Real-time Feedback on Nonverbal Clinical Communication

Theoretical Framework and Clinician Acceptance of Ambient Visual Design

A. L. Hartzler; R. A. Patel; M. Czerwinski; W. Pratt; A. Roseway; N. Chandrasekaran; A. Back

1 The Information School, University of Washington, Seattle, Washington, USA;
2 Department of Biomedical Informatics and Medical Education, Division of Biomedical and Health Informatics, University of Washington, Seattle, Washington, USA;
3 Microsoft Research, Redmond, Washington, USA;
4 Department of Medicine, Division of Oncology, University of Washington, Seattle, Washington, USA

Keywords
Nonverbal communication, patient-provider communication, information display, social signal processing, formative evaluation, feedback, user-computer interface

Summary
Introduction: This article is part of the Focus Theme of Methods of Information in Medicine on “Pervasive Intelligent Technologies for Health”.

Background: Effective nonverbal communication between patients and clinicians fosters both the delivery of empathic patient-centered care and positive patient outcomes. Although nonverbal skill training is a recognized need, few efforts to enhance patient-clinician communication provide visual feedback on nonverbal aspects of the clinical encounter.

Objectives: We describe a novel approach that uses social signal processing technology (SSP) to capture nonverbal cues in real time and to display ambient visual feedback on control and affiliation—two primary, yet distinct dimensions of interpersonal nonverbal communication. To examine the design and clinician acceptance of ambient visual feedback on nonverbal communication, we 1) formulated a model of relational communication to ground SSP and 2) conducted a formative user study using mixed methods to explore the design of visual feedback.

Methods: Based on a model of relational communication, we reviewed interpersonal communication research to map nonverbal cues to signals of affiliation and control evidenced in patient-clinician interaction. Corresponding with our formulation of this theoretical framework, we designed ambient real-time visualizations that reflect variations of affiliation and control. To explore clinicians’ acceptance of this visual feedback, we conducted a lab study using the Wizard-of-Oz technique to simulate system use with 16 healthcare professionals. We followed up with seven of those participants through interviews to iterate on the design with a revised visualization that addressed emergent design considerations.

Results: Ambient visual feedback on nonverbal communication provides a theoretically grounded and acceptable way to provide clinicians with awareness of their nonverbal communication style. We provide implications for the design of such visual feedback that encourages empathic patient-centered communication and include considerations of metaphor, color, size, position, and timing of feedback.

Conclusions: Ambient visual feedback from SSP holds promise as an acceptable means for facilitating empathic patient-centered nonverbal communication.

1. Introduction

Skilled bedside manner requires not only speaking clearly and avoiding jargon, but also competency in nonverbal communication. The ability to understand and convey nonverbal signals is essential to forming empathic relationships in patient-centered care [1–3]. In particular, nonverbal cues, such as voice tone, body movement, touch, and facial expression, link to important patient outcomes [4–6], such as patient satisfaction [7], adherence to medication [8], and clinical outcomes of care [9]. Nonverbal behavior is particularly effective at communicating emotional messages, establishing rapport, and facilitating patient comprehension of complex concepts in clinical settings [10–13]. These nonverbal cues often provide a more powerful signal of one’s internal state and perceptions than verbal communication alone [14–15].

Traditional clinical communication training lacks specificity on nonverbal competencies, using only general directives such as “offer empathy in greeting” [16]. Some training instructs clinicians to use touch, proxim-
A. L. Hartzler et al.: Real-time Feedback on Nonverbal Clinical Communication

Methods Inf Med 5/2014

© Schattauer 2014

We address the following research questions through a model formulation and formative user study using mixed-methods. We reviewed research on nonverbal communication to formulate a theoretical model that maps nonverbal cues to relational signals. We also designed ambient real-time visualizations that reflect variations of affiliation and control as relational signals in nonverbal patient-clinician communication. To examine clinician acceptance of this visual feedback, we conducted a lab study using the Wizard-of-Oz approach [24–28] with 16 healthcare professionals. Based on formative feedback obtained, we subsequently followed up using interviews with 7 of those health professionals to obtain additional input on a revised visual design. Findings from this two-part formative user study generate implications to consider when designing visual feedback for clinicians regarding their nonverbal communication skills.

2. Related Work

Social signal processing technologies (SSP) automatically sense and interpret social signals – complex aggregates of behaviors through which individuals express their attitudes towards other participants in the current social context [29, 30]. Thus, SSP can detect nonverbal behavioral cues associated with communicative signals – such as empathy – in social interactions. Researchers have developed SSP systems that use machine learning to identify nonverbal cues. For example, Pentland’s “sociometer badge” consists of a digital sensor that classifies nonverbal cues (e.g., physical proximity, pitch variation) into one of four “honest signals”, including activity, consistency, influence, or mimicry [15]. Mimicry, for example, occurs when an individual matches the pitch and speech rate of a conversational partner. Byun and colleagues augmented a videoconferencing system to detect nonverbal cues without physical sensors and successfully classified those cues into honest signals [22]. We extend this work by grounding our nonverbal feedback in studies of relational communication [31–33], particularly between clinicians and patients [10, 23, 34–37].

Some researchers focus specifically on the application of sensors and other SSP techniques to capture and analyze social signals [29, 30], but few focus on visually displaying those signals as reflective feedback to influence behavior. For example, a number of studies apply conversational analysis to model small group interaction [38], while other researchers have used SSP to influence interactions in small groups through both public and private displays of behavior-based feedback. Using the “Meeting Mediator” system, Kim and colleagues captured nonverbal cues with sociometer badges and visualized this data on mobile phones to reduce the amount of overlapping speaking time and to increase the interaction in team meetings [39]. Other researchers have used peripheral displays to promote awareness of both speaking time and eye gazing in small groups to encourage collaboration [40–42]. Vocal features, body language cues, eye gaze, and other cues have fed into such automated feedback displays. Evaluations indicate that real-time feedback on speaking activity can balance participation among meeting participants. However, none of these projects explored design implications for the visual elements of the feedback in depth, nor has work explored such displays in the context of clinical patient encounters.

In the patient-clinician communication literature, researchers have historically used self-report and observational methods to assess empathy and nonverbal communication. For example, one approach relies on the patient to rate how empathic the clinician was during an encounter [43]. Alternatively, observational methods involve third-party observers watching videos of clinical encounters. These methods include labeling nonverbal behaviors (e.g., talk time, speech rate, nodding, gesture) [44] or assigning utterances to particular types of socio-emotional or task-oriented exchanges [45]. However, labeling

RQ1: How do nonverbal cues map to relational signals of control and affiliation?
RQ2: Is real-time ambient visual feedback on nonverbal communication acceptable to clinicians?
RQ3: What are design considerations of such ambient visual feedback developed for clinicians?

We answered these research questions through a model formulation and formative user study using mixed-methods. We reviewed research on nonverbal communication to formulate a theoretical model that maps nonverbal cues to relational signals. We also designed ambient real-time visualizations that reflect variations of affiliation and control as relational signals in nonverbal patient-clinician communication. To examine clinician acceptance of this visual feedback, we conducted a lab study using the Wizard-of-Oz approach [24–28] with 16 healthcare professionals. Based on formative feedback obtained, we subsequently followed up using interviews with 7 of those health professionals to obtain additional input on a revised visual design. Findings from this two-part formative user study generate implications to consider when designing visual feedback for clinicians regarding their nonverbal communication skills.

2. Related Work

Social signal processing technologies (SSP) automatically sense and interpret social signals – complex aggregates of behaviors through which individuals express their attitudes towards other participants in the current social context [29, 30]. Thus, SSP can detect nonverbal behavioral cues associated with communicative signals – such as empathy – in social interactions. Researchers have developed SSP systems that use machine learning to identify nonverbal cues. For example, Pentland’s “sociometer badge” consists of a digital sensor that classifies nonverbal cues (e.g., physical proximity, pitch variation) into one of four “honest signals”, including activity, consistency, influence, or mimicry [15]. Mimicry, for example, occurs when an individual matches the pitch and speech rate of a conversational partner. Byun and colleagues augmented a videoconferencing system to detect nonverbal cues without physical sensors and successfully classified those cues into honest signals [22]. We extend this work by grounding our nonverbal feedback in studies of relational communication [31–33], particularly between clinicians and patients [10, 23, 34–37].

