) and heavily filtered post-hoc (with roughly the following measures) – *detrending* (removal of low frequency signal artifacts and drift), *smoothing* (removal of systemic artifacts such as cardiac pulsations, respiration, and Mayer waves), *motion correction* (reduction of motion artifacts), and *data reduction* (removal of noisy or corrupt trials; averaging over repetitions of a task and/or truncation of the signal to reduce temporal variation; using summary statistics, e.g. area-under-the-curve, percent signal change to represent the overall hemodynamic response) – see [5, 6, 9, 20, 25, 29, 34, 46]. Such processing can result in dramatic reductions of signal noise, however, in online, passive settings, signal processing faces

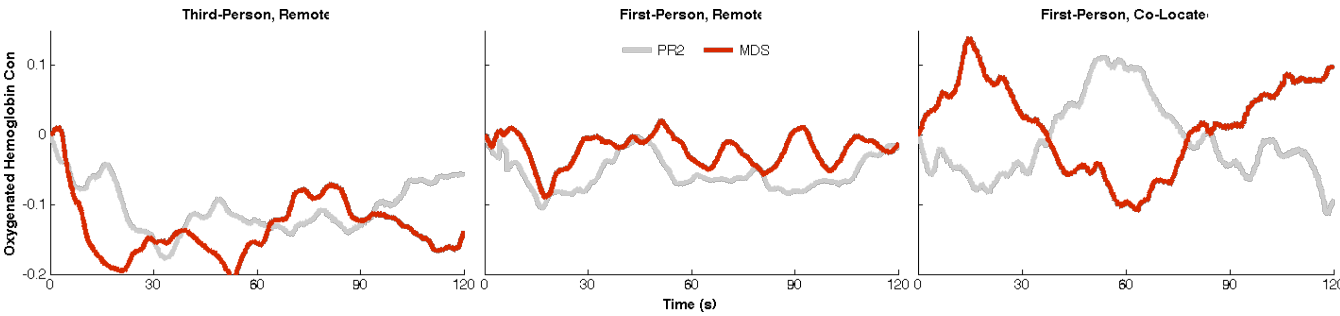

Figure 3: Average (N=15) hemodynamic activity sampled from the left prefrontal cortex in response to interactions with the MDS (red) and PR2 (gray) robots, mediated by situational context [41]. Left: third-person observation of video-recorded interactions. Center: remote human-robot interaction (occured over Skype). Right: co-located interaction. Image reproduced from [44] with permission.

substantial limitations [6, 35, 40].

Participant mobility. Motion artifacts degrade and transform the NIRS signal. These artifacts can be caused by various sources including movement of the sensors on the skin, facial expressions, head orientation, and more [6, 27, 31]. Attempts at combating these effects include chin rests and mechanical supports [8] (which are not particularly reasonable in realistic HRI settings) as well as specific signal processing filters. There are, however, several recent and promising proposals for real-time motion artifact correction in natural environments [1, 5, 22].

Task-unrelated activity. Separating task-related from unrelated cortical activity and signal noise is also difficult. Task-unrelated activity such as resting-state fluctuations [20, 21] or whole brain activity [26] can degrade the signal quality. Additionally, the prefrontal cortex, for example, is a "bustling metropolis of executive functions" [6], hence it is unlikely that it is possible to discriminate among the various parallel ongoing processes in this area. Moreover, NIRS is sensitive to systemic physiological artifacts such as Mayer waves and those from respiration, blood pressure, and cardiac pulsation (e.g., [9, 11, 27, 49]). Hence, all of these factors confound inferences about characteristics of the signal, specifically characteristics attributable to the task (i.e. not present during rest). This sensitivity can obfuscate task-related activity and potentially lead to incorrect interpretations of the signal.

Hemodynamic response. It is unclear how reliably and quickly any such distinction could be made since even the onset of vascular change lags 1-2 seconds behind the neural activation that caused it (e.g., [10]). Realtime systems which intend to dynamically act on rapid cognitive changes (e.g., [26, 37]) must address this inherent trait since it otherwise limits the applicability of NIRS to slow-changing (on the order of several seconds) cognitive phenomena. Another potential limitation is the way the hemodynamic response changes over longer periods of time. Most previous NIRS research in HCI and HRI has been limited to task-onset detection in 10-45 second windows. The signal over longer time spans has gone largely uninvestigated.

