**Project Description**

1. **Introduction**

Discovery, learning and innovation in science and engineering rely heavily on collaborations, often among research teams and resources distributed across institutional and geographic boundaries. The recent increase in multi-disciplinary and multi-institutional research and education has reinforced the need for transformational technological solutions that enrich collaborative interactions across distances. NSF’s mission for CyberInfrastructure (CI) aims to deliver to the science and engineering communities access to world-class CI tools and services that connect researchers across geographical boundaries through virtual organizations (VO) [1]. While virtual organizations and collaboratories are expected to be *highly interactive* and *widely accessible* to the participating entities, existing efforts have largely focused on the development of CI tools for data storage, analysis, retrieval and exchange in remote collaborations with limited efforts to enhance the interaction capabilities of technological frameworks. Since existing VOs lack the rich social experience that is present in face-to-face interactions, scientific communities still rely on face-to-face meetings among collaborators to advance research and pedagogy. This continued need for travel and face-to-face interactions has a direct impact on the carbon foot print. There is therefore a strong need to guide the development of new technologies that enrich remote research interactions in an ecologically sustainable manner.

It is widely accepted that collaborative efforts improve the quality of research [2]. In particular, collaborations among people of diverse backgrounds, expertise and personalities show enhanced creativity and decision-making relative to more homogeneous teams [3]. Effective collaboration requires frequent communication in different formats (e.g., group conference, lecture, one-on-one meeting, etc.), among different kinds of participants (faculty, researchers, students) located in different places. Thus, telecommunication technologies have become integral to collaborative teamwork. Remote interactions in inter-disciplinary collaborations assume different forms: (i) limited-bandwidth, limited-modality communications such as text messaging, emails and instant messaging, (ii) collaborative editing solutions for document management (i.e, Google docs); (iii) scheduled one-on-one audio conversations; (iv) one-to-one desktop conferencing/few-to-few video conferencing solutions; (v) immersive group videoconferencing solutions, such as telepresence; and (vi) avatar mediated virtual presence solutions.

While these technologies have served the needs of scientific communities in recent decades, “virtual collaboration” has severe limitations. It is difficult to develop group norms, trust, and team identification in virtual teams, relative to face-to-face [4,5,6]. Managing conflict is especially difficult in virtual teams, as initial misunderstandings may go undetected and opportunities for resolution are limited [7]. Confusion is more likely to occur, and less likely to be resolved, because nonverbal feedback cues that regulate verbal information exchange are blocked or distorted [8]. In particular, decision-making processes are impaired by virtual communication. All of these problems interfere with productivity in virtual teams. Multiple reviews of the empirical literature have concluded that in comparison to face-to-face teams, computer-mediated teams take more time to make decisions, exchange lesser information, and report lower satisfaction with the process [9].

There is therefore an impending need for *human-centric* communication technologies, where technology adapts to the needs of participants in order to enrich the experience of interaction. In this proposal, we seek to bridge the divide between existing communication frameworks for remote interactions and the rich capabilities present in face-to-face interactions. We propose a human-centered computational framework based on principles of socioemotional psychology and ecological design that enriches research advancements through collaborative technologies (e-REACH).

1. **Rationale and Significance**

Discovery, innovation and learning are advanced through spontaneous and unstructured interactions, where participants interact in a natural setting promoting creativity and free expression. However, existing technology solutions that are closest to simulating face-to-face interactions (such as telepresence environments) are limited to scheduled, highly structured, and formal interactions. Furthermore, all current telecommunication technologies that support virtual collaborations suffer from **emotional impoverishment**. Emotions, and emotional communication, are deeply and inextricably woven into the fabric of human social interaction [10]. Analyses of the limitations of existing technologies for virtual collaboration often demonstrate their failure to communicate nonverbal cues of crucial psychological states such as attention, understanding, and emotion [11]. We propose that these limitations are due to a fundamental misunderstanding of the process by which these socioemotional states (abbreviated to “emotion”) are communicated. In Figure 1, the sender feels and displays annoyance, the receiver perceives this annoyance, and the receiver feels and displays embarrassment and a willingness to reconcile. Although extreme emotions can sometimes interfere with healthy communication, socioemotional communication more often regulates face-to-face interaction in a dynamic, natural way.

This model may, on the surface, seem absurdly simple. However, certain features of this process are widely misunderstood, in ways that limit the development of emotionally rich communication technologies: (1) Displays of attention, emotion and the like are subtle and complex, typically involving interacting signals among several non-verbal “channels” rather than a strong signal in one channel [12,13,14,15]. (2) Perception of another’s socioemotional display is typically not a logical, deliberative process. These perceptions are instinctive, requiring little conscious cognitive effort, and we are often unaware of the nonverbal cues driving our perception of others’ states [15,16,17]. (3) Appropriate responses to a sender’s emotions are best facilitated not by *knowing* that the sender is feeling an emotion, but by *feeling* an appropriate, corresponding emotion in response [10].

Figure 1: Communication of socioemotional states between interaction participants

Even the most sophisticated existing communication technologies interfere with this process. Transmission of the nonverbal cues is extremely limited and often grossly distorted. Text-based technologies such as email and messaging completely block facial expression, tone of voice, posture, and gaze. Voice-based technologies often accommodate only one speaker at a time and time delays and altered tone quality result in stilted conversation. Even video conferencing limits natural movement, distorts the time course of facial expressions and acoustic properties of the voice, and requires a tradeoff between visual precision and breadth. As a result, many of the nonverbal cues critical to socioemotional communication simply fail to reach the receiver. In recent years, a few interactive technologies have attempted to identify the socioemotional state of a sender, and to communicate this information to a receiver explicitly through the use of media rich means like virtual agents. Unfortunately, productive virtual collaborations cannot be achieved by substituting face-to-face interactions with virtual visual avatars since this violates another principle of the Socioemotional Communication Model – perceptions of another’s state should be *instinctive* and *“natural”*, thereby evoking a corresponding emotional response [18,19]. The manner in which cutting edge technologies deliver information about a sender’s socioemotional state is typically cognitively demanding, placing added strain on the receiver’s attention, and having little emotional impact. The immersive telepresence solution, addresses these issues sufficiently and provides a seamless audio-visual platform for interaction partners in geographically diverse locations. However, the costs associated with the infrastructure for installation, very high bandwidth channels, the loss of flexibility and structured nature of interaction limits the adoption of this technology to corporate environments. In addition, we believe that it misses a significant opportunity to augment the interaction experience through other multi-modal cues (i.e. touch, etc.).

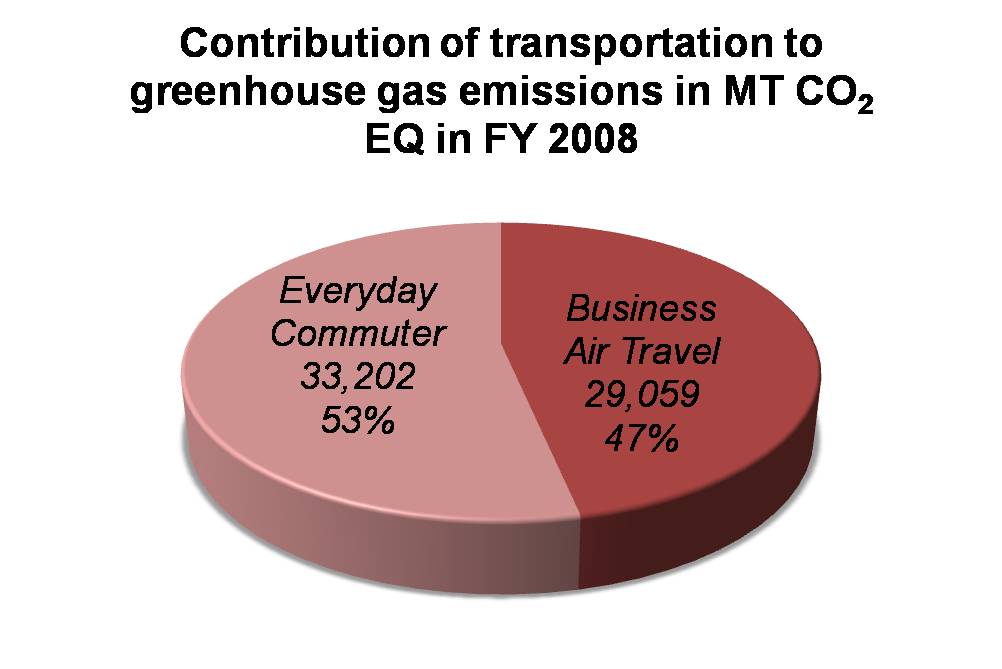
The lack of collaborative technologies which provide a rich interaction experience (both structured and unstructured) has resulted in a continued need for travel and face-to-face interactions to advance research and academic collaborations. This in turn has a direct impact on the collaborators’ carbon footprint. There is increasing awareness of the environmental impacts caused by this need for travel and face-to-face interactions. Although travelling to research meetings (by airline or automobile) is not explicated as an activity contributing to overall anthropogenic global warming gases by the IPCC (Intergovernmental Panel on Climate Change), commuting to work by fossil fuel-powered automobile for collaborative meetings and travelling by airline to research meetings undoubtedly emit substantial quantities of global warming gases into the atmosphere. For example, shows the total contribution of travel to greenhouse gas emissions at Arizona State University in FY 2008. This chart compares the contributions (from student, faculty and staff) of travel by everyday commuters, against air travel needed to attend research meetings, and illustrates the significant effects of travel for research meetings on the carbon footprint. While this observation is intuitive, this demonstrates the strong need to guide the development of new technologies that enrich remote research interactions in an ecologically responsible and sustainable manner.