Some researchers focus specifically on the application of sensors and other SSP techniques to capture and analyze social signals [29, 30], but few focus on visually displaying those signals as reflective feedback to influence behavior. For example, a number of studies apply conversational analysis to model small group interaction [38], while other researchers have used SSP to influence interactions in small groups through both public and private displays of behavior-based feedback. Using the “Meeting Mediator” system, Kim and colleagues captured nonverbal cues with sociometer badges and visualized this data on mobile phones to reduce the amount of overlapping speaking time and to increase the interaction in team meetings [39]. Other researchers have used peripheral displays to promote awareness of both speaking time and eye gazing in small groups to encourage collaboration [40–42]. Vocal features, body language cues, eye gaze, and other cues have fed into such automated feedback displays. Evaluations indicate that real-time feedback on speaking activity can balance participation among meeting participants. However, none of these projects explored design implications for the visual elements of the feedback in depth, nor has work explored such displays in the context of clinical patient encounters.

In the patient-clinician communication literature, researchers have historically used self-report and observational methods to assess empathy and nonverbal communication. For example, one approach relies on the patient to rate how empathic the clinician was during an encounter [43]. Alternatively, observational methods involve third-party observers watching videos of clinical encounters. These methods include labeling nonverbal behaviors (e.g., talk time, speech rate, nodding, gesture) [44] or assigning utterances to particular types of socio-emotional or task-oriented exchanges [45]. However, labeling...
cues requires considerable time and training. Thus, researchers might label only a "thin slice" of the clinic encounter and attempt to predict clinicians' overall communication [44, 46]. With SSP, we can move beyond the limitations of self-report questionnaires, ratings that require trained observers during a clinic encounter, and labor-intensive coding processes.

Additionally, SSP is being explored in clinical settings to fulfill a range of needs. For example, a text analysis tool called "Discursus" processes transcripts from patient encounters to produce a visual account of the temporal content and turn-taking dynamics for later use in clinician communication training [47]. Another example is the web-based training system "Star Workshop" that encourages role-play practice for medical students through scenarios with illustrative patients. The system processes text communication from these online patient-clinician simulations in real-time and provides the user with tailored feedback on their communication skills from an automated coach [48].

Griol and colleagues extend these text-based approaches by incorporating context-aware conversational agents that recognize the emotional state of users [49]. In the context of adaptive home monitoring systems for patients with chronic pulmonary diseases, a conversational agent greets the patient and provides personalized feedback and support based on the patient's oxygen level, heart rate, and blood pressure. This context-aware agent is also tuned to recognize the patient's emotional state (e.g., anger, boredom) and adjust its interaction in response. Context-aware systems have also been examined for sensing and monitoring in hospitals, where clinicians show favorable reception to awareness displays that improve their coordination and their ability to locate work resources [50]. Other systems employ ambient displays for patient monitoring that alert nurses to changes in patient status [51]. Surprisingly little prior work has applied similar contextual feedback displays to patient-clinician communication during clinical patient encounters.

Other researchers have examined nonverbal communication patterns through the analysis of video captured from patient encounters. For example, Montague and colleagues [52] examined gaze patterns between patients and clinicians. In addition, Stepanikova and colleagues [53] examined the role of race in patient-doctor communication. To compare communication patterns with and without the aid of English interpreters, Weibel and colleagues [54] captured speech and body movements of clinicians and patients with Microsoft Kinect. For these studies, researchers analyzed data collected from patient encounters after the fact to enhance our understanding of clinical interaction, promote nonverbal skills, and inform tools to facilitate communication [55]. We expand upon this work by adding real-time ambient feedback of communication patterns during patient encounters.

### 3. Relational Communication Framework

Reflecting real-time visual feedback on communication requires that we first conceptualize the mapping of nonverbal cues to relational signals associated with empathic, patient-centered care. To provide this grounding, we devised a mapping framework based upon a validated model of relational communication that has been applied in clinical settings [34–35] and was derived more generally from interpersonal communication research [31–33]. The model suggests that interpersonal communication reflects intimacy (i.e., extent of immediacy and affection, trusted receptivity, and depth of similarity), dominance (i.e., extent of equality, assertiveness, control, and influence), composure (i.e., extent of feeling calm and poised versus tense), and formality (i.e., extent of formal versus informal interaction) [34–35]. Of these dimensions, prior research points to two central types of relational signaling – affiliation and control – that comprise core dimensions of nonverbal communication in both general settings [14, 56, 57] and medical contexts [10, 23, 36, 37].

Mapping to intimacy, affiliation reflects interpersonal warmth, trust, and rapport. Mapping to dominance, control reflects inequality, influence, and authority. These relational signals show polarity and can vary in strength along a continuum. Therefore, we focused on these two most important dimensions in our designs.

Based on these dimensions, we developed a framework to map types of nonverbal cues to the two core types of relational signals – affiliation and control (Table 1). Each mapping represents supporting research evidence that associates a given nonverbal cue and relational signal. Although the most relevant evidence comes from studies on patient-clinician interaction, we also provide evidence from interpersonal communication research generally. For each mapping, we note the direction of the association (i.e., positive or negative), the research setting from which evidence is drawn (i.e., clinical or general), and favorability of the associated outcome (i.e., favorable, unfavorable, mixed – reflecting studies with both favorable and unfavorable results). These theoretical underpinnings point to a wide range of nonverbal cues that SSP can target. We next detail this evidence base of relational signals and nonverbal cues that underlies our relational communication framework.

### 3.1 Relational Signals

The core relational signals of communicative affiliation and control are strong predictors of patient satisfaction with health care [7]. In particular, patients report greater satisfaction when clinicians use affiliative and less domineering communication [23, 34]. Although the importance of clinicians' communication in patients' evaluations of care can vary with illness severity, age, specialty, and number of prior visits [23], patients generally prefer affiliative clinicians who are attentive, responsive, and establish rapport, while patients disapprove of clinicians who are overly domineering and argumentative [37]. Clinicians often control interactions more than patients by producing more questions, interruptions, and topic initiations and by talking for a greater proportion of time [10, 37]. Although patients and clinicians show more congruence among affiliative than controlling behaviors, patients tend to tolerate greater communicative control with clinicians who are open and engaging [37]. Exercising some communicative control is linked to patient adherence [10, 58], but patients are...
### Table 1  Relationship between nonverbal cues and relational signals of affiliation and control (*used in lab study)

<table>
<thead>
<tr>
<th>Nonverbal Cue</th>
<th>Affiliation</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rapport</td>
<td>Trust</td>
</tr>
<tr>
<td><strong>Haptics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social touch (polite, friendly, ritualistic e.g., handshake)</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>Instrumental touch (task-oriented e.g., physical exam)</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Self-touch (e.g., touch own face, rub hands together)</td>
<td>+</td>
<td>(-)</td>
</tr>
<tr>
<td>Comfort touch (reassuring e.g., hug or pat on shoulder)</td>
<td>(-)</td>
<td>(+)</td>
</tr>
<tr>
<td>Exclamatory touch (encouraging or complimenting pat)</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>Persuasive touch (influencing pat to gain attention)</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>Accidental touch (e.g., bumping into another person)</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>Touch avoidance</td>
<td>(-)</td>
<td>(-)</td>
</tr>
<tr>
<td>Excessive touch</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Kinesics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mouth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smile</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Lip licking</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lip pursing</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>Eyes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Making eye contact</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Increase in eye contact/gaze by patient*</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Moderate (not excessive) increase in gaze by clinician*</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Mutual gaze</td>
<td>-</td>
<td>(+)</td>
</tr>
<tr>
<td>Looking at others while speaking</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>Looking at others while listening</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shifting gaze</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>Eye blinks</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>Glare</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td><strong>Brow</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyebrow raise</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>Eyebrow lowered</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td><strong>Face</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relaxed facial expression</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Pleasant facial expression</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Angry facial expression</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Positive facial expressivity</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td><strong>Head</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Head nodding*</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Head shaking*</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Fluid head turns</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td><strong>Trunk</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward lean</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Sideways lean</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Shoulder shrug</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td><strong>Hands &amp; limbs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Active gesturing</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>More active gesturing by clinician than patient</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Reciprocal gesturing</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Expansive gesturing</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Illustrator gestures</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Expressive gesture</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Open hand and arm position</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Limb symmetry (side-by-side)</td>
<td>(+)</td>
<td>(+)</td>
</tr>
</tbody>
</table>
### Table 1  Continued