Probe placement and reproducibility. Sensors must be attached securely in order to avoid creating motion artifacts that degrade the NIRS signal. One research group reported that the method of connecting the optodes to the subject's head actually had the greatest bearing on system performance because of its substantial effect on signal quality [8]. The quality of the NIRS signal is additionally affected by dark skin pigmentation [48] and hair caught between the sensor and scalp [8]. While a first step toward imaging specific regions of the brain with NIRS is to use a probe-placement framework like the international 10-20 system for EEG [13], there is no guarantee precisely which regions are measured [20, 30]. These frameworkis rely on anatomical landmarks for placing sensors and so are inherently inexact across participants. There is also variation in precise function-location mappings between partici- pants. Furthermore, re-positioning the probe set on an individual is prone to error and even millime- ter movements of the probe set can lead to centimeter shifts in whole channel position [30].

Hardware and Environment Additional complications arise due to NIRS hardware and environmental conditions. For one, because probes need to be placed on the subject's head and connected to a processing device (either through wires or through a wireless tether), there are range and mobility limitations imposed on the kinds of interactions that can be performed. This is especially limiting of proximate interactions (e.g., situated social robotics experiments). Moreover, in some cases, the subject's hair may entirely prevent appropriate probe placements and thus prevent NIRS from being used (as shaving the head is usually not an option). Ambient light and changes of lighting conditions in the environment also poses potential problems as they can alter the signal-to-noise ratio. In most cases this can be addressed with thick opaque covers (e.g., headbands, helmets, etc.) that protect the probes.

4. DISCUSSION

The aim of this paper was to provide (1) an overview of what we can do with NIRS in HRI and (2) a discussion of the relevant considerations for its usage in more realistic settings. We described two ways of utilizing NIRS: as a tool for evaluation of interactions and as a passive input for adapting robot behavior. The prototypical application of NIRS as an evaluation tool is as an offline post-hoc analysis of a signal recorded during some interaction. Whereas, the usage of NIRS as a passive BCI has emerged as realtime monitoring of a person's cognitive state, and with its emergence, a number of considerations have followed. The exploration of NIRS as a tool in HRI research and the evaluation of its potential have already begun, as demonstrated by the above exemplary studies. However, given the many challenges to NIRS that remain to be addressed, it seems premature to consider NIRS ready for more realistic HRI applications. Hence here we summarize best practices, research avenues, and paradigms.

The following list suggests paradigm-independent best research practices for any application of NIRS to HRI (and related fields):

- Probe placement: effective, consistent placement that ensures no hair caught between the probes and the forehead, no movement of the probes, and constant contact with scalp. Application should also adhere to a standard framework for probe placement (e.g., 10-20 system for EEG).
- Limited subject movement and motion: support for the subject's head as well as verbal instructions to minimize movement. Moreover, facial expressions should be recorded for post-hoc filtering as contortion of the facial muscles (e.g., frowning, smiling, talking, etc.) can produce signal artifacts (e.g., [16]).
- Appropriate signal processing: filtering of regular artifacts from the NIRS signal, including those caused by Mayer waves, cardiac and respiration patterns, and subject mobility (including facial artifacts).
- Statistical inference: use of adequate control conditions (e.g., samples should be recorded of the participant frowning if the task involves negatively-valenced stimuli that might induce frowning).

The above guidelines are important for optimizing the signal-to-noise ratio, reliability of data, and legitimacy of inferences about those data.

While the utility of NIRS for robot adaptivity based on mental workload states remains to be determined, there are some domains where NIRS can already be rather straightforwardly deployed. Candidates include any type of HRI domain where human motion can be significantly restricted and environmental conditions can be controlled (such as in the case of a human operator situated at a computer screen in an indoor environment) and where monitoring the human can lead to some performance improvement (either by just collecting data about the human for post-hoc processing or by directly adapting robot behavior during task performance). This includes mixed-initiative human-robot teams where robots would benefit from being better-informed about the physiological state of their human teammates allowing them to make better and more informed decisions (e.g., regarding task allocations, behavior adaptations, and verbal communication). Such tasks include coordinated search and rescue missions, remote deep-sea and space exploration, military exercises, unmanned aerial vehicle (UAV) operation, and many others [15] for a complete survey of HRI problem domains). In particular, NIRS might be able to help address one of the important open problems in HRI: how to provide interfaces and mechanisms to allow a single operator to control multiple robots (e.g., human cognitive workload measured by NIRS during multi-robot interactions could be used be used to dynamically adjust the autonomy of these robots; for early efforts, see [38]). Currently, however, NIRS might find wider and more robust use as a component in evaluative studies. Evaluation studies do not necessitate realtime processing of the NIRS signal and thus the signal can be processed offline, after the interaction, which does not impose as many challenges.