Figure 2: Contribution of travel to greenhouse gas emissions at ASU

In this proposal, we integrate three distinct research perspectives namely: socioemotional modeling, human-centered multimedia computing and sustainable design. While the proposed project will result in transformational research in each of the component disciplines, the most important contribution is in the paradigm-shifting transformational research at the confluence of the three disciplines. The outcome of this project (e-REACH) will be a creative, human-centered computational framework, based on socioemotional models, that is evaluated on stringent measures of human factors and sustainable design.

e-REACH will be coordinated by the Center for Cognitive Ubiquitous Computing (CUbiC) at Arizona State University (ASU). The spirit of this project aligns with the goals of ASU, as defined by President Crow for a New American University based on the concept of ‘One University in Many Places’, trans-disciplinary research and education, and sustainable development. ASU’s several thousand faculty, researchers and students across geographically dispersed campuses require technologies like the proposed **e-REACH** system to advance research and education. The strong commitment of ASU to global leadership in sustainability (through the establishment of the Global Institute of Sustainability) includes substantial efforts to reduce the carbon footprint. Hence, this project has strong support from the University (please see enclosed letter of support from University President Michael Crow), and the institution has committed to being a test bed. More relevantly, the University Technology Office (UTO) has agreed to provide us with access to the immersive telepresence system installed on campus for the proposed research.

1. **Research Objectives**

We propose to design, develop and evaluate a human-centered computational framework that enriches interactions among academic collaborators across disciplines and distances. The goals of **e-REACH** will be achieved through the pursuit of the following fundamental research objectives:

1. ***Investigate the fundamental behavioral elements of effective collaborative interactions, with a focus on non-verbal cues of socioemotional states exchanged during face-to-face interactions.***
2. ***Design, develop and evaluate a human-centered computational framework for collaborative technologies using innovative fusion of multi-modal on-body and ambient sensor/actuator technologies that capture and deliver interaction primitives in a seamless, unobtrusive manner.***
3. ***Evaluate and optimize e-REACH technology prototypes using ergonomic design principles and study the ecological impact of the collaborative technologies.***

e-REACH draws upon intellectual synergies among experts in socioemotional communication (Michelle Shiota, Social Psychology); human-centered multimedia computing, pattern recognition and multimodal interfaces (Sethuraman Panchanathan, Computer Science and Engineering); and sustainable ecological product design and human factor impact analysis (Philip White, College of Design). The proposed research will result in the advancement of computational thinking at the intersection of the component disciplines, namely social psychology (through the introduction of novel methodologies for computational analysis and evaluation of social behavior in human interactions), multimedia computing (through the design of a novel human-centered computational framework driven by principles of socioemotional psychology), and ergonomic and sustainable product design (through the generation of new evaluation strategies and design impact factors – both ergonomic and environmental – for collaborative technologies). Our industry partners include Freescale (who have expressed their continuing support in providing state-of-the-art wearable sensor technologies) and Motorola (who have expressed interest in sharing their expertise on haptic actuator technologies and mobile computing solutions). The letters of support from both Freescale and Motorola are attached to this proposal.

1. **Results of Prior NSF Support**

In the past, CUbiC@ASU has been successful in the design, development and deployment of innovative multimedia technologies and systems through support from various NSF programs, as listed below:

Table 1: Prior NSF Support received by the PI

|  |
| --- |
| **Title:** HCC: The iCare Ambient Interactive Shopping Environment (Award ID: 0739744), **Amount:** $156,927, **PI:** T. Hedgpeth, S. Panchanathan, **Duration:** October 2007 - September 2008 |
| **Title:** ITR: iLearn: IT-enabled Ubiquitous Access to Educational Opportunities for Blind Students (Award #: IIS 0326544), **Amount:** $1,112,212, PIs: S. Panchanathan, K. S. Candan, F. Golshani, T.Hedgpeth, P. Green, **Duration:** September 2003 - August 2008. |
| **Title:** SGER: Incorporation of a psychological basis of haptics in the design of assistive haptic user interfaces (Award ID : 0554698), **Amount:** $196,574, PIs: S. Panchanathan, D Homa, D. Hansford T. Hedgpeth, **Duration:** November 2005 - October 2006. |
| **Title:** NSF-PPD-FRI: Ubiquitous Environment to Facilitate Engineering Education for Blind Persons (Award #: HRD 0333452), **Amount:** $172,538, PIs: S. Panchanathan, K. S. Candan, M. E. Donderler, T.Hedgpeth, and P. Green, **Duration:** October 2003 - September 2005. |

As mandated by NSF guidelines, we elaborate the results of the recently concluded successful NSF ITR iLearn (Award #: IIS 0326544) project which directly led to the conceptualization of the proposed work. The iLearn project was focused on the development of a suite of assistive systems to help students with visual impairments succeed in their educational endeavors. These assistive systems included the *iLearn Reader*, the *iLearn NoteTaking Device,* and the *iLearn Information Assistant*. The success of the *iLearn* systems led to the expansion of its scope to support other activities of everyday living under the umbrella of the broader iCARE (**i**nformation Technology **C**entric **A**ssistive and **R**ehabilitative **E**nvironments) project, wherein a pilot prototype of the *iCARE Interaction Assistant* was designed and developed. The Interaction Assistant was primarily intended to enhance the social interaction experiences of individuals with sensory/perceptual and cognitive disabilities, such as blindness and autism.

We believe that blindness, while generally perceived as a disability, is in fact a psychologically meaningful concept pertinent to everyone (e.g., remotely located interaction participants are blind to the socioemotional signals of their counterparts).Therefore, the challenges faced in the iCARE Interaction Assistant directly relate to the objectives of the proposed e-REACH project, and have served as a significant motivator for the proposed research, where there is a need for technologies to assist in delivering rich social experiences across distances. Further, the challenges of prototyping ergonomic, unobtrusive and socially non-distracting technologies for individuals with disabilities have revealed the need to integrate stringent design guidelines as an integral component of any socioemotional technology augmentation. These experiences gained in designing the iCARE Interaction Assistant led to the conceptualization of this inter-disciplinary project, intended to explore new paradigms for interaction between remotely located collaborators while adhering to stringent ergonomic design guidelines. In addition, we believe that the strong orientation of projects at CUbiC to serve individuals with disabilities will be enriched further through the proposed effort, and will also ensure the participation of underrepresented groups at different stages of this project.

***Research Products:*** A brief description of the significant research products of the iLearn project is given in . A detailed report of the results and achievements of each of these research products are provided in Appendix A. A list of publications resulting from this research is presented in Appendix B.

Table 2: Research products of the iLearn Project

|  |
| --- |
| The ***iLearn Reader*** [20,21] was designed to provide individuals who are blind with the experience of reading books and other non-electronic material, with the same ease as their sighted counterparts. |
| The ***iLearn NoteTaking Device*** [22] focused on the development of a portable device intended to assist students with low vision (STEM students, in particular) for taking notes in their classrooms. |
| The ***iLearn Information Assistant*** [23,24] assisted students who are blind or visually impaired to navigate, access and annotate online course and curriculum materials on the Blackboard system. |
| The ***iCARE Interaction Assistant*** project [25,26,27,28] was focused on the initial design of a wearable system for individuals who are blind and visually impaired to recognize individuals in their proximity. |

**Contributions to Human Resource Development:** Two post-doctoral researchers, 5 PhD, 8 Masters and more than 10 undergraduate students gained valuable experience under the mentorship of the PI. More details of our contributions are presented in Appendix A.

In summary, the Interaction Assistant project provided insights into the concept of blindness, resulting in the strong motivation for exploring new paradigms for interaction on remote collaboration platforms. In addition, our prior successes in transdisciplinary research led to the formation of novel synergistic knowledge partnerships among social psychology, human-centered multimedia computing and sustainable product design/evaluation. We believe that the proposed project will result in transformational outcomes that advance research and education across disciplines and geographical boundaries.

1. **Research Plan**

While the efforts in each of the research objectives will be coordinated by one of our team members, the investigators and research personnel will work together as a team to realize the goals of every objective. Across the objectives, expertise in experimental and empirical activities will be provided by Co-PI Shiota; expertise in computational tools and frameworks will be provided by PI Panchanathan; and the expertise in human factor impact/design analysis will be provided by Co-PI White. The research objectives along with relevant prior work that has been performed by the project investigators are detailed below.

***Research Objective 1: Investigate the fundamental behavioral elements of effective collaborative interactions, with a focus on non-verbal cues of socioemotional states exchanged during face-to-face interactions. (Coordinator: Shiota, Social Psychology)***

Our experiences in trans-disciplinary research have shown the importance of face-to-face interaction for effective research collaboration. Face-to-face interactions are complex processes with several simultaneous channels of communication, including speech, tone of voice, facial expression, physiological responses (e.g., sweating, blushing), posture, and gesture [29]. While we are most aware of verbal speech, nonverbal cues convey a large portion of the information in a typical face-to-face interaction, and social psychologists have documented critical roles of several nonverbal cues in communicating fairly specific messages.

The proposed research will develop new computational frameworks to address both “sides” of the communication issue, (a) *Encoding* – identifying the non-verbal cues most important for effective face-to-face collaborative decision-making, and developing technologies for detecting these cues, and (b) *Decoding* – delivering information about these cues to the receiver in a way that is processed naturally, requires little cognitive effort, and leads to greater emotional impact on the receiver. Both of these aims are extremely innovative, at the cutting edge of both social psychology and human-computer interaction.