<table>
<thead>
<tr>
<th>Nonverbal Cue</th>
<th>Affiliation</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Rapport</td>
<td>Trust</td>
</tr>
<tr>
<td>Kinesics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Posture &amp; body</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Open body position/posture</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Erect posture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relaxed posture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Still posture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Postural sway</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vigorous and fluid body movement and shifts</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Close proximal distance</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>Proxemics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Direct body orientation (vs. indirect “cold shoulder”)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Invasion of space</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocalics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Friendly vocal tones</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Concerned vocal tones</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Angry vocal tones</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Relaxed resonant and rhythmic vocal tones</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Passive voice</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faster louder speech*</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>More vocally active clinician than patient*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Vocal variety (e.g., volume, pitch, intonation, tempo)*</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>Lowered pitch*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Increased pitch*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Response latencies</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Close listening</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pauses while speaking</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>Silence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocal fluency</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>Coordinated speech</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Balanced talk time*</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Balanced turn taking</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Matched speech latencies</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Vocal back channeling</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Shared talk time*</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Increased talk time*</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Interruptions*</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Greater conversational control</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

*+: Positive/ negative relationship in clinical context; (+/-): positive/negative relationship in general context; ■■■■: favorable outcome; ■■■■: unfavorable outcome; ■■■■: mixed outcome with both favorable and unfavorable aspects
less satisfied with highly controlling styles [23]. Thus, affiliation and control are two key factors that influence the quality of patient-clinician communication. Next, we describe how our framework delineates each of these core relational signals.

### 3.1.1 Affiliation

Affiliative communication reflects interpersonal warmth, trust, and rapport. Report refers to attentive interest, engaging approachability, and conversational involvement through willingness to listen and responsiveness. Trust refers to openness and feeling at ease that encourages disclosure. Warmth refers to immediacy and closeness, friendly affect, positivity, and likability, as well as empathic comfort and emotional caring. There is a strong positive relationship between perceptions of clinician affiliation and patient satisfaction [37]. For example, clinician warmth and listening are associated with patient satisfaction [4]. An affiliative style helps to reduce anxiety and facilitate patient disclosure [23]. More broadly, when clinicians show more nonverbal immediacy patients have greater trust and satisfaction, more disclosure, understanding, and retention of medical information, and are more motivated to follow medical advice [59].

### 3.1.2 Control

Controlling communication reflects dominance, influence, and authority. Dominance refers to energetic enthusiasm, a take-charge attitude, conversational management that engenders respect, and even aggression.

Influence refers to power, and persuasive capacity to gain attention and impact the behavior of others. Authority refers to poised self-confidence, charismatic panache, credibility, responsibility, and expert competence impressed by status markers. Because of the medical expertise that clinicians hold, patients and clinicians generally differ in power, and this difference is typically reflected in their nonverbal communication patterns [13, 37]. Clinicians tend to exercise more control in consultations than patients. For example, clinicians communicate social power by having unquestioned access to the patient’s body, controlling communicative exchanges, and regulating the degree of interaction and involvement on their “turf” [10]. These differences are reinforced by status differentials, such as dress (e.g., clinicians in white coats and patients in gowns) [13]. Although highly controlling styles have shown a negative relationship with patient satisfaction [23], exercising some communicative control is linked to patient adherence [10, 58].

### 3.2 Nonverbal Cues

Nonverbal communication cues can be divided into four major categories: *haptics* – touch, such as a handshake or pat on the shoulder; *kinetics* – body movements, such as facial expression and gesture; *proxemics* – spatial cues, such as body orientation and conversational distance; and *vocalics* – non-linguistic vocal cues, such as speech tempo and pitch) [14]. Combinations of these nonverbal cues are powerful indicators associated with desirable patient-clinician communication. For example, direct body orientation (i.e., a proxemic cue) and minimizing talk time (i.e., a vocalic cue) indicate involvement and approachability, and a smile (i.e., a kinesic cue) indicates warmth and likability [10, 14]. Reciprocating nonverbal cues through mimicry and matching further builds rapport and increased trust [10, 15]. In contrast, clinicians’ indirect body orientation, frequent interrupts, and long speaking turns are associated with a dominating influence that inhibits patient participation [10, 60].

Our relational communication framework lists numerous examples of nonverbal cues for each category that are prevalent in clinical settings as well as their mappings to the core relational signals of affiliation and control based on the literature (Table 1). Such cues can be processed by SSP to capture and provide real-time visual feedback about communicative signals that can encourage empathic patient-centered communication. We next walk through the evidence base of nonverbal cues that map to relational signals of affiliation and control.

### 3.2.1 Haptics: Communicating through Touch

Communicating through touch, referred to as haptics, helps to make interactions engaging and to promote health and well-being [61, 62]. Yet, touch can lead to either positive or negative outcomes depending upon the form (e.g., intensity, duration, frequency) and type (e.g., social touch, persuasive touch, task touch, self touch). Touch that is friendly, comforting, exclamationary, or even accidental can signal immediacy and warmth [63]. However, physicians often exert dominance and authority by using more social touch than patients do [37]. For example, comfort touch (e.g., pat on the shoulder) can indicate emotional support and sympathy, but can also be used persuasively to influence and signal control [14]. Physicians’ use of task touch (e.g., touching patient as part of examination procedures) is associated with patients’ perceptions of dominance [37], yet patients can perceive excessive use of touch as non-affiliative and aggressive [37, 64]. Although use of self touch outside the clinical context may be a signal of deceit [63], use of self touch by both patients and clinicians has been associated with expressive and warm interaction, [65] suggesting they are paying close attention and building rapport [66]. In contrast, touch avoidance is often negatively associated with immediacy, trust, and rapport [63].

### 3.2.2 Kinetics: Communicating through Body Movement

Body movements, referred to as kinetics, comprise the most influential nonverbal cues revealed through movements of the face, head, trunk, hands and limbs, and through bodily posture.

*Facial movements* comprise a dense source of nonverbal cues by revealing expressivity from mouth to brow. Signaling warmth, likeability and trust, [14] and patient desire for involvement [13], smiling is associated with attentive patient-clinician rapport [67, 68]. A pleasant and smiling expression signals openness, relaxation, and authoritative poise [14]; it positively reinforces affiliation between patients and clinicians [10]. Yet, smiling is nuanced – a
wide-open smile can signal authoritative poise and an appeasing smile can signal submissiveness [14]. In contrast, lack of smiling could reflect social influence or coercion, and an angry facial expression signals intimidating threat [14]. People with greater facial relaxation are rated as having more expert power and influence [69], whereas people with less power status tend to mirror the amount of relaxation shown by those with higher rank [14]. Even lip pursing and lack of facial movement are associated with deception. In addition, lip licking signals tension and lack of poise [63].