The properties of the hemodynamic response should also influence application avenues for NIRS in HRI. Since the onset of the vascular response lags behind the neural stimulus by several seconds, NIRS is not suited for rapid (subsecond time scale) reaction to cognitive events. Instead, NIRS might be better suited to detection of minute-by-minute state changes over sustained periods (i.e., slower-changing cognitive phenomena). However, there are still many obstacles to overcome even for this kind of application, and it is necessary to integrate knowledge of regional hemodynamics over time to design appropriate applications.

5. CONCLUSIONS

Functional near-infrared spectroscopy is a promising new technology for the HRI community and there is a wide range of areas to which brain-imaging seems to be applicable within and outside HRI. NIRS has already been used for HRI research (such as robot-assisted therapy, engagement, attention, limb-control and rehabilitation, intention, and others), both as a brain-robot interface to control robot behavior and as an evaluation tool for ascertaining the effectiveness of human-robot interactions. However, several significant challenges remain to be addressed before NIRS can become a more widely useful, practical tool for HRI research. These challenges include context-dependent hardware concerns, signal inference, interface design, robust signal processing, properties of the hemodynamic response, effective probe placement, and developing statistical analysis and inference tools. While many of the difficulties with NIRS have been recognized, there is currently a dearth of follow-up studies to address them. It is thus our hope that this survey will prompt researchers to actively engage in NIRS-related HRI research that can help overcome the current challenges to make NIRS a genuinely and consistently useful tool to the HRI community.

6. REFERENCES

[1] A. Abdelnour and T. Huppert. Real-time imaging of human brain function by near-infrared spectroscopy using an adaptive general linear model. *Neuroimage*, 46(1):133–143, 2009.

[2] R. Aoki, H. Sato, T. Katura, R. Matsuda, and H. Koizumi. Correlation between prefrontal cortex activity during working memory tasks and natural mood independent of personality effects: an optical topography study. *Psychiatry Research: Neuroimaging*, 212(1):79–87, 2013.

[3] R. Aoki, H. Sato, T. Katura, K. Utsugi, H. Koizumi, R. Matsuda, and A. Maki. Relationship of negative

mood with prefrontal cortex activity during working memory tasks: an optical topography study. *Neuroscience Research*, 70(2):189–196, 2011.

[4] W. A. Bainbridge, S. Nozawa, R. Ueda, K. Okada, and M. Inaba. A methodological outline and utility assessment of sensor-based biosignal measurement in human-robot interaction. *International Journal of Social Robotics*, 4, 2012.

[5] S. Brigadoi, L. Ceccherini, S. Cutini, F. Scarpa, P. Scatturin, J. Selb, L. Gagnon, D. Boas, and R. Cooper. Motion artifacts in functional near-infrared spectroscopy: a comparison of motion correction techniques applied to real cognitive data. *NeuroImage*, 85(1):181–191, 2014.

[6] C. Canning and M. Scheutz. Function near-infrared spectroscopy in human-robot interaction. *Journal of Human-Robot Interaction*, 2:62–84, 2013.

[7] E. Coffey, A.-M. Brouwer, and J. van Erp. Measuring workload using a combination of electroencephalography and near infrared spectroscopy. In *HFES*, 2012.

[8] S. Coyle, T. Ward, and C. Markham. Brain-computer interface using a simplified functional near-infrared spectroscopy system. *Journal of Neural Engineering*, 4(3):219, 2007.

[9] X. Cui, S. Bray, and A. Reiss. Functional near infrared spectroscopy (NIRS) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *NeuroImage*, 49(4):3039–3046, 2010.

[10] X. Cui, S. Bray, and A. Reiss. Speeded near infrared spectroscopy (NIRS) response detection. *PLoS ONE*, 5(11):e15474, 2010.