On the encoding side, we propose to employ a frequently used, laboratory-based task to examine the specific nonverbal cues that best predict effective collaborative decision-making. It is clear from prior research that collaborative decision-making is best performed via face-to-face interactions, but the exact nonverbal cues that facilitate this are as yet unknown. The metrics for predicting effective collaborative relationship development (such as trust building, accurate personality assessment, and interaction enjoyment) involve cues beyond our current knowledge. However, prior research offers evidence of specific nonverbal cues that help structure social interaction in dyads, and suggests a few possibilities. Smiles are universally recognized as signs of pleasure and welcome [30], and smiling and laughter are often used to elicit and/or offer social support [31], or to communicate social deference [32]. Other facial expressions and vocal properties reliably communicate several emotions important for social interaction, including anger, sadness, fear, embarrassment, confusion, and surprise [30]. The gaze direction indicates where one’s attention is focused: looking at a speaker indicates full attention, looking away indicates wandering attention, and a sudden change in eye gaze direction prompts partners to follow [16]. Head nods communicate a listener’s ability to follow the speaker’s message over time [33]. Posture can communicate dominance, submissiveness, and liking, among other states [15,17]. Although the role of these cues in facilitating dyad-level social communication has received some attention, researchers have not addressed the cues’ role in larger groups collaborating toward an instrumental goal. In phase 1 of the proposed research, we will assess the role of the nonverbal cues in effective face-to-face collaborative decision-making, and will also use an exploratory approach to search for other, as yet undocumented cues. We have already performed extensive research [10] on the nonverbal elements of unstructured interactions, with a focus on: (i) Facial expressions as measures of emotion in laboratory tasks, in studies of emotion regulation and positive emotion experience; (ii) Human facial and postural expressions of several specific positive emotions, including awe, amusement, pride, nurturant love, and anticipation, defined using Ekman-Friesen’s Facial Affect Coding System [34]; and (iii) Psychophysiological aspects of emotion (e.g., galvanic skin response, heart rate) in social interaction.

On the decoding side, we propose to develop technologies that deliver information about socioemotional states in a more naturalistic way, requiring little cognitive effort by the receiver, and evoking a natural emotional response. A rich body of research documents point to simple stimulus features that are instinctively interpreted as having emotional tone [35,36]. For example, a recent study by Williams and Bargh (2008) [37] found that participants’ ratings of an ambiguous target’s friendliness were influenced by the temperature of a cup of coffee they were asked to hold a few minutes earlier, i.e., feeling physical warmth actually made participants perceive greater “warmth” in others. Stimulus features such as this will be used in developing new ways to communicate sender socioemotional states to interaction partners (see for examples). We have extensive experience both in developing emotion-eliciting stimuli for use in laboratory research, and in developing unobtrusive signals that deliver information about the social environment [38,39].

Table 3: Stimulus Features Reliably Associated with Socioemotional States

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| *Stimulus Feature* | *Socioemotional State* | | | | | | |
| *Happy* | *Angry* | *Sad* | *Affectionate/Loving* | *Sympathetic* | *Dominant/*  *Proud* | *Submissive/Shy* |
| Color | Yellow | DarkRed  (unsaturated) | Gray | Rose Red  (saturated) | Green | Purple | White |
| Sound | High, variable pitch; high volume; rich harmonics; rapid tempo | High, variable pitch; high volume; narrow harmonics; rapid tempo | Low, static pitch; low volume; narrow harmonics; slow tempo | Unknown | Low, variable pitch; low volume; rich harmonics; slow tempo | High, variable volume | Low, static volume |
| Haptic | Swinging | Sharp Blow | Withdrawal  of pressure | Warmth, Stroking | Patting | Unknown | Unknown |

The proposed research would integrate these two areas of expertise, leading to the development of unobtrusive yet emotionally evocative technology-based signals that communicate socioemotional content. This research objective will be addressed in 2 phases:

***Phase 1:*** *What non-verbal cues of socioemotional states best predict effective decision-making in face-to-face collaborations?* This question will be addressed using the following study:

*Sample*: The sample will consist of 50 teams of 4 participants. All participants will be students currently enrolled in a doctoral program in engineering or in a physical or social science at ASU. Thus, the sample will be composed of young researchers from fields likely to benefit from collaborative research. Fifty percent of participants will be female, and approximately 30% of participants will be of Hispanic, Asian, or African-American ethnicity. Teams will be mixed-sex, mixed-ethnicity, and mixed-discipline, reflecting the types of collaborative research teams likely to make use of the developed technologies. Participants will be recruited via flyers posted in various locations on the ASU campus, as well as by word of mouth.

*Procedure:* Decision-making processes will be studied in the context of the “Desert Survival Task”, a structured problem-solving exercise used frequently in real-world team development as well as in research on group dynamics [8,40,41]. This laboratory-based task requires teams to make a difficult decision when facing a novel challenge, and assesses the quality of the team’s decision both in absolute terms, and relative to decisions made by individuals prior to collaboration. Because the task is “content-free” and is likely to be outside the experience of individuals contemplating the scenario, variability in performance strongly reflects the team’s overall problem-solving process. In the task, groups of four individuals imagine a scenario in which their airplane has crashed in the middle of the Arizona Sonoran Desert, in August. According to the scenario, the team has salvaged 15 items from the wreckage of the airplane, and must rank these items in terms of their usefulness in the current situation, in order to decide which to take and which to leave behind. First, individual team members produce their own rankings. Then the team is asked to discuss the problem and agree upon a team-level solution.

*Measures:* Non-verbal behaviors during the Desert Survival Task will be assessed via coding. High-fidelity audio-visual recording of each team member will be collected throughout the task. These recordings will be coded by research assistants trained and certified in Ekman and Friesen’s Facial Action Coding System for the following non-verbal cues of socioemotional states: (a) smiling; (b) laughter; (c) facial displays of anger, sadness, fear/anxiety, embarrassment, confusion, and surprise; (d) non-verbal vocal displays of excitement, interest, anger, sadness, fear/anxiety, contempt/sarcasm, whining, dominance, and submissiveness; (e) gaze direction toward the speaker versus away/drifting; (f) head nods during another’s speech; (g) dominant/expanded versus submissive/constricted posture; (h) leaning toward another participant; and (i) touching another participant. Prior to coding, Co-PI Shiota will examine the audio-visual records in an exploratory manner, looking for previously unrecognized nonverbal cues contributing to collaborative success; if identified, these cues will be added to the coding system to be applied independently by the research assistants. Codes will be summed for each participant across the task, and also summed for each team. 20% of the data will be coded by both/all research assistants, and intra-class correlations calculated, to ensure that coders generally agree with each other.

With respect to decision-making quality, individual and team solutions to the Desert Survival Problem are scored against a “correct” expert solution, producing a quantitative measure of decision-making quality. The difference between the team score and the expert rating provides an index of overall quality; the difference between the team score and the average individual score provides an index of the effectiveness of collaboration in improving upon individual decision-making quality. In addition to these core outcome variables, we will assess time needed to reach a team-level solution; participant ratings of task enjoyment; and participant ratings of team cohesiveness.

*Data Analysis:* A multiple regression approach will be used to examine the relative importance of each of the coded non-verbal cues in predicting each of the five outcome variables: (a) absolute team-level solution quality on the Desert Survival Task, (b) increment in solution quality associated with collaboration, as compared with individual-level solutions, (c) time needed to reach a team-level solution, (d) ratings of task enjoyment, and (e) ratings of team cohesiveness. Preliminary analyses will enter all non-verbal cues simultaneously in predicting each outcome; subsequent analyses will use an iterative process to identify the most efficient (i.e. balancing magnitude of prediction effect with parsimony in number of cues) combination of non-verbal cues predicting solution quality and other outcomes. Computational frameworks from machine learning such as clustering (techniques such as spectral clustering) and non-linear techniques (such as manifold learning) will be used to analyze the results. To the best of our knowledge***,*** the use of machine learning frameworks for analysis in social psychology tasks will be the first of its kind, and will no doubt open up new vistas for computational thinking in social psychology*.*

***Phase 2:*** *How do the proposed technology solutions compare with face-to-face communication and existing telecommunications technologies in terms of the ratio of decision-making effectiveness to environmental impact?*This question will be addressed using the following study:

*Sample:* The sample will consist of 100 teams of 4 participants each. As in Phase 1, all participants will be students enrolled in a doctoral program in engineering or in a physical or social science at ASU; 50% of participants will be female; approximately 30% of participants will be ethnic minorities; and teams will be mixed-sex, mixed-ethnicity, and mixed-discipline. Teams will be randomly assigned to one of five Interaction Medium conditions: (1) face-to-face interaction; (2) interaction via proposed augmentations of technology solutions; (3) interaction via computer-mediated instant messaging; (4) interaction via telephone conference call; or (5) interaction via desktop video conferencing.

*Procedures and Measures:* Collaborative teams will complete the Desert Survival Task used in Year 1 to measure decision-making quality.

*Data Analysis:* We will evaluate how proposed technology augmentations lead to improved decision-making relative to any of the existing technology controls, and whether the new technology still shows significant deficits in facilitating virtual collaboration relative to face-to-face communication. Planned comparisons within an Analysis of Variance approach will also be used to assess the statistical significance of differences between the new technology and existing technologies using the following criteria: (a) absolute team-level solution quality on the Desert Survival Task, and (b) increment in solution quality associated with collaboration, as compared with individual-level solutions, (c) time needed to reach a team-level solution, (d) ratings of task enjoyment, and (e) ratings of team cohesiveness. In addition to effectiveness of decision-making, we will evaluate the environmental impact of each of the technology augmentations, as captured by the Life Cycle Assessments (LCA)[42] detailed in Objective 3.