**Eye movements** also provide strong relational signals. Making eye contact signals patients’ desire for involvement [13] and signals clinicians’ nonverbal immediacy and attentiveness [10, 70]. Clinicians who gaze into patients’ faces show affiliation [10], which leads to increased patient satisfaction [60, 64]. In contrast, patients who make little eye contact could feel anxious [68] and lack involvement in their care [13]. Patient disclosure is less likely when clinicians break or use indirect eye contact while speaking [67]. Similarly, increased gaze generally signals affiliation [10, 14], whereas distancing gaze signals reduced patient-clinician rapport [71]. However, researchers found that increased patient gaze toward clinicians correlated with reduced immediacy [64]. Other researchers found that physicians rated high for rapport show moderate but less extensive eye contact than physicians rated low for rapport [66]. Concurrently, gaining and maintaining eye contact while speaking can also signal dominance and influence [14]. For example, powerful people tend to look more at others while speaking than they do while listening. In addition, less powerful people look at superiors while listening to signal respect [14]. High levels of maintained eye contact with listeners is generally associated with trustworthiness, but can also signal authoritarianiveness [14]. Mutual gaze in which partners look at one another ‘eye-to-eye’ can be an affiliative cue of interactional synchrony that helps even unbalanced power [37, 63]. But mutual gaze can also signal domineering clinician control [10, 66, 67, 72], perhaps owing to a limit on clinician involvement above which clinicians may be unfavorably evaluated [10]. Harrigan and colleagues [66, 72] offer support for such nuanced limitations in their finding that clinicians who gaze more at patients, but engage in less mutual gaze, establish more rapport with patients. Further, shifting eyes can signal lack of composure and poise [14], eye blinks can signal deceit [63], and a steady glare signals an intimidating threat [14]. Even brow movements provide relational signals – a raised brow reflects authoritative charisma whereas a lowered brow tends to communicate coercive influence and intimidation [14].

**Head movements** serve as one of the most salient nonverbal cues in patient-clinician interactions. In general, frequent head nodding reinforces affiliation [14]. Head nodding by clinicians is associated with warm empathy and rapport [10, 68, 72] that encourages patient involvement [70] and disclosure [67], whereas head nodding by patients signals desire for involvement [13]. Head nodding also reflects authoritative panache and conversational control [14]. Head shaking is associated with energetic authority [14] and could indicate rapport [66]. Finally, head turns that are fluid rather than stiff are associated with charismatic authority [14].

**Trunk movements** – such as a forward lean toward a conversational partner – provide a common signal of warm immediacy and trust [14, 63, 67] and are one of the most indicative cues to patient-clinician affiliation [68, 70, 72]. A lean of the patient toward the clinician signals desire for involvement [13]. When met with resistance, patients might lean further toward the clinician to bind them to the encounter or signal non-adherence [13]. In contrast, a backward lean of the clinician away from the patient can reduce patient satisfaction and understanding, both cues to lower affiliation [64]. Both forward and sideways leaning is associated with poised self-assurance and dynamism, signal relaxation and authority [14]. Even shoulder shrugs could link to negative relational signals, such as deceit [63].

**Hand and limb movements**, while serving as emblematic social rituals (e.g., wave goodbye), provide a primary means of nonverbal communication. For example, illustrator gestures (e.g., pointing to emphasize spoken words) are indicative of interest, expressiveness, and likability [14]. Active gesturing generally signals affiliation, whereas expansive gesturing that fills space signals relaxed charismatic authority.

In addition, expressive gesturing signals authoritative likability [14]. In patient-clinician interaction, gesture has been linked to warm and expressive likability [65] and helps to build rapport as well as serve as a sign of nonverbal encouragement to patients [70]. Clinicians’ use of gesture could even enhance patient memory for medical information [10]. Furthermore, open hand and arm positions can signal trustworthiness [14] and warm clinician empathy [72]. Whereas open, symmetrical arms signal clinician rapport [66], asymmetrical arms and legs signal authoritative- ness [14]. Clinicians and patients who reciprocrate gestures signal affiliation, yet clinicians who are more gesturally active than patients are perceived as less affiliative [37].

**Body movements and posture** also signal relational messages. For example, postural openness is associated with affiliative communication between patients and clinicians [10]. Body relaxation can also signal warm immediacy [63], whereas erect posture, open body position, and vigorous yet fluid body shifts can signal energetic and authoritative charisma [14]. Similarly, vigorous but controlled body movement shows kinesic expressivity and reflects rapport [63]. In contrast, a closed and defensive body position gives a communication partner more space and signals lack of power [14], submissiveness [13], and lack of empathy [65]. Postural stillness can signal lack of trust, whereas postural sway can signal tension and lack of poise [14]. Postural sway, including swiveling, can indicate lack of composure [14] and even deceit when sway is exaggerated through body rocking and twisting [63].

### 3.2.3 Proxemics: Communicating through Space and Distance

Proxemics refers to communication through the perception, use, or structuring of space [14]. Space helps people maintain privacy and shield themselves from emo-
physical discomfort, harm, and threat. For example, close conversational distance is common in intimate interactions, whereas increased distance provides greater ‘personal space’. Close proximal distance, in which personal space is shared, is associated with warm, immediate, and trusting relationships [14], including patient-clinician relationships [10, 70]. Patients report greater satisfaction when clinicians sit close to them [73] and lean toward them [64]. In contrast, increased proximal distance signals dominating conversational control but reduced levels of authority and influence [14]. Similarly, invasion of personal space by leaning toward a conversational partner while the partner leans away signals dominance [14].

Body orientation also contributes to relational signaling. Direct body positioning is generally associated with approachability [63], authoritativeness, and social influence [14]. For example, when clinicians face patients directly, patients have greater immediacy, trust, and satisfaction with care [10, 14]. Seeing eye-to-eye on the same physical plane as a conversational partner can also signal immediacy [63]. Direct body orientation is associated with clinician immediacy [64] and rapport [65, 66, 68, 72], whereas indirect body position is associated with patients’ perception of clinician dominance [37] and lack of rapport [66, 72, 74].

### 3.2.4 Vocalics: Communicating through Vocal Cues

Vocalics refer to non-linguistic speech that provides significant communicative power through variations in intonation, volume, tempo, pitch, and speech patterns, all of which help to shape communication style [14].

Vocal expressivity signals rapport [63], patient desire for involvement [13], and persuasive influence [14]. In general, vocal variety is a cue to trustworthiness [14] that increases attention and rapport to enhance patient comprehension [10]. Warm, friendly vocal tones signal immediacy, trust, and rapport in patient-clinician interaction [14, 67]. In addition, a caring tone signals warm concern for the patient [10, 68]. Resonant and rhythmic voices with relaxed laughter signal rapport [10] and authority [14]. In contrast, an angry voice reflects social influence to persuade patients to follow medical advice [75]. Patients who use a passive voice-tone with a closed body position and make little eye contact signal signal submissiveness [13]. Louder and faster speech is nuanced, sometimes signaling patient-clinician rapport and involvement [10], and other times signaling authority and influence, or even intimidation [14]. For example, clinicians who are more vocally active than patients are perceived as less affiliative [37]. In contrast, not talking too much and listening closely by granting patients the floor helps to build rapport [11]. Similarly, response latencies, in which time elapses before speaking, signal clinician affiliation [37, 60]. Vocal signals of patient involvement and rapport range from lower pitch to greater pitch variety and relaxed laughter [10]. Lower pitch is associated with rapport [10] whereas increased pitch can signal deceit [63]. Verbal fluency corresponds with open expressiveness, positive affect, conversational control, and relaxed poise, whereas non-fluencies (e.g., stutters, choppiness, “er”, “um”) can signal reduced trustworthiness [63] and lack of composure and influence [14].

Speech patterns, such as vocal synchrony, including coordinated speech, comparable turn durations, matched speech latencies and vocal backchannels (e.g., “uh hhm” to indicate listening) are associated with less domineering and more responsive patient-clinician interactions [10]. Although shared talk time is a strong sign of warm immediacy and balanced influence [63], increased talk time of a conversational partner is a sign of poised authority and dominance [14]. Longer speaking turns, more pauses while speaking, more interruptions, and greater talk time signal dominance and authoritative leadership [14], all of which are associated with clinician dominance [37] and less patient speech time [60]. Clinicians who keep tight conversational control and do not let patients speak exhibit this controlling style, referred to as ‘clinical hypocompetence’ [5]. Although pauses can encourage rapport by providing conversational acknowledgment, such as granting the floor [14], fewer pauses and response latencies have been associated with patient-clinician involvement and rapport [10]. Furthermore, too much silence can signal intimidating threat, lack of acknowledgement, and reinforce power [14].