[11] T. Fekete, D. Rubin, J. Carlson, and L. Mujica-Parodi. The NIRS analysis package: Noise reduction and statistical inference. *PloS one*, 6(9):e24322, 2011.

[12] M. Ferrari and V. Quaresima. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*, 63(2):921–935, 2012.

[13] J. Frey, C. MÃijhl, F. Lotte, and M. Hachet. Review of the use of electroencephalography as an evaluation method for human-computer interaction. In *International Conference on Physiological Computing Systems (PhyCS)*, 2014.

[14] A. Girouard, E. Solovey, and R. Jacob. Designing a passive brain computer interface using real time classification of functional near-infrared spectroscopy. *International Journal of Autonomous and Adaptive Communications Systems*, 6(1):26–44, 2013.

[15] M. A. Goodrich and A. C. Schultz. Human-robot interaction: A survey. *Foundations and Trends in Human-Computer Interaction*, 1(3):203–275, 2007.

[16] D. Heger, R. Mutter, C. Herff, F. Putze, and T. Schultz. Continuous recognition of affective states by functional near infrared spectroscopy signals. In *Affective Computing and Intelligent Interaction (ACII)*, 2013.

[17] C. Herff, D. Heger, O. Fortmann, J. Hennrich, F. Putze, and T. Schultz. Mental workload during n-back task - quantified in the prefrontal cortex using fNIRS. *Frontiers in Hum. Neur.*, 7(935), 2013.

[18] L. Hirshfield, K. Chauncey, R. Gulotta, A. Girouard, E. Solovey, R. Jacob, A. Sassaroli, and S. Fantini. Combining electroencephalograph and functional near infrared spectroscopy to explore users' mental workload. In *HCII*, pages 239–247, 2009.

[19] L. Hirshfield, R. Gulotta, S. Hirshfield, S. Hincks, M. Russell, R. Ward, T. Williams, and R. Jacob. This is your brain on interfaces: enhancing usability testing with functional near-infrared spectroscopy. In *CHI*, pages 373–382, 2011.

[20] Y. Hoshi. Towards the next generation of near-infrared spectroscopy. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1955):4425–4439, 2011.

[21] X. Hu, K. Hong, and S. Ge. Reduction of trial-to-trial variability in functional near-infrared spectroscopy signals by accounting for resting-state functional connectivity. *Journal of Biomedical Optics*, 18(1), 2013.

[22] M. Izzetoglu, P. Chitrapu, S. Bunce, and B. Onaral. Motion artifact cancellation in nir spectroscopy using discrete kalman filtering. *Biomedical engineering online*, 9:16, 2010.

[23] Y. Kawaguchi, K. Wada, M. Okamoto, T. Tsujii, T. Shibata, and K. Sakatani. Investigation of brain activity after interaction with seal robot measured by fnirs. In *Robot and Human Interactive Communication,*, pages 308–313, 2011.

[24] S. Lloyd-Fox, A. Blasi, and C. Elwell. Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neuroscience and biobehavioral reviews*, 34(3):269–84, 2010.

[25] N. Logothetis. What we can do and what we cannot do with fmri. *Nature*, 453:869–878, 2008.

[26] H. Matsuyama, H. Asama, and M. Otake. Design of differential near-infrared spectroscopy based brain machine interface. In *Robot and Human Interactive Communication*, pages 775–780, 2009.

[27] F. Matthews and B. Pearlmutter. Hemodynamics for brain-computer interfaces. *Signal Processing*, pages 87–94, 2008.

[28] T. Nozawa. Autonomous adaptive agent with intrinsic motivation for sustainable hai. *Journal of Intelligent Learning Systems and Applications*, 2(4):167–178, 2010.

[29] F. Oriheula-Espina, D. Leff, D. James, A. Darzi, and G. Yang. Quality control and assurance in functional near infrared spectroscopy (fNIRS) experimentation. *Physics in Medicine and Biology*, 55(13), 2010.

[30] M. Plichta, M. Herrmann, C. Baehne, A. Ehlis, M. Richter, P. Pauli, and A. Fallgatter. Event-related functional near-infrared spectroscopy (fNIRS) based on craniocerebral correlations: reproducibility of activation? *Human Brain Mapping*, 28(8):733–41, 2007.

[31] F. C. Robertson, T. S. Douglas, and E. M. Meintjes. Motion artifact removal for functional near infrared spectroscopy: a comparison of methods. *IEEE transactions on bio-medical engineering*, 57(6):1377–87, 2010.