***Research Objective 2: Design, develop and evaluate a human-centered computational framework for collaborative technologies, using innovative fusion of multi-modal on-body and ambient sensor/actuator technologies that capture and deliver interaction primitives in a seamless and unobtrusive manner. (Coordinator: Panchanathan, Human-Centered Multimedia Computing)***

Socioemotional primitives that are essential for effective collaborative interactions (as identified by Research Objective 1) need to be unobtrusively captured, seamlessly communicated, and discreetly presented to the participant(s) in an interaction. This first entails innovative fusion of multi-modal sensor channels, and extracting socioemotional primitives using adaptive computational learning techniques. These primitives are then translated into delivery cues that can be presented using multi-modal (haptics, audio, visual) actuators. The proposed framework will be employed in collaborative technology configurations that are informed by ergonomic analyses and modified, where applicable, to minimize overall environmental impacts (reduced carbon footprint). Fusion of multiple channels of non-verbal cues to extract individual primitives (such as anger expressed by a raised lower eyelid, increased vocal volume and higher heart rate); the need to understand and extract subtler interpersonal primitives (such as dominance or liking), which often even elude human judgment; and the need to adapt to stringent design guidelines of naturalistic interactions under varying contexts all make this work transformational in its purpose, paradigm-shifting in its design, and path-breaking in its evaluation process.

The proposed work will be anchored to the Desert Survival Task context specified in Research Objective 1. The manual coding carried out as part of the psychology studies will provide the necessary ground truth for the collected data streams. The computational framework, along with the sensors/actuators, will be evaluated within the experimental study detailed in Phase 2 of the Desert Survival Task.

The proposed computational framework will address several fundamental research challenges with the potential to transform traditional pattern recognition and machine learning approaches, moving towards human-centeredness in multimedia computing. The key research challenges include:

* Identifying the optimal placement of sensors, mapping sensor data streams to socioemotional primitives (see below), and establishing confidence levels to ensure graceful degradation of performance in the context of missing sensor information (or other causes of uncertainty).
* Fusing raw data streams from multiple sensors and extracting high-level features (such as Fourier descriptors from inertial sensor signals) that correspond to specific socioemotional primitives.
* Designing dimensionality reduction techniques that derive low-dimensional embeddings from high dimensional data streams (eg. video) that represent socioemotional primitives in interactions.
* Designing kernel learning functions that project data streams (eg. accelerometer data) into feature spaces that represent socioemotional primitives in interactions.
* Designing delivery cues that translate socioemotional primitives into multimodal (haptics, audio and/or visual) actuations that evoke *naturalistic* emotional responses in recipients with minimal cognitive load.

The above research challenges will be discussed as part of the *capture*, *processing* and *delivery* modules of the proposed computational framework.

***Capture Module:*** With recent advances in technology, state-of-the art sensors assume miniaturized form factors, consume low power and are wireless, thereby providing a robust platform for gathering information about the various nonverbal cues (constituent components of socioemotional primitives) in a non-intrusive manner. On-body sensors collect different physiological signals from a first person perspective of the interaction partner while the ambient sensors offer a third person perspective into the interaction scene thereby assessing interpersonal interactions. lists the different on-body and ambient sensors along with their potential use in extracting nonverbal cues.

Table 4: Different on-body and ambient sensors with the corresponding nonverbal cues

|  |
| --- |
| ***Inertial sensors (Accelerometers, Gyroscopes & Magnetometers):*** Body posture, gestures, head movements conveying socioemotional primitives |
| ***Galvanic Skin Response:*** Perspiration associated with sympathetic nervous system activation (characteristic of fear, anger, startle response, orienting response and enthusiasm) |
| ***Heart Rate/Blood pressure:*** Cardiovascular response (characteristic of several socioemotional states, such as anger, fear, and excitement) |
| ***Temperature:*** Peripheral vascular response (characteristic of socioemotional states such as embarrassment and anxiety) |
| ***Microphone:*** Prosody (vocal tone variation), paralinguistic vocal cues (e.g., “uh-huh,” “hmmm”) |
| ***Pressure Sensors:*** Posture, gestures conveying socioemotional primitives |
| ***Cameras:*** Ambient information, facial expression, body posture, gestures, head movements. |

The PIs have extensive experience in the use of various sensors for capturing body movement, facial information, physiological activity, and emotion psychophysiology. The PI (Panchanathan) has worked on detecting and recognizing motion patterns in activities ranging from simple ambulation like walking or sitting [43] to complex activities like making a glass of tea [44] using inertial sensors (accelerometers, gyroscopes, etc.); detecting specific motion patterns such as stereotypic body rocking using triaxial accelerometers [45]; capturing facial landmarks [46], head orientation [47], person identity [48], and interpersonal distances [49] using wearable camera systems; understanding muscular activity using electromyography (EMG) signals; and understanding user intent using eye-tracking technologies [50]. The Co-PI, Shiota, has worked extensively with sensors for capturing bioelectrical measures of emotion (such as galvanic skin response, heart rate and skin temperature). We will use our combined expertise in critically assessing the relative merits of each sensor and combination of sensors for capturing the various socioemotional primitives identified in *Research Objective 1*. In addition, based on our prior experience on investigating strategic placement of accelerometers for human activity recognition, we will identify the optimal placement of sensors as required by the different interaction mediums in the Desert Survival Task (emails, cell phones, audio/video conferences, etc.).

Simple socioemotional primitives such as smiling are typically captured using one sensor modality, namely video. However, complex primitives such as anger, dominance, etc. require integration across multiple sensor modalities, namely video, accelerometers, audio (voice prosody), pressure sensors, etc. In such situations, the reliability and information content of each of these sensors may be different, and hence the combined sensor data streams may need to be treated with varying levels of confidence. In addition, missing sensor data and the failure of sensors can result in degradations of performance. There is therefore a need to determine not only the level of confidence in the result but also ensure graceful degradation in performance. The PI has worked on fundamental approaches to estimating confidence levels from multiple data streams in the context of head pose estimation [47]. We plan to extend this work to derive confidence levels for each of the sensor input streams, thereby allowing for suitable adaptations to offset missing sensor inputs, and to handle scenarios with high levels of uncertainty.

***Processing Module:***Socioemotional primitives of participants in interactions can be broadly divided into individual primitives (such as smiling, anger, confusion, etc.) and interpersonal primitives (such as dominance, liking, trust, etc). Existing efforts in affective computing and related multimedia computing fields [51] aim to capture straightforward individual primitives using unimodal (vision-based) sensors. However, the interplay of individual emotions in group interactions (smiles exchanged between two participants) as well as increasing subtlety of primitives involved in interpersonal collaborative interactions (dominance of a participant) presents new challenges that have hitherto not been attempted.

The significant need to process data streams from multiple sensor inputs and extract the aforementioned primitives in real-time requires a human-centered thinking towards developing a computational framework. Based on our experience in working with high-dimensional data (image and video) in applications such as face recognition, activity recognition and head pose estimation, we strongly believe that the ambitious goals of the proposed work (as mandated by the computational constraints and challenges of different interaction mediums specified in the Desert Survival Task) will be achieved only when: (i) salient features that are highly representative of the socioemotional primitives are extracted not just individually from each sensor input, but simultaneously from appropriate multi-sensor sources; (ii) feature streams that correspond to each of these primitives are transformed into meaningful low-dimensional representations (embeddings) that results in a lower computational complexity; and (iii) data and high-level feature streams are mapped into discriminative feature spaces using kernel functions which capture the semantics of the primitives. While individual socioemotional primitives (such as enthusiasm, anger, confusion, etc.) will be analyzed from sensor inputs from respective individuals, we plan to use the sensor inputs from multiple individuals along with ambient sensors to understand and extract subtler interpersonal socioemotional primitives such as dominance, liking, etc. This approach of processing sensor data using techniques informed by sociopsychology principles forms the backbone of the proposed *human-centered* computational framework.

*Fusing raw data streams from multiple sensors and extracting high-level features that correspond to specific socioemotional primitives:* Selection of appropriate features (such as Fourier descriptors from inertial sensor signals) from the sensor channels is critical for ensuring reliable performance of the processing algorithms. Existing feature selection algorithms (both feature ranking, where a metric is used to rank features of data streams, and subset selection approaches, where an optimal subset of features is obtained from a set of possible features) rely on filter metrics such as correlation, mutual information, class separability, and other optimality criteria such as the Akaike Information Criterion [52] and the Bayesian Information Criterion [53]. However, in order to incorporate the understanding of sociopsychology principles in selecting features, we plan to design appropriate filter metrics and optimality criteria that measure the extent to which the selected features mimic the distinct expressive properties of respective sensor channels. For example, features from the mouth region in face images will be evaluated using filter metrics that ensure a high correlation between the feature vectors for wry, gentle and broad smiles, but a low correlation between a gentle smile and pursed lips. Over the last 2 years, the PI has worked on techniques for feature selection that capture the essential movement primitives for human activity recognition [43], which directly relates to this objective.

Secondly, socioemotional primitives such as anger are conveyed through multiple cues such as increased vocal volume, raised lower eyelids and higher heart rate. None of these cues alone indicates anger, but in combination, they clearly do so. This necessitates extracting features from data aggregated across appropriate sensor channels. We note that there are no straightforward solutions in existing literature on multimedia fusion to address this problem. We propose to use the “minimum-redundancy-maximum-relevance” (mRmR) scheme [54], with appropriate definitions of redundancy and relevance derived from sociopsychological principles, to select the features using the combined sensor inputs, thereby generating succinct representations of voluminous data from multiple sources.