### 3.2.5 Other Nonverbal Communication Cues

Other nonverbal cues, such as physiological readings, have received less attention in past research. Yet, advances in biosensing technology make it possible to capture heart rate, blood pressure, or even skin response as cues in relational communication [76]. Such cues may be closely tied to the relational signal of composure, which reflects the extent of feeling calm and relaxed versus anxious and tense in patient-clinician interaction [34, 35]. Other nonverbal cues, such as attire [77], may also provide insights into patient-clinician communication (e.g., clinician’s white coat signals authority and formality) [13].

### 4. Methods

Grounded in our theoretical framework of relational communication, we designed ambient real-time visualizations that reflect variations in affiliation and control as relational signals in nonverbal patient-clinician communication. To examine clinician acceptance of this visual feedback, we conducted a two-part formative user study of clinician acceptance comprised of a lab study (part 1) and follow up interviews (part 2) with health professionals. Next, we describe the procedures and results from each part of this mixed-methods study.

### 5. Part 1: Lab Study

We explored the acceptability and design considerations of ambient, real-time, nonverbal communication feedback for clinicians. We conducted a lab study using “Wizard of Oz” [26] to simulate system use and focused data collection on both real-time and conceptual properties of our initial design of visual feedback. Wizard of Oz is a rapid-prototyping method that simulates system responses in situations...
where the system is complex and costly to build [24–28]. A primary motivation for using this approach is to anticipate design issues through user testing during early stages of system development and then iteratively refine the user interface before committing resources to fully implement the underlying technology. Wizard of Oz is typically conducted in an experimental setting in which participants are given the impression that they are interacting with a fully functioning system. The researcher acts as a “wizard” to surreptitiously intercept the participant’s response and feed this input to the system. Given the breadth of nonverbal cues from our theoretical framework and costs of classifying those cues for visual feedback, we chose Wizard of Oz as a first step to examine a subset of nonverbal cues, which the researcher (RP) intercepted to manually adjust the visual display. The kinds of nonverbal cues the researcher focused on during the lab study are marked with asterisks in Figure 1, including head movements and gaze, variations in pitch and tempo, relative talk time, and interruptions. Through this simulation, we sought emerging design considerations to drive our iterative refinement of Entendre’s ambient visual feedback.

In the lab study, we employed a role-play scenario with healthcare professionals to gather their perceptions and to determine acceptance of our initial design of the reflective visual display. The task for each participant was to be as empathic as possible with an actress playing the role of a patient (“Alicia”) and to incorporate feedback from the ambient visual display as much as s/he chose. Participants engaged with a professional actress who had improvisational experience simulating a patient in medical school exams. We designed a “getting to know you” counseling scenario to enable recruitment from a broad range of health professions. The counseling scenario involved meeting with a shy young woman played by the actress. Participants were free to incorporate details from their own lives. The participant and actress were videotaped as they carried out the role-play scenario in an observation room with one-way glass, through which the researcher observed and logged nonverbal cues, and then adjusted the visual feedback display in response. Participants were screened for prior counseling experience to help ensure a natural role-play setup. We conducted two pilot sessions with people from the general population to train the actress and modify the scenario to help make the task more natural.

### 5.1 Initial Visual Design

Our initial visual design juxtaposed two separate visual elements representing affiliation and control, which change in real-time (Figure 1). A sun-moon element represents changes in affiliation, becoming large and yellow to indicate high affiliation and becoming small and blue to indicate low affiliation. The second visual element, a seesaw, represents changes in control with a ball that rolls from one side labeled “You”, the clinician, to the other side labeled “Alicia”, the patient. By observing the participant’s nonverbal cues from behind the one-way glass, the researcher acted as the “wizard” to adjust the visual feedback to corresponding levels of affiliation and control on a 7-point scale. For example, when the health professional signaled affiliation, such as head nodding, the researcher increased the affiliation level of the visual display, making it more sun-like by becoming larger and more yellow in color. When the health professional signaled control, such as interrupting the patient, the researcher increased the control level by making the ball on the seesaw roll away from “Alicia”, the patient, and toward “You”, the clinician.

### 5.2 Participants

We recruited 16 healthcare professionals (P1–P16) whose specialties are listed in Table 2. Participants ranged in age from 25–55 (mean 40, median 42), and nine were men. We selected young to middle-aged healthcare professionals to obtain the perspective of those who are more used to technology in their everyday lives. All participants were compensated with a software gratuity.

#### Table 2 Participants’ professional specialties

<table>
<thead>
<tr>
<th>Professional Specialty</th>
<th>Number of Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emergency Medical Technician (EMT)</td>
<td>5</td>
</tr>
<tr>
<td>Nurse</td>
<td>4</td>
</tr>
<tr>
<td>Certified Nursing Assistant (CNA)</td>
<td>2</td>
</tr>
<tr>
<td>Speech-Language Pathologist</td>
<td>2</td>
</tr>
<tr>
<td>Physician</td>
<td>1</td>
</tr>
<tr>
<td>Dentist</td>
<td>1</td>
</tr>
<tr>
<td>Chiropractor</td>
<td>1</td>
</tr>
</tbody>
</table>
5.3 Procedures

Prior to leading participants to the observation room, the researcher played a training video and gave participants a handout that explained the meaning of the visual feedback. The handout described what the concepts of control and affiliation were, and what cues mapped to the high and low signals of each. The researcher also shared role-play instructions for the “getting to know you” counseling scenario. The actress was the same for all participants to keep the scenario experience consistent. We did not inform the actress of specific goals or expected outcomes. Finally, the researcher did not reveal that she, rather than automated SSP, would be generating the visual feedback for this study.

The observation room was set up as shown in Figure 2. Two auxiliary video cameras captured the head and shoulders of the participant and of the actress. The first 8 participants received visual feedback from a 17-inch display placed on the table two feet to the left in between both parties and tilted toward the participant. The last 8 participants received visual feedback from the display placed two feet behind the actress and slightly to the left as shown in Figure 2. In both cases, the actress could not view the display from where she was sitting.

The role-play conversation lasted 10–12 minutes, with the display showing no feedback for the initial 2½ minutes. We created this initial no-feedback period to help the participant get immediately acquainted with the actress without distractions. The researcher who controlled the visual feedback referred to live video streams of the participant, the actress, and the visual feedback display. The researcher rated the affiliation and control of participants on a 7-point scale, basing her overall judgment by focusing on cues marked in Table 1. On average, the researcher updated the continuously playing visual feedback every 6.5 seconds. We logged timestamps for changes that the researcher made to the control and affiliation levels in a database for later analysis.

After the role-play conversation was over, the researcher administered written questionnaires to both the health professional participant and actress to gauge the participant’s acceptability of the visual design and to gauge the actress’ perception of empathy exhibited by the participant. In a brief exit interview, the researcher also asked for further comments from participants about their attitude toward the technology and its impact on their interaction.

5.4 Measures and Analysis

Our mixed methods data collection included qualitative and quantitative assessments of feedback consistency, clinician acceptability of the visual feedback, and perceived impact of the visual feedback on participants’ behavior.

5.4.1 Feedback Consistency

The actress completed the Consultation and Relational Empathy Measure (CARE) with slightly modified wording for the role-play scenario. CARE is a 10-item questionnaire that measures patients’ perceptions of relational empathy in clinical consultations [43]. We used CARE to compare the actress’ perception of the participant’s empathy to scores from the researcher-controlled feedback logged in the database. We assessed the relationship between CARE empathy scores and average time-weighted control and affiliation levels using a Pearson’s correlation. Specifically, time-weighted control and affiliation scores were calculated as follows: the level (i.e., score between 1 and 7) was multiplied by the time in seconds at that level, and then all seven levels were added together and divided by the total number of seconds the participant was shown the visualization during the role-play.