[32] A. M. Rosenthal von der Putten, N. C. Kramer, L. Hoffmann, S. Sobieraj, and S. C. Eimler. An

experimental study on emotional reactions towards a robot. *International Journal of Social Robotics*, 5:17–34, 2013.

[33] A. M. Rosenthal von der Putten, F. P. Schulte, S. C. Eimler, L. Hoffmann, S. Sobieraj, S. Maderwald, N. C. Kramer, and M. Brand. Neural correlates of empathy towards robots. In *Human-Robot Interaction (HRI)*, pages 215–216, 2013.

[34] F. Scholkmann, S. Kleiser, A. Metz, R. Zimmermann, J. Pavi, U. Wolf, and M. Wolf. A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology. *NeuroImage*, 85(1):6–27, 2014.

[35] L. Schudlo and T. Chau. Dynamic topographical pattern classification of multichannel prefrontal NIRS signals: online differentiation of mental arithmetic and rest. *Journal of Neural Engineering*, 11(1), 2014.

[36] T. Shibata. Therapeutic seal robot as biofeedback medical device: qualitative and quantitative evaluations of robot therapy in dementia care. *Proceedings of the IEEE*, 100(8):2527–2538, 2012.

[37] E. Solovey, P. Schermerhorn, M. Scheutz, A. Sassaroli, S. Fantini, and R. Jacob. Brainput: enhancing interactive systems with streaming fNIRS brain input. In *CHI*, pages 2193–2202, 2012.

[38] E. T. Solovey. *Real-Time fNIRS Brain Input for Enhancing Interactive Systems*. PhD thesis, Tufts University, 2012.

[39] M. Strait, G. Briggs, and M. Scheutz. Some correlates of agency ascription and emotional value and their effects on decision- making. In *Affective Computing and Intelligent Interaction (ACII)*, 2013.

[40] M. Strait, C. Canning, and M. Scheutz. Limitations of NIRS-based BCI for realistic applications in human-computer interaction. In *BCI Meeting*, pages 6–7, 2013.

[41] M. Strait, C. Canning, and M. Scheutz. Let me tell you! investigating the effects of robot communication strategies in advice- giving situations based on robot appearance, interaction modality and distance. In *Human-Robot Interaction (HRI)*, 2014.

[42] M. Strait, C. Canning, and M. Scheutz. Reliability of nirs-based bcis: a placebo-controlled replication and reanalysis of brainput. In *Alt.CHI (in review)*, 2014.

[43] M. Strait and M. Scheutz. Using functional near infrared spectroscopy to measure moral decision-making: Effects of agency, emotional value, and monetary incentive! In *Brain-Computer Interfaces*, 2014.

[44] M. Strait and M. Scheutz. Using near infrared spectroscopy to index temporal changes in affect in realistic human- robot interactions. In *Physiological Computing Systems (PhyCS), Special Session on Recognition of Affect Signals from Physiological Data for Social Robots*, 2014.

[45] D. Szafir and B. Mutlu. Pay attention! designing adaptive agents that monitor and improve user engagement. In *Conference on Human Factors in Computing Systems (CHI)*, pages 11–20, 2012.

[46] S. Tak and J. Ye. Statistical analysis of fNIRS data: a comprehensive review. *Neuroimage*, 85(1):92–103, 2014.

[47] S. Tupak, T. Dresler, A. Guhn, A. Ehlis, A. Fallgatter, P. Pauli, and M. Herrmann. Implicit emotion regulation in the presence of threat: neural and autonomic correlates. *Neuroimage*, 85(1):372–379, 2014.

[48] E. Wassenaar and J. Van den Brand. Reliability of near-infrared spectroscopy in people with dark skin pigmentation. *Journal of Clinical Monitoring and Computing*, 19(3):195–199, 2005.

[49] J. Ye, S. Tak, K. Jang, J. Jung, and J. Jang. NIRS-SPM: statistical parametric mapping for near-infrared spectroscopy. *Neuroimage*, 44(2):428–447, 2009.

[50] T. Zander and S. Jatzev. Detecting affective covert user states with passive brain-computer interfaces. In *Affective Computing and Intelligent Interaction (ACII)*, pages 1–9, 2009.