All the above feature selection methods will be applied directly on the signal streams in the case of sensors such as accelerometers, gyroscopes, skin conductance sensors, and pressure sensors. For audio and video, we plan to use commercially available state-of-the-art voice stress analysis packages to extract low-level audio features from audio sensor data (like syllable length, volume, pitch, formant frequencies of speech sounds, rhythm, length, and tension), and low-level visual features from camera sensor data (such as eyes, eyebrows, lips, and ears) using face processing tools such as FaceVACS and FaceAPI. These low-level features will be used as inputs to our techniques to select high-level features that correspond to specific socioemotional primitives.

*Designing dimensionality reduction techniques that derive low-dimensional embeddings from high dimensional data streams that represent socioemotional primitives in interactions:* Recent efforts in applying non-linear dimensionality reduction techniques (manifold learning) to human movement analysis [43] have suggested that data streams generated from specific movements can essentially be represented as low-dimensional representations (embeddings) where the number of dimensions reflect the degrees of freedom in the source. This serves as the motivation to design techniques that generate low-dimensional embeddings from high-dimensional data to reduce the computational complexity, while retaining the meaning of the socioemotional primitives involved. As an example, it may be possible that features extracted from the mouth region from face images (which could be of the order of 10 or 20 dimensions) may be represented as a 5-dimensional feature vector that captures the essential movements of the mouth.

Traditional dimensionality reduction techniques (including linear methods such as principal component analysis, factor analysis, projection pursuit, independent component analysis, and non-linear methods such as multi-dimensional scaling, self-organizing maps, principal curves, manifold learning, Gaussian process latent variable models, and other non-linear variants of the linear methods) rely on objective functions that optimize the low-dimensional embeddings based on covariance and inter-data distance matrices. However, we propose to design appropriate objective functions that, in addition to using data, are modulated using relationships between primitives identified from sociopsychology. For example, the correlation between anger and confusion primitives (obtained from literature and empirical studies) will be used to compute the distance between the corresponding feature representations, before deriving the low-dimensional embeddings for each of these socioemotional primitives. Such objective functions can be generically used across all the aforementioned dimensionality reduction techniques, thereby providing a fundamental approach to address this challenge. The PI recently has developed the Biased Manifold Embedding framework [47] that represents high-dimensional data from face images with varying head poses as a 3-dimensional manifold that captures the pitch, yaw and roll of the head movement, using techniques such as Isomap [55], Locally Linear Embedding (LLE) [56] and Laplacian Eigenmaps [57]. This approach has been promising, and will be explored further to design techniques to derive low-dimensional embeddings that represent socioemotional primitives in interactions.

*Designing kernel learning functions that project data streams into feature spaces that represent socioemotional primitives in interactions:* Over the last decade, kernel-based methods have grown to be extremely effective in machine learning. Kernel functions map data from a given space into a discriminative feature space (generally of higher dimensions), where the understanding and analysis of data is simpler, more effective and more purposeful. In addition, computational simplifications (called the kernel trick) make kernel functions efficient, and almost all processing algorithms (such as Support Vector Machines, Fisher's linear discriminant analysis, principal components analysis (PCA), ridge regression, spectral clustering, etc.) can be ‘*kernelized’*. In earlier work, the PI has extensively applied kernel-based pattern recognition methods for multimedia data on problems such as human activity recognition [43], face recognition [46], environment perception [58], and clinical decision support [59].

Kernel functions effectively capture the inner product or the distance between data representations in the feature space. We propose to design kernel functions (satisfying the mathematical constraints of the Mercer’s theorem) that are modulated using relationships between socioemotional primitives, similar to the techniques for dimensionality reduction. While dimensionality reduction scales the data down to dimensions in turn lowering the computational complexity, kernel functions provide a means to use algorithms (such as neural networks) to identify the primitives from the low-dimensional embeddings.

***Delivery Module:*** Presenting socioemotional primitives to the receiver in a collaborative interaction necessitates reformatting the primitives (or the corresponding low-dimensional embeddings) into actuations (through haptics, audio and visual modalities) that evoke a corresponding emotional response. In addition to audio tones and speech, recent advancements in actuator technologies (such as vibrators, solenoids, air bladders, Peltier junctions) provide the motivation to use vibrotactile cues (vibratory signals defined by signal frequency, intensity, rhythm, and duration of the vibration) through haptic devices for discreet presentation of information [38] about interaction primitives. Also, recent technologies such as the hug shirt provide creative opportunities to deliver social touch cues in interactions, which are evidently missing in the most sophisticated of existing virtual communication environments. Further, psychology research supports the use of cues such as color and temperature to evoke emotional responses. shows potential actuator technologies with corresponding stimuli and interaction cues that will be used to evoke emotional responses among remote partners.

Table 5: On-body and ambient actuator technologies for conveying interaction cues

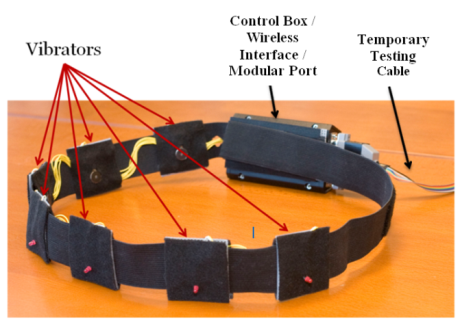
|  |
| --- |
| ***Vibration motor (Vibration) and Solenoids (Pressure, vibration***): Social touch cues (e.g., hugs, taps), nonverbal cues (e.g., emotions, body language), general communication (e.g., tactile icons) |
| ***Piezoelectric motor (Lateral skin stretch):*** Social touch cues (e.g., tap, pinch, scratch), non-verbal cues and general communication (e.g., haptic icons) |
| ***Heating /Cooling using Peltier junctions, heating coils (Temperature variations):*** Interaction cues (e.g., body warmth) |
| ***Force-feedback devices (Kinesthetic):*** Social touch cues (e.g., handshake), non-verbal cues and general communication (e.g., haptic icons) |
| ***Color of the ambient settings or surroundings (Visual):*** Colors can evoke direct response of emotions like happiness, anger, sadness etc, or indirectly evoke emotional response by coding temperature like red of heat, blue for cool etc. |
| ***Tones and music in the surroundings (Audio):*** Similar to subtle visual cues, audio can be used for evoking emotional states in interaction partners. Stereoscopic manipulation of tone can convey sense of space and direction. |

We will address the following specific challenges in this module for each of the identified primitives from Research Objective 1: (i) the translation of feature representations of these primitives into sentences of actuations that are delivered through multiple modalities like haptics, visual and audio; (ii) the design of schemes of presentation that not only deliver the primitive, but evoke naturalistic responses in the recipient (for example, liking could be conveyed through warmth, or through vibrotactile stimulations); and (iii) the identification of the most suitable presentation scheme (for example, a happy response could best be evoked as a change in the color of the visual field to yellow, rather than haptic actuations).

Existing telecommunication media are primarily audio-visual in nature. Hence, the use of audio and visual modalities for delivering cues will have to be carefully handled to prevent cognitive overload on the user with respect to these modalities. However, haptics (proprioception and tactile) provide an alternate modality which is underutilized in these media. Hence, we plan to focus our efforts in delivering the non-verbal cues primarily using the haptic modality, supplemented with appropriate audio-visual cues. In the last few years, the PI has worked on designing haptic delivery devices (such as the haptic belt, ; Provisional US Patent Case # M9‐051P) for discreet communication of non-verbal cues such as interpersonal distances and direction to user(s) in an interaction (in the context of development of assistive technologies for individuals who are blind, as stated in Section 4). These devices have been designed in collaboration with the College of Design at ASU, and were refined using feedback from several focus group studies, design prototypes and evaluation methods (detailed in Research Objective 3).

We have studied methods for inter-modal information transfer using haptic devices (the Immersion CyberGlove, SensAble Phantom and Falcon) to deliver complex visual concepts like color through vibrotactile actuations. This cueing methodology was tested with several subjects (individuals who are blind and sighted-blindfolded individuals). In addition to learning to recognize colors that were used in the training phase, the intuitiveness of the cueing methodology enabled the participants of these experiments to recognize newer colors (that were not used during training) accurately. Our multi-dimensional scaling analysis of this study revealed a one-to-one mapping between the color space (HSV) and the dimensions of the vibrotactile cues. This research study provided us with numerous insights into the methodology required to design naturalistic delivery cues for complex concepts. We will leverage this knowledge in the proposed work to deliver socioemotional primitives in collaborative interactions.

Figure 3: Haptic Belt developed at CUbiC

While we have explored the use of a haptic belt as a delivery device, other form factors such as haptic suit, haptic glove and haptic objects (like balls or grippers) will be explored for effective presentation of interaction primitives. In a more recent study, we used vibrotactile rhythms to convey interpersonal distances and directions in interactions. However, our current efforts in using vibrotactile cues for delivery have been exploratory, and we now plan to design and develop a formal cueing framework that uses variations in signal frequency, rhythms, tones and durations of vibrotactile cues to deliver the socioemotional primitives. In the past, we have also studied the conflicts of cross-modal delivery of interaction cues (for example, since a primitive like anger may be delivered through vibrotactile cues or changes in audio tones, we need to understand how one modality of presentation may conflict with another modality that is being used during the user activity) from a cognitive perspective [60]. We plan to extend this work to achieve effective cross modal communication, through automatic adaptation of presentation media, based on the availability of technologies at the user’s end.