5.4.2 Clinician Acceptability

The written questionnaire administered to participants captured 1) design ratings about the extent to which they perceived the visual feedback as informative, interesting, helpful, distracting, and confusing on a 5-point Likert scale, with 1 = “not at all” and 5 = “very much”, 2) likes and dislikes of the sun-moon representation of affiliation and the seesaw representation of control (open-ended qualitative response), and 3) whether in the future they would be open to using the visual feedback in a professional setting (yes/no/unsure). The verbal exit interview gave participants the opportunity to provide additional open-ended feedback on their attitudes towards acceptance of the technology. Common themes in open-ended qualitative responses emerged by grouping together similar responses across participants in a grounded manner. For quantitative data, we calculated mean Likert scores and calculated the proportion of “yes”, “no”, and “unsure” responses across participants.

5.4.3 Behavioral Impact

In the written questionnaire, we also asked participants how they thought the feedback display influenced their interaction with the actress (open-ended qualitative response) and to self-report the frequency of glancing at the visualization (1–2 times...
during the session, 3–6 times during the session, once per minute, more than once per minute). We also counted the number of times participants glanced at the visual display, which we compared to their self-reported frequency of glancing. The verbal exit interview gave participants the opportunity to provide additional feedback on the perceived impact of the feedback on their behavior. Again, common themes in open-ended qualitative responses emerged by grouping together similar responses across participants in a grounded manner.

5.5 Lab Study Results

Findings from the lab study reflect consistency between patient-reported empathy and observed nonverbal affiliation, solid signs of clinician acceptance of the visual feedback with formative design refinements for future use, and trends toward positive impact of the visual feedback on behavior.

5.5.1 Feedback Consistency

There was a positive correlation between the participants’ CARE scores and their time-weighted affiliation levels (Pearson’s r = 0.815, p < .001). However, no relationship was found between CARE scores and time-weighted control levels. Still, the actress engaged participants enough to make the role-play feel realistic for the participants, which they consistently expressed in exit interviews. This finding points to consistency between patient-reported empathy and observed nonverbal affiliation.

5.5.2 Clinician Acceptability

Design ratings: Participants’ ratings about the extent to which they found the overall visual feedback informative, interesting, helpful, distracting, and confusing are shown in Figure 3. Mean scores for the positive aspects (informative: 3.8, interesting: 3.6, helpful: 3.3) were higher than the negative aspects (distracting: 3.0, confusing: 2.3). However, these differences were not significant.

Likes and dislikes: In their open-ended qualitative responses, participants described a number of things they liked or disliked about the visual feedback, which fell into two main themes: interpretability and actionability:

Interpretability, including the ability to understand and use the visual feedback, was one major theme that emerged for participants. Five participants stated that the visualization was straightforward to understand because they resonated with the metaphor. Participants appreciated the “simplicity of the feedback” (P7) and that the display provided “an easy-to-understand concept” (P14). P13 told us that the “sun and moon are good representations” that make the feedback on affiliation “easy to quickly understand” and P10 told us that the seesaw “shows clearly the balance [in control]”. Thus, use of metaphor appears to be an important factor that supported interpretability of the visual feedback.

In contrast, other participants found the visual feedback harder to interpret, such as P3 who told us “I don’t understand all the colors.” Four participants expressed confusion about how they could make the display change, perhaps because changes in the visual feedback were not tightly coupled to changes in their behavior, as described by P2: “at times, it [the feedback] didn’t seem to reflect what was going on.” Similarly, P13 told us “there was a bit of a time delay.” Other participants raised further issues that limited their use of the feedback, such as as poor placement and distraction of the display. For example, P7 told us the monitor was “too far to the left” to be usable while still engaging in the conversation, whereas P9 thought changes in the affiliation were detectable because “I could see size change out of the corner of my eye”. Three participants said that they barely used the display because the “feedback was small” (P12) and that the feedback “could be bigger” (P10). For these participants, color, timing, placement, and size of the visual feedback were all important factors that limited their ability to understand and readily use the ambient feedback.

Actionability of the feedback, resulting in increased self-awareness to either reinforce or change one’s interpersonal behavior, was perceived as beneficial by some participants, but as challenging for many others. For example P4 told us that the display “gives visual feedback that can be quickly acted upon.” Several participants liked the self-awareness provided by the display about their communication style because: “I could tell if one of us was talking too much” (P8) or the display “made me aware if I wasn’t listening enough” (P11). P9 told us “I liked the reminder to not dominate the entire conversation” and “I liked knowing I was being warm.” Although P10 glanced at the feedback only three times, she felt that just having the display available reminded her to be warm and friendly. Participants varied in their preference for
granular updates that provided awareness of what happened in the moment to change their course versus broader checks of the interaction on the whole. Two participants found the use of metaphor particularly helpful for providing self-awareness through the visual feedback. In particular, P1 liked seeing the warm sun confirming that he was being engaging because that feedback "let me know the conversation was going okay." P11 found that the seesaw representation provided useful reminders because it "told me when to ask more personal questions to draw her out."

However, the actionability of the visual feedback came with a cost to several participants. P12 told us the feedback "made me a little nervous." P14 became self-conscious because the visual feedback made it "really hard to know what to do to control the conversation and still be natural." Five participants mentioned the display created distraction that could disrupt the interaction. For example, P16 told us "If I looked too much [at the display], it felt like I wasn't paying attention to her." P11 told us the display "made me work at keeping eye contact," and P15 told us "when I noticed it [the seesaw] was on me, I felt I needed to stop talking and not necessarily finish." P13 also felt the visual feedback could discourage him from finishing his speaking turn "as soon as I started talking the seesaw would tilt towards me being dominant." Given participants' responses about these costs, coupled with problems that emerged around understanding and use, it was clear that iterating on the visual design could improve both the interpretability and actionability of the reflective feedback.

Future use: When asked if they would be open to using the visual feedback in a professional healthcare setting, ten participants (63%) said "yes", while only 2 participants (12%) said "no", and 4 participants (25%) were "unsure". In the exit interview, several participants further validated their desire to use such a tool in the future. A physician noted the feedback display is "something I could get used to since I'm already referring to monitors and clocks [in the exam room]" (P14). A nurse confirmed that the visual feedback could be helpful because "when patients tell their story, time pressure often makes you less empathic in real situations" (P11). One certified nursing assistant was more ambivalent about the visualization, saying that it "feels so subjective" (P16). Overall, participants confirmed the potential for targeting clinicians' improvement of nonverbal communication through visual feedback.

5.5.3 Behavioral Impact

In their response to the question "How did the ambient feedback influence your interaction with the patient?" 8 participants (50%) indicated that the feedback had a positive effect on the interaction with the actress, while 3 (19%) thought that the effect was negative, and 5 (31%) reported little or no perceived effect on the interaction. Participants' responses reinforced the types of likes and dislikes raised related to metaphor, timing, placement, color, and size of the visual display and its influence on the interpretability and actionability of the ambient feedback.

Seven participants (44%) reported that they glanced at the visualization 3 – 6 times and 5 participants (31%) reported they did so every minute or more. At the low end, 4 participants (25%) reported that they glanced at the visualization only 1–2 times. Based on observation, participants glanced at the visualization every 90 seconds on average. Eight participants closely estimated the actual number of times they glanced at the visualization. In contrast, nearly half of participants underestimated their glancing behavior, perhaps owing to less perceived distraction. Data for one participant was omitted due to poor video quality.

5.6 Design Implications

Based on findings from the lab study, we identified a number of design considerations to inform the next iteration of a revised visual design, including feedback timing through the update rate, size and position of the display, color, and visual metaphor.

