Towards Automated Pain Assessment using Embodied Conversational Agents

Prasanth Murali
Northeastern University
Boston, MA, USA
murali.pr@northeastern.edu

Mehdi Arjmand
Northeastern University
Boston, MA, USA
arjmand.me@northeastern.edu

Matias Volonte
Clemson University
Charleston, SC, USA
volontematias@gmail.com

Zixi Li
Northeastern University
Boston, MA, USA
li.zixi1@northeastern.edu

James Griffith
Northwestern University
Chicago, IL, USA
j-griffith@northwestern.edu

Michael Paasche-Orlow
Tufts Medical Center
Boston, MA, USA
michael.paasche-orlow@tuftsmedicine.org

Timothy Bickmore
Northeastern University
Boston, MA, USA
t.bickmore@northeastern.edu