***Research Objective 3: Evaluate and optimize e-REACH technology prototypes using ergonomic design principles and study the ecological impact of collaborative technologies (Coordinator: White, Sustainable Product Design)***

This objective has two sub-components: ergonomic design and environmental impact. The motivation for stringent guidelines of ergonomic design in interaction media is derived from our prior experience in designing the iCARE Interaction Assistant, as explained in Section 4. Furthermore, since the primary purpose of e-REACH is to reduce carbon footprint through face-to-face meetings, the proposed technology augmentations have to be evaluated on environmental impact criteria, against travel for face-to-face interactions. We discuss these sub-components below:

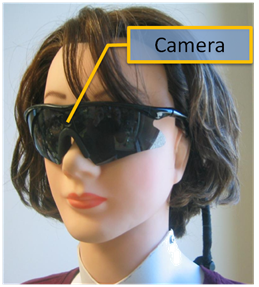
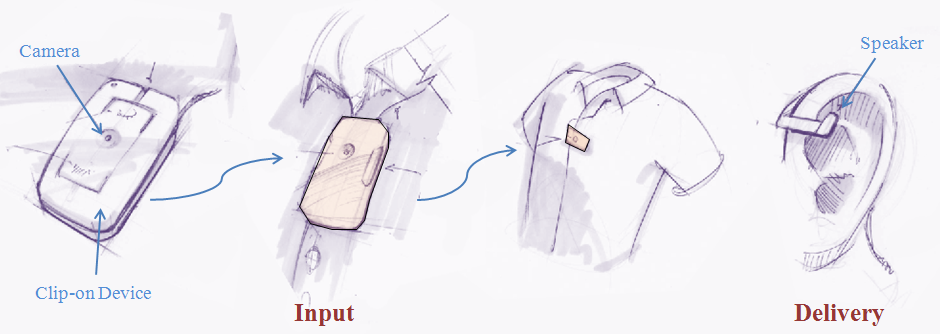
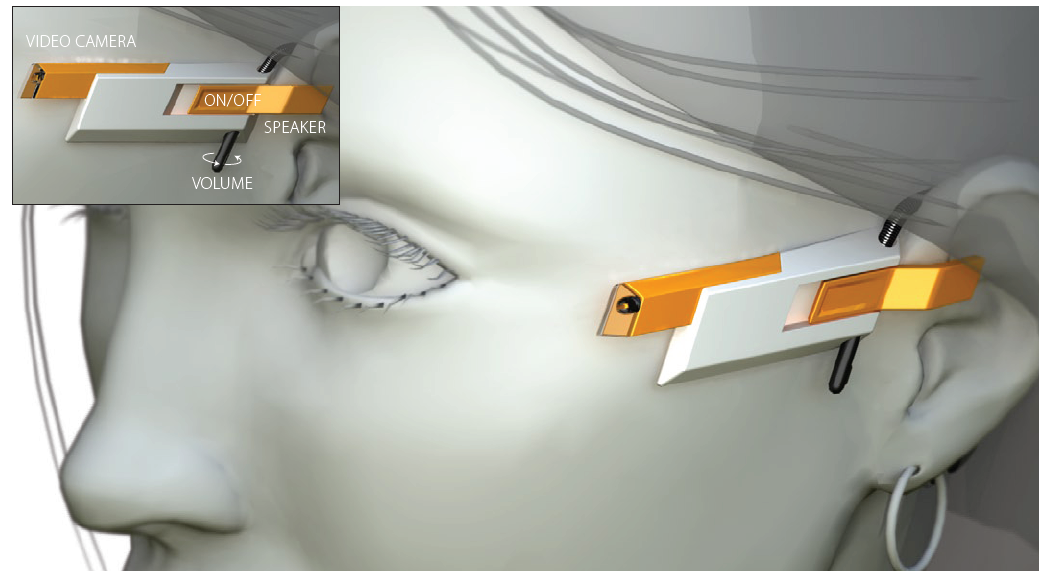
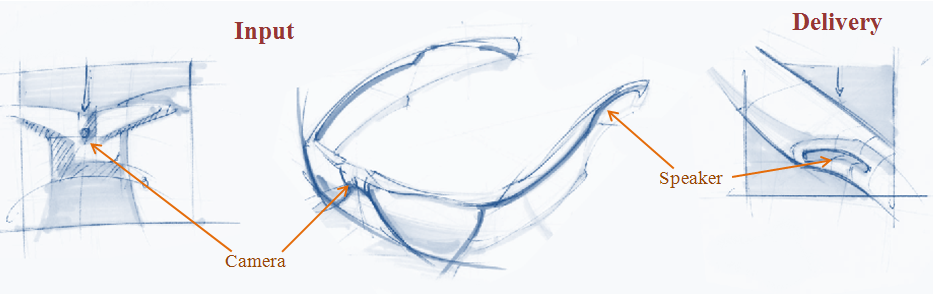
***Ergonomic Design:*** The e-REACH technology augmentations must offer an experience that is as comfortable and natural as typical face-to-face interactions. This requirement for naturalness (design, aesthetics and usability) is pertinent to all sensor/actuator technologies used in the capture and delivery components of Research Objective 2. Failure to cater to any of these requirements will stretch a user beyond his/her comfort perimeter will not only induce physical or cognitive overload, but may be counterproductive to the purpose of the interaction [61]. A straight forward approach to ensuring user acceptance is to meet strict ergonomic design criteria of social acceptance. That is, a technology introduced to mediate social interactions should carefully consider various design requirements like shape, form factor, placement, etc. to ensure that no undesired social artifact is created during the interaction. Hence, it is of utmost importance to consider social factors while developing metrics to evaluate the design prototypes. While product design has typically focused on aesthetics, there exists no principled approach to design technology prototypes for collaborative interactions and social mediation.

Ergonomic development and evaluations of prototypes are typically done through observations and interviews. This will be executed in 3 stages:

1. The first step will assess the usefulness of the target socioemotional primitives with the feasibility of sensing and delivery using state-of-the-art technologies. This work will be in collaboration with both psychology and human-centered multimedia teams.
2. The next step will be the development of 2-D visual concepts via sketching or CAD modeling to convey complex ideas about physical product configurations. Once the participants of the experiments have completed the DS tasks (in Research Objective 1), the 2-D visual concepts will be presented with detailed explanation of the potential use for such technologies. Feedback about the visual concepts will be collected from the participants (belonging to different age groups, gender, cultural and professional backgrounds). Typical questions include: “Would you be comfortable using the prototype on a scale of 1-5?”, “How satisfied are you with the form factor, etc. This feedback will be used to further refine the 2-D visual concepts leading to the design of 3-D prototypes. In this step, the usability and design team will work closely with the psychology team.
3. In the third step, the 2-D visual concepts will be turned into 3-D prototypes with ergonomic design factors that consider humans and their environments, anthropometry (anatomical geometries distributed according to age, gender, or race), biomechanics, kinesiology, and psychology. Design prototypes will be assessed through interviews and expert observations in mock remote communication scenarios enacted by users. Functional prototypes will be assessed through task analysis where feedback from observing experts and users will be used to choose the technologies.

In the past, we have designed prototypes for people who are blind and visually impaired to engage in tasks such as reading, recognizing people, etc. shows 2D design concepts used for the design of a wearable camera system to assist individuals who are blind to recognize people in front of them. The 3D prototype for the wearable social interaction assistant (Fig. 4) was developed based on the feedback received about the 2D visual concepts. The rigorous design evaluation proved to be a significant factor for the technology to be well-received by the target population.

Figure 4: The 2-D sketches of the prototype (on the right) and the 3-D prototype (on the left)



***Environmental Impact:*** Recent research in “greener” engineering and design offers new options for sensor and telecommunication equipment manufacturers [62]. But, environmental impact analysis that takes into account the production, maintenance and end of life treatment of sensors, devices and telecommunication infrastructures has not been modeled in a systematic manner. We address this important issue which will be a major contribution of this project. For example, there is a thorough understanding of how cellular telephones contribute to a wide range of environmental and human health impacts over their life cycles, but there is much less understanding about impacts of maintaining enormous telecommunication infrastructure that make cellular phone communications possible.

Since e-REACH would offer technology augmentations to existing telecommunication equipment and services, understanding the sustainability impact of e-REACH necessitates life cycle assessments (LCA) of both sensor and communication devices as well as complete telecommunication infrastructures. Such life cycle assessment offers (i) system level modeling of environmental and human health factors for various telecommunication services, and (ii) offers a base model to compare any new technology development. To this end, we propose to use Life Cycle Assessment (LCA) as a means of modeling the environmental impact of new technology developments and hence offer a comparative scale where travel for face-to-face interactions and use of telecommunication systems can be mapped.

*Procedure:* LCA follows the steps of 1) Process inventory collection, 2) Impact Characterization, 3) Normalization and 4) Results interpretation. As the first step, we will collect the “system bill of materials” needed to conduct the Life Cycle Inventory (LCI) of the various communication systems (including infrastructural elements). These will be combined with the device level LCI data [63], which are more readily available than the infrastructural system data, to model the overall system impacts. We will identify appropriate system boundaries and LCI data which apply the boundaries consistently, so that the results of the assessments are comparable. As an example, we have recently developed a comprehensive technique for road designers to model the life cycle climate change impacts of different pavement types including those that use recycled waste tires. The process employs variables that can be modified by the designer to customize for specific road configurations and road materials. Overall, the method allows engineers and planners to model the direct life cycle CO2 emissions related to pavement design. We propose to create analogous integrated assessment models for telecommunication technologies, namely email [64], cellular telephone [65], desktop video conferencing [66, 67], etc.