5.6.1 Update Rate

Participants were divided about how distracting they found the visualization, and this could be due to the frequent update rate (i.e., every 6.5 seconds on average). However, several participants commented on perceived delays and mismatches in their behavior and the timing of the feedback, making the feedback difficult to interpret. In the next design iteration, we considered using a slower update rate, but to maintain visible traces of changes in affiliation and control levels in the display over time. In this way, we would allow those who found the display distracting or misaligned to adapt to a subtler refresh.

5.6.2 Display Size and Position

Several participants described problems interpreting the ambient feedback because the display was poorly positioned to be easily viewed and could end up creating strain. Compared to the first eight participants with the display close to them, the last eight participants who had the display behind the actress were more likely to say that the visualization was small and hard to see. In the next design iteration, we considered ways to size and position the feedback to help improve interpretability.

5.6.3 Color

At least three participants said that it was difficult to detect and interpret the meaning of the colors for the sun-moon, which changed simultaneously with size. In addition, one participant indicated that the sun-moon color change only became easy to see because the sun was large. We wanted to make sure that the color changes were easier to detect and facilitated interpretability of the feedback. As a result, we considered varying the saturation of a single color along a gradient in the next iteration of the design.

5.6.4 Visual Metaphor

Overall, participants found that the use of metaphor was helpful for interpreting the ambient feedback. Our initial design displayed control and affiliation as separate graphical elements. Many participants found the seesaw tilt a particularly straightforward way to understand who was dominating the conversation. Still, we realized that this particular metaphor has the potential to discourage the completion of a
6. Part 2: Follow up Interviews

Based on formative feedback obtained from the lab study, we revised the design of the ambient visual feedback and conducted follow up interviews with seven participants from the lab study to obtain additional design feedback.

6.1 Revised Visual Design

For comparison with the initial sun-moon/seesaw visualization, we selected a lotus flower metaphor to represent control and affiliation in an artistic composite visual element. Not only could we leverage this design to combine control and affiliation in a single element, but we could also use petal size and color to represent changes in the level of those signals. Because of problems raised in the lab study around timing and persistence of feedback, we also wanted to test the idea of visually animating the current state of dialogue while preserving visual traces of recent dialogue from the past few minutes.

In the revised visual design (Figure 4) each pair of flower petals represents a one-minute segment of dialogue in which nonverbal cues are captured between two people. We designed the lotus flower to be split in half vertically, with the clinician's nonverbal signals represented on the left and the patient's on the right. After one minute of dialogue, a new petal appears on each side at the base of the animated flower. Five sets of petals are displayed at a time to communicate the past few minutes of conversation. As the forward-most petal begins to appear with each new minute, the backward-most petal slowly fades. The size of each petal indicates how much control (large petal) or how little control (smaller petal) each user exhibited in the previous minute. Similarly, the petal color changes with each minute showing how much or how little affiliation each user exhibited in the previous minute. Initially, we chose to vary color from bright blue (i.e., high affiliation) to grey (i.e., low affiliation) (Figure 4 left). However, with feedback received through the first 3 follow up interviews, we revised the design to vary color more saliently from bright orange (i.e., high affiliation) to bright blue (i.e., low affiliation) through a neutral grey (Figure 4 right).

6.2 Participants

We brought back 7 of the 16 healthcare professionals who participated in the lab study. We purposively selected participants who could provide diverse opinions and represent a broad range of specialties. These participants included 1 physician, 1 dentist, 2 certified nursing assistants, 1 nurse, and 2 emergency medical technicians. Five men and 2 women returned, ranging in age from 25 to 51 years (mean 39). All participants were compensated for their time with a software gratuity. Five returning participants reported in the lab study that they would use the initial sun/moon/seesaw design in a professional setting. One participant was unsure and one said she would not. Thus, we included a range of accepting and skeptical participants.

6.3 Procedures

We conducted follow-up interviews four weeks after the lab study. Each participant took part in a 45-minute session. We confirmed participants’ responses to the initial visual design and introduced the new lotus flower visualization. We trained participants to interpret the new visual feedback by giving them a handout that displayed the meaning of the control and affiliation dimensions and playing a training video showing the full range of petal size and color.

To gather participants’ perceptions about acceptability of the revised visual feedback, we asked participants to respond to a questionnaire and interview questions after viewing the lotus flower visualization alongside the video from the lab study session in which they participated. We replayed the participant’s role-play video from the lab study on a 24-inch computer display. The role-play video showed the participant (shot head-on from one camera) and the actress (shot head-on from another camera). Alongside the display of this video, a 30-inch display simultaneously played the lotus flower visualization seeded with the participant’s control and affiliation log data from the lab study. To better understand the optimal size and position of the visual feedback, we chose this larger 30-inch display compared to the 17-inch display used in the lab study.

6.4 Measures and Analysis

After watching the video and the visual feedback together, we gave participants a written acceptability questionnaire that captured 1) preference for lotus flower or sun-moon/seesaw design, 2) openness to future use of the technology in a professional setting (yes/no/unsure), and 3) reactions to specific design revisions, including update rate, display size and position, color, and visual metaphor (open-ended qualitative response). We then interviewed par-
Participants to draw out design implications from their responses and ended the session with a debriefing. We calculated the proportion of participants who preferred each design. We also calculated the proportion of “yes”, “no”, and “unsure” responses about future use across participants. For the open-ended qualitative responses, we grouped quotes according to the recurrent design themes that emerged in the lab study (i.e., update rate, display size and position, color, and visual metaphor).

6.5 Follow Up Interview Results

6.5.1 Metaphor Design Preferences
Six of the seven participants (86%) preferred the lotus flower design to the sun-moon/seesaw design. These participants liked seeing the conversation shift over time through changes in the lotus flower. The one participant who preferred the sun-moon/seesaw felt the sun-moon/seesaw was more “objective and obvious” than the lotus flower (P8). One participant said that he thought the lotus flower was “way better” than the sun-moon/seesaw visualization because “it was more fluid, less distracting, subtle, and effective” (P1). He liked seeing the course of the conversation by looking at all the visible petals from the past few minutes as the forward-most petal changed in real-time. Compared to the sun-moon/seesaw, an oncology nurse thought the flower petals provided “more data to process at a time” (P4), but that the lotus flower feedback would be easier for her to understand with more training.

6.5.2 Future Use
At first, all seven participants had the same response about use of visual nonverbal communication feedback in a professional setting as they reported in the lab study. But during the interview, P16 – who initially did not think she would use it at all – said she could see use by healthcare professionals, particularly in oncology and mental health who are motivated to be particularly empathic. She was the only participant who was not enthusiastic about using ambient visual feedback as a training tool in healthcare. She thought that people who were not empathic would need such a tool, and her assumption was that they also would not be motivated to use it. However, all of the other participants believed that the ambient feedback could provide an effective way to learn about nonverbal cues and empathy. Thus, participants were overall quite open to the use of this technology in professionals settings in the future.

6.5.3 Reactions to Design Revisions

Update rate: The update rate proved crucial in determining the acceptability of the lotus visualization. Despite concerns from the lab study participants about potential distractions from the frequent update rate, the first three participants in follow-up interviews wanted an instant snapshot of the feedback and did not want to wait a whole minute to see their control and affiliation levels change. P7 told us “it was hard to remember how the conversation was going in the last minute.” An emergency medical technician agreed, saying that with the one-minute update he lost the opportunity to find out “what was I talking about when that [change in feedback] happened?” (P3). As a result, we decided to test a revised version of the lotus flower visualization with the last four participants. In this revised version, the newest petal updated in real time to provide instant feedback rather than appearing after 1 minute of dialogue. Once this change was made, only P16 had an issue with the update rate. She had more trouble trusting the data in the moment but could still see using the lotus flower visualization in some situations. The remaining three participants found the real-time update informative.

Display size and position: When asked about the ideal size and position of the visual feedback, most participants thought that the larger 30-inch display could make changes easy to detect in a clinical setting, but two participants said they thought the 17-inch display used in the lab study could be just as acceptable. Placement seems to be a matter of room configuration and individual preference.