Measures: Environmental impacts will be quantified using established mid-point impact characterization methods, namely TRACI [68], and ReCiPe [69].

Table 6: Mid-point Impact Characterization methods

|  |  |  |
| --- | --- | --- |
|  | **TRACI (North American [68], US EPA)** | **ReCiPe (European [69])** |
| ***Ecological Health*** | Global warming (climate change), Acidification, Ecotoxicity, Stratoshpheric ozone layer depletion, Photochemical smog, Water eutrophication | Global warming (climate change), Freshwater, marine and terrestrial ecotoxicity, Stratospheric ozone creation, Freshwater and Marine eutrophication |
| ***Human Health*** | Human cancer, Human respiratory (criteria air pollutants), Human toxicity | Ionized Radiation, Particulate formulation, Human toxicity |
| ***Resource Depletion*** | Fossil fuel depletion | Mineral depletion, Water consumption, natural land transfer, urban land occupation, agricultural land occupation |

Characterized impacts, including global climate change potential (measured in pounds of CO2 equivalent), will be normalized globally and compared in appropriate functional units and clearly interpreted in conformance with ISO 14040 [70] series LCA standards. At the completion of the model, we will apply the same LCA methodology on proposed technology augmentations, as defined by the Desert Survival Task (DST). Normalization of the characterized values of each of the impact categories allows us to create a single score (global per person impact per person-year) that represents the overall environmental impact. This normalized single factor environmental impact score will be used as the common denominator in ratio of decision making effectiveness and environmental impact. This ratio will be the functional unit of the assessment for comparing different telecommunication systems with the environmental impacts of travel for face-to-face interactions [71].

**6. Evaluation Plan**

The empirical nature of the research carried out as part of Objectives 1 and 3 necessitate the inclusion of evaluation as a primary component of the research tasks themselves. However, in addition to these evaluations, formative evaluations will be carried out incrementally at different stages of the project, and a thorough summative evaluation will be rigorously performed at the end of the proposed work. Co-PI Shiota will guide the evaluations from the perspectives of socioemotional psychology, and Co-PI White will guide the evaluations from both ergonomic and sustainability perspectives.

*Formative Evaluation:*While Research Objectives 1 and 3 focus on evaluating the project at a system level through its different stages, the individual components of the proposed computational framework will also be evaluated at periodic intervals, as part of our formative evaluation process. A use-case based approach will be used to evaluate the performance of each of the components (capture, processing and delivery modules) of the computational framework. These use cases will be derived from specific interaction contexts of the Desert Survival task. Stringent performance metrics that are motivated by the requirement for low computational complexity and real-time performance will be used to evaluate the algorithms used in the proposed framework at every step of development. In addition to satisfying specified design criteria and meeting the need to extract appropriate socioemotional primitives, the sensor technologies used in the capture module will be regularly assessed using metrics of reliability and fidelity of the sensor streams. Metrics that capture the naturalness of delivery and the cognitive overload due to training will be used to evaluate multimodal actuator technologies and delivery strategies. The results of the formative evaluation process will be continuously integrated into design decisions of subsequent stages of the project.

*Summative Evaluation:* At the end of the project, a comprehensive summative evaluation will be carried jointly by the teams working under each of the investigators. This evaluation of the system will be performed to study the (1) usability of the proposed collaborative technology augmentations; (2) productivity in academic collaborations; (3) efficacy of interaction cues, and (4) environmental impact. System-level evaluations will be carried out between faculty collaborators located in four different campuses of ASU. External evaluators will conduct learner assessments and go over these system-level evaluations to quantitatively measure the efficacy of the technology augmentations under each of the criteria listed above.

# 7. Project Timeline and Schedule

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Tasks*** | ***Year 1*** | | | | ***Year 2*** | | | | ***Year 3*** | | | | ***Year 4*** | | | | |
| **Research Objective 1: Identification of socioemotional primitives** | | | | | | | | | | | | | | | | | |
| Face-to-face DST studies (Phase 1 study) |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Existing communication media DST studies (Phase 2 study) |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| e-REACH augmentation DST studies (Phase 2 study) |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| **Research Objective 2: Computational framework for capture, processing and delivery** | | | | | | | | | | | | | | | | | |
| Capture Module: Sensor selection and configuration |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Capture Module: Reliability analysis |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Processing Module: Fusion |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Processing Module: Dimensionality reduction |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Processing Module: Kernel machines |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Delivery Module: Actuator selection and configuration |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Delivery Module: Cueing framework and delivery strategies |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| **Research Objective 3: Ergonomics and Environmental Impact (EI) assessment** | | | | | | | | | | | | | | | | | |
| EI assessment of existing media |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Ergonomic designs for e-REACH technology augmentations |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| EI assessment of e-REACH augmentations |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| **Evaluation** | | | | | | | | | | | | | | | | | | |
| Formative Evaluation |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |
| Summative Evaluation |  |  |  |  |  |  |  |  |  |  |  |  | |  |  |  |  | |

# 8. Intellectual Merit

# The proposed interdisciplinary project exemplifies the goals of the NSF CDI program by developing new paradigms for next generation communication technologies with a focus on advancing research collaborations. The intellectual merit of the proposed e-REACH project lies not only in answering fundamental research questions in the fields of social psychology, multimedia processing, ergonomic human-computer interfaces and sustainable product design but more importantly opens up new research directions at the intersections of these disciplines. In the process of developing technology augmentations to capture and deliver socioemotional primitives, the understanding of socioemotional psychology itself will be advanced by newer insights derived from the computational analysis. Life Cycle Assessments (LCAs) of sensors, devices and telecommunication infrastructures in measuring their environmental impact offers a base model for any new technology development in this domain. This project in itself exemplifies the spirit of the proposal by bringing together teams of researchers across disciplines, collaborating to understand the limitations of existing technologies, and to propose novel solutions that incorporate transformational paradigms of interaction that will advance the frontiers of virtual organizations and collaboratories.

# 9. Broader Impacts

The proposed research project has the potential for significant impact in various applications, including social robotics, assistive and rehabilitative technologies, cyber-communication systems, online education portals and telemedicine. While the e-REACH project intends to integrate new paradigms of interactions in the context of remote collaborations, this research will directly impact all forms of human interactions across a range of electronic media platforms. This project will catalyze trans-disciplinary research, enhancing the infrastructure to share rich interaction experiences across different geographical locations, simultaneously resulting in potential reduction of carbon footprint. Workshops and special sessions will be periodically conducted to increase the awareness in different disciplines of the technology augmentations, as well as the underlying computational framework. The strong orientation of current and past projects of the PI to cater to individuals with disabilities will be continued to ensure the participation of underrepresented groups at different stages of this project. The need for active participation of faculty, researchers and students from different disciplines in the evaluation of the system will also be extended to promote teaching, training and learning in STEM education. Further, we will work closely with the Department of Education at ASU to conduct special workshops for K-12 students to understand and appreciate the potential of STEM fields in the development of a sustainable environment that can transform the way individuals interact with each other across distances.

**Coordination Plan**

The coordination plan will discuss the roles of the personnel involved in the project, project management strategies, coordination mechanisms, mentoring activities for the post doctoral researcher, outreach activities and the dissemination plan.

# Personnel Roles: The expertise and experience of the PIs with respect to this project, as well as their roles and responsibilities in the project, are as follows:

The Lead PI, ***Dr. Sethuraman Panchanathan*** is the director of the Center for Cognitive Ubiquitous Computing at ASU. He has extensive experience in the area of human-centered multimedia computing. His primary research interests are in the areas of computer vision, image and video processing, face/gait analysis and recognition, haptic and tactile user interfaces, and ubiquitous computing environments for people with sensory/perceptual/cognitive disabilities. He has a successful history as PI and Co-PI of research projects of all sizes with different funding agencies, and has published over 300 journals and conference papers. He was the lead PI of the NSF-ITR project focused on the design of assistive devices and technologies for enhancing educational opportunities for individuals who are blind and visually impaired. This project won the Governor’s Innovator of the Year Award in Arizona in 2004. Dr. Panchanathan has years of experience in working on large interdisciplinary teams, including successful collaborations with clinical partners such as Mayo Clinic. As the PI of this project, Dr. Panchanathan will provide leadership, mentor the Postdoctoral Researcher, supervise PhD students, execute the design and development of the Social Interaction Assistant (Research Objective 2) component, and provide the computing expertise in all aspects of the project.

***Dr. Michelle Shiota*,** Co-PI**,** is an Assistant Professor of Social Psychology in the Department of Psychology at Arizona State University. She specializes in the area of emotion and social bonding, integrating psychophysiology, behavioral, cognitive, narrative, and self-report measures. Dr. Shiota has published fundamental work in leading journals such as *Emotion* and *Cognition and Emotion*, including a textbook, *Emotion*, published by Thompson Wadsworth. Dr. Shiota will supervise the Ph.D. student in Psychology identifying the key non-verbal cues in social interactions (Research Objective 1), and assist with both formative and summative evaluations of the project.

***Philip White***, Co-PI, is an Assistant Professor of Industrial Design in the College of Design Innovation and the Global School of Sustainability at Arizona State University and is also a principal of Orb Analysis for Design in Phoenix, Arizona. He is the primary author of the “Okala Design Guide”, a curricular text used in the majority of schools of industrial design in North America. He has developed the Okala Impact Factors, and continues to refine these factors for designing more environmentally friendly systems. He is a contributor to the United Nations Environmental Program Lifecycle Initiative. White will co-supervise a PhD student in modeling sustainability and human health factors for existing telecommunication technologies and thereby building a baseline for comparison of any technology developed as a part of the e-REACH project.

***Dr. Gaurav N. Pradhan*,** is a Postdoctoral Research Associate at Arizona State University, under the mentorship of the PI, Dr. Panchanathan. His primary research interests are assistive technologies, pattern recognition, feature extraction in multimedia data, data integration, data mining and analysis, and human-centered multimedia computing. Dr. Pradhan will assist the PI in all stages of design and development of this project, co-supervise the computer science students. He will co-ordinate all data collection efforts, manage archival of data collections, and actively lead the dissemination efforts of this project.

A ***Research Engineer*** will be hired on this project for developing prototypes of any e-REACH technology designed. These prototypes will be used by the psychology team led by Shiota to determine the efficacy of e-REACH technologies in enhancing remote research collaborations. Further, the research engineer will coordinate with the design team led by White towards choosing ergonomic designs for various sensor and actuator technologies. The engineer will work bridge the technology development across the three teams. The research engineer will possess interdisciplinary expertise in computer science/electrical engineering and usability engineering.

***Ph.D. Students*** (3) will conduct focused research in the collaborative areas of computer science, social psychology and sustainability engineering towards designing the various functional components. The students will work with the three PIs, the Postdoctoral Researcher and the research engineer to carry out all research necessary for the design and development of e-REACH technologies.

**Project Management:** The PI (Panchanathan) and Co-PIs (Shiota and White) of this CDI project have past experience working with diverse interdisciplinary teams on various long term and short term projects. Primary project management responsibilities will reside with the lead PI (Panchanathan) who will also coordinate primary decision-making and will have overall responsibility on budgetary issues. All data generated from the proposed work and results obtained will be shared by the PIs. The PIs and key personnel will communicate to one another through scheduled and ad-hoc meetings, phone calls and emails on all aspects of the project. Publication authorship will be based on the relative scientific contributions of the PIs and key personnel. Efforts will be made to identify interdisciplinary dissemination means and any intellectual property related to the proposed work will be shared between the PIs and the key personnel. If a potential conflict develops, the PIs shall meet and attempt to resolve the dispute. If they fail to resolve the dispute, the disagreement shall be referred to an arbitration committee within the university.

Project management will also be carried out through the use of new technologies and software tools that allow PIs to track the progress of their counterparts and the entire project as a whole. We have identified the purchase of such software tools and associated training involved in the budget with an annual spending cost of approximately $2000 for the first three years. Graduate students, post doc researcher and the research engineer will be trained in effective use of these tools so that periodic updates become an essential part of their everyday work. This will allow the PIs to coordinate their own team and also assess their relative progress with respect to the entire project. Periodic evaluations between PIs will be done to ensure the progress in all three research fronts are according to set time schedules. This CDI project being research centric could result in unexpected delays in certain aspects of the project. Necessary actions will be taken to ensure that goals set by the PI and their graduate students are not impacted by these unforeseen hindrances.

**Coordination Mechanisms:** Coordination will be achieved through monthly meetings and scheduled telephone conversations between the PIs. To facilitate coordination across all of the other personnel, we will rely on many traditional project management techniques, including monthly open conference calls, jointly constructed web-based resources (Wiki, blog, and web site), annual open house and cross-disciplinary workshops, dedicated funding for student interactions between disciplines, and more frequent informal group meetings. The collaborative research discussions will serve as a test-bed of opportunity for the proposed e-REACH framework.

Joint publications will be authored by all three PIs in major conferences and journals (listed in the dissemination plan) within the respective areas of expertise. The PIs will also consider coordinating special issues in premier journals in all three areas. Ideas generated through such special issues will be used towards starting new journals and/or conferences that target interdisciplinary research highlighted in this proposal. Further, the PIs will work towards offering modules and courses at ASU that specifically highlight some of the issues in this proposal. We expect that the course materials will form the foundation for a possible monograph detailing the research findings of the e-REACH project.

We will also conduct annual workshops with the aim of educating the community on the research outcomes of the project and for motivating students to take up STEM fields of study. The participation in the workshop is expected to increase on a yearly basis. In the *first year*, we intend the workshop to be carried out within ASU between the various participating departments. Focus will be directed towards attracting undergraduate students from ASU to attend sessions where specific training about collaborative projects will be highlighted. In the future years, the workshops will be opened up for students from other institutions*. Year two and three* will focus on getting students from other institutions in the State of Arizona. This will include University of Arizona in Tucson, Northern Arizona University in Flagstaff, Mayo Clinic in Scottsdale, Banner Health Center in Mesa and Barrow Neurological Institute in Phoenix. This will provide an opportunity for the researchers of this CDI project to receive inputs from various researchers from different disciplines in science and medicine. The PI has already worked with most of these institutions at an executive level in establishing interdisciplinary centers of research at Arizona State University. We will leverage his expertise to evolve a successful outreach program. In *Year four*, we will establish an annual training summit to provide critical cross-disciplinary training to students nationwide. The summit will be a week-long course held in late summer that provides detailed description of the e-REACH project and various findings in the associated fields of human centered multimedia computing, socioemotional psychology and sustainability issues in telecommunications. The goal of the summit is to allow students with appropriate undergraduate training in computer science, social and behavioral psychology, human factors design, or sustainability engineering to enter into graduate-level work in green communication technologies for the future. For supporting these workshops, we have identified required funding within the proposed budget. The expected increase in the size of the participation is reflected in the increasing funding request for conducting these workshops.

**Mentoring Activities for Postdoctoral Researcher**: Efforts will be taken by the PI to actively involve the Postdoctoral Researcher in all stages of this project, providing extensive experience in using relevant research methods, tools, and skills for human-centered multimedia computing technologies, as well as the acquisition of knowledge of the scientific literature in related research areas. In addition, the Postdoctoral Researcher will be strongly encouraged to attend specialized training programs on grant management, laboratory management, ethics in research, communication skills, career planning, and professional development, conducted by the Center for Learning and Teaching Excellence and the Office of Research and Sponsored Projects Administration at ASU. The Postdoctoral Researcher will be required to assist the PI in coordination of various research activities and supervise the PhD students in the design and development of the functional components discussed for the e-REACH project. The PI and the Postdoctoral Researcher have already discussed each other’s commitments to ensure a productive postdoctoral experience, and in advancing the latter’s short-term and long-term career goals. The Postdoctoral Researcher will be required to prepare and present oral and poster presentations for lab meetings, department seminars, and on-campus cross-disciplinary seminars, in addition to publications in conferences and journals. The PI will also require the Postdoctoral Researcher to formulate his own research agenda and will encourage him to actively collaborate with other faculty mentors outside CUbiC@ASU.

**Outreach**: This CDI project will increase the awareness of green and sustainability thinking within the research and scientific community. We believe strongly that green practices are yet to permeate the research community and we identify air travel for remote collaborations as one such major green house gas contributor. The e-REACH project will be the first of many projects that will alter the ways in which researchers and scientists work in the future. On a broader scale, this project will prepare a new generation of scientists and researchers who think green in their professional careers. Our outreach will focus on (1) providing appropriate role models for women and underrepresented minorities (Dr. Shiota, a leading researcher in the area of Social Psychology), (2) demonstration of the significance of the development of green communication technologies as a means of exciting K-12 students in STEM, and (3) early integration of undergraduates into research experiences – CUbiC has been instrumental in getting undergraduates interested in open research problems through capstone projects, undergraduate research initiatives, hourly jobs and credit based independent studies. Working with CUbiC, undergraduates have also been successful in publishing papers in conferences and journals.

# Dissemination Plan: The progress on each phase of the project will be regularly updated on our website ([http://cubic.asu.edu](http://cubic.asu.edu/)) with appropriate descriptions and links to publications. In the past, we have collaborated with students from the Mass Media Communication department at ASU to create video news articles and bulletins, which will be continued to publicize the e-REACH efforts. Researchers at CUbiC have also been instrumental in organizing leading conferences (including the ACM ASSETS held in Tempe, Arizona in 2007, and the BodyNets conference at ASU in 2008), and we will continue to organize focused workshops/tutorial sessions based on the results of the e-REACH project in premier conferences. We plan to target the following premier venues to maximize the impact in each of the individual fields: (i) Computer Science: Trans. on Pattern Analysis and Machine Intelligence, Intl. Journal of Computer Vision, Trans. on Image Processing, Computer Vision and Pattern Recognition conference, Intl. Conf. on Computer Vision, Intl. Conf. on Pattern Recognition, Intl. Conf. on Image Processing, and Intl. Conf. on Machine Learning, Trans. on Computer-Human Interaction, ACM SIGCHI Conference, HCI International, World Haptics Conference, (ii) Socio-behavioral Psychology: Journal of Personality and Social Psychology, Human Communication Research, Cognition and Emotion, Journal of Positive Psychology, (iii) Design for Sustainability: Intl. Journal of the Industrial Designers Society of America, Intl. Journal of Industrial Ecology, Intl. Journal of Life Cycle Assessment, Intl. Journal of Sustainable Product Design, Intl. Journal of Sustainable Engineering, and (iv) Collaborative Computing: IEEE Trans. On Professional Communication, Journal of Collaborative Computing, Intl. Conf. on Computer Supported Cooperative Work, the Virtual Workplace.