Color: The first three participants found it challenging to differentiate between the five blue saturation levels that we used to represent changes in affiliation. Because the saturation difference was so subtle, P8 commented that he did not feel like he was getting enough feedback. To him, the colors represented something that was far more relaxing than the task at hand. Given the first three participants’ views on color, we changed the color spectrum to be more obvious for the remaining four participants. We used bright orange (to represent warmth) and bright blue (to represent hostility) at each end of the affiliation spectrum, with a neutral gray in the middle. One satisfied participant commented that the mapping of the new palette to affiliation was quite effective. Nevertheless, P4 said she could easily detect petal size differences to interpret control but still needed to learn what the different colors meant before being able interpret affiliation as quickly.

Visual metaphor: Three participants debated about whether the subtle artistry of the lotus flower representation detracted from its purpose as a communication training tool. P8, a male certified nursing assistant, felt that while the revised visualization had a “pleasing appearance and may not interfere with the conversation as much,” the lotus flower was not as obvious of a metaphor for communication. He was the only healthcare professional who preferred the sun/moon and seesaw, saying that it was “more objective and obvious to me” (P8). P7, a dentist, suggested that a design with abstract shapes might be easier to focus on for real-time feedback about control and affiliation than the lotus petal representation. Despite this suggestion, he ultimately preferred the lotus visualization over the sun-moon and seesaw, since he could interpret control and affiliation together in a single visual element.

7. Discussion
Results from this work indicate that ambient visual feedback from SSP could encourage empathic nonverbal communication. In particular, Entendre could serve as a valuable training tool for patient-clinician interaction. Still, motivation to change – or simply to be aware of – one’s nonverbal communication style is a key ingredient for training healthcare professionals with Entendre. Over the long term,
Entendre could be used to assess communication skills and remind clinicians to maintain and improve empathic nonverbal skills, either intermittently, continuously, or reflectively.

Through our model formulation, we delineated mappings between nonverbal cues and primary relational signals for interpersonal communication evidenced in the extent literature. We generated an extensive framework that lists a full range of nonverbal cues that can be captured by SSP to drive visual feedback on affiliation and control exhibited in interpersonal interaction. Through our 2-part formative user study, we learned that clinician acceptability for ambient visual feedback shows promise. Our design iterations point to important design considerations, including metaphor, color, size, position, and timing for conveying effective ambient feedback to clinicians.

By pulling together a wealth of evidence on patient–clinician communication, our theoretical model provides a conceptual space from which to consider the design of both SSP that captures nonverbal behavior and displays that reflect that behavior. We have demonstrated the utility of our framework for informing ambient displays that illustrate primary constructs in interpersonal communication. In the future we will further explore how our framework can guide the selection of priority cues to drive SSP. Although empirical validation of our framework is an important next step, it already expands the existing literature with a detailed analytical tool for analyzing nonverbal communication.

Our formative user study identified a number of key considerations for designing ambient visual feedback that is instrumental in helping clinicians to reflect on their communication practices. In particular, interpretability and actionability of feedback is greatly shaped by feedback timing, size, position, color, and visual metaphor. Although providing feedback in real time can enhance self-awareness by tightly coupling changes in the display to behavior, we learned that real-time deployment also has the potential to distract clinicians from patient care. Future work could examine variations in display update rate, size, and position to better understand these trade-offs. Alternative deployments, such as training settings without potential negative implications for direct patient care could also be explored.

We also learned that while careful selection of update rate, color, and visual metaphor helps to promote interpretability of ambient visual feedback, individual preferences vary. Clinicians may also vary in how actionable they prefer the display. Participants who wanted more granular updates liked gaining awareness of what happened in the moment to change course, while others wanted to just check on the interaction on the whole. Preference for visual metaphor may also vary. For example one participant suggested less abstract and more familiar visualizations than the lotus flower, such as charts and graphs. We also found that preferences varied for nearby displays that are easy to refer to versus displays set in the background to provide a subtle reminder. Furthermore, preference for the salience of feedback might vary by level of clinical experience. Future work could explore customizing feedback to match a clinician’s preferences and training goals. For example, individuals could configure the update rate to strike a balance between being useful in the moment and being overwhelming to interpret. Some could choose a metaphor that provides a salient reminder or that blends into the background like artwork. In the future, we could explore other design interventions, such as providing feedback retrospectively for clinicians to review outside of patient care or providing tactile feedback that would be imperceptible to the conversational partner and serve an alternative modality for feedback.

The purpose of this work was to establish nonverbal cues indicative of patient-centered communication and then to determine clinician acceptance of proof of concept designs that reflect feedback on a subset of cues through a formative user study. Although our findings are promising, our exploratory study design has limitations that we will address in future work. First, we employed Wizard of Oz as a low-cost technique to drive reflective feedback through a scenario-based lab study. Although this approach enabled us to simulate real-time feedback through role-play, it provided potential users with feedback that captured only the essence of the researcher’s perceptions of their nonverbal communication style. Thus, simulated use within a role-play scenario could have only approximated real world use where a full range of nonverbal cues could be captured automatically. We are currently building Entendre to analyze audio and video streams in real-time to drive the visual feedback. Our findings from the reported study help to inform this system development by addressing emergent design considerations early. Follow-up studies with a fully implemented system in real-world settings will allow users to gauge its utility as well as measure levels of affiliation, control, and distraction of real-time ambient feedback. In particular, future evaluations could explore the real-world use of Entendre in clinical settings, as well as other settings in which effective interpersonal communication is paramount.

Second, our exploratory work relied on the manual capture of nonverbal cues and adjustment of reflective feedback. While this approach enabled us to easily test Entendre as a proof of concept with a small subset of nonverbal cues, manual classification by a researcher could have been biased toward capturing more observable and salient cues that resulted in feedback that reflected only part of the nonverbal communication occurring. With this limitation, our small sample size may have contributed to poor alignment of patient-reported control scores and actual conversational control. In our ongoing work, we are developing a classifier that will automate the mapping process, which we can train and test for accuracy to empirically validate our theoretical framework. Through this future work, we will examine the extent to which Entendre corresponds with other measures of empathic patient-centered communication, such as users’ self report and human coding of nonverbal behavior. Because individual variation in communication style and contextual factors (e.g., task-related versus social interaction, depth of relationship) can influence the usefulness and fit of reflective feedback, it is important to account for these factors and examine the ways they shape how Entendre performs.
We did not examine learning effects in this study. The training session for the lab study was brief yet adequate. Had the visualization required learning over time, the effect could be quite different and impact the optimal design. What this means in terms of design tradeoffs has yet to be determined. Perhaps an abstract visualization like the lotus flower might take longer to learn, but become less distracting over time with practice. In contrast, simple charts and graphs might be easier to learn but become more distracting.

Despite these limitations, findings from this work show substantial promise of ambient visual feedback on nonverbal communication for clinicians and patients. Not only could Entende be used as a training tool for real-time feedback, but captured data could also be examined reflectively for development of nonverbal communication skills over time. Because Entende uses audio and video streams, this approach could be applied to both face-to-face interactions and remote telehealth interactions. Entende could also be useful in a range of non-clinical settings in which interpersonal communication is important, such as teaching or videoconferencing.

8. Conclusion

We presented a novel approach for reflecting ambient feedback on communication style. Using SSP, nonverbal assessment can be particularly valuable for encouraging empathic, patient-centered communication by clinicians. This initial work provides early evidence that reflective visualizations of nonverbal behavior can be theoretically-rooted and acceptable to clinicians given consideration of important design choices, including metaphor, color, size, placement, and timing of feedback. Tools, such as Entende, could enhance clinicians’ empathic patient-centered communication skills and lead to greater patient involvement in health care and improved health outcomes in the future.

Acknowledgments

We would like to thank the healthcare professionals and actors who participated in our study. We also wish to thank Valerie Manuso and acknowledge Roland Fernandez and Jessica Herron for production assistance.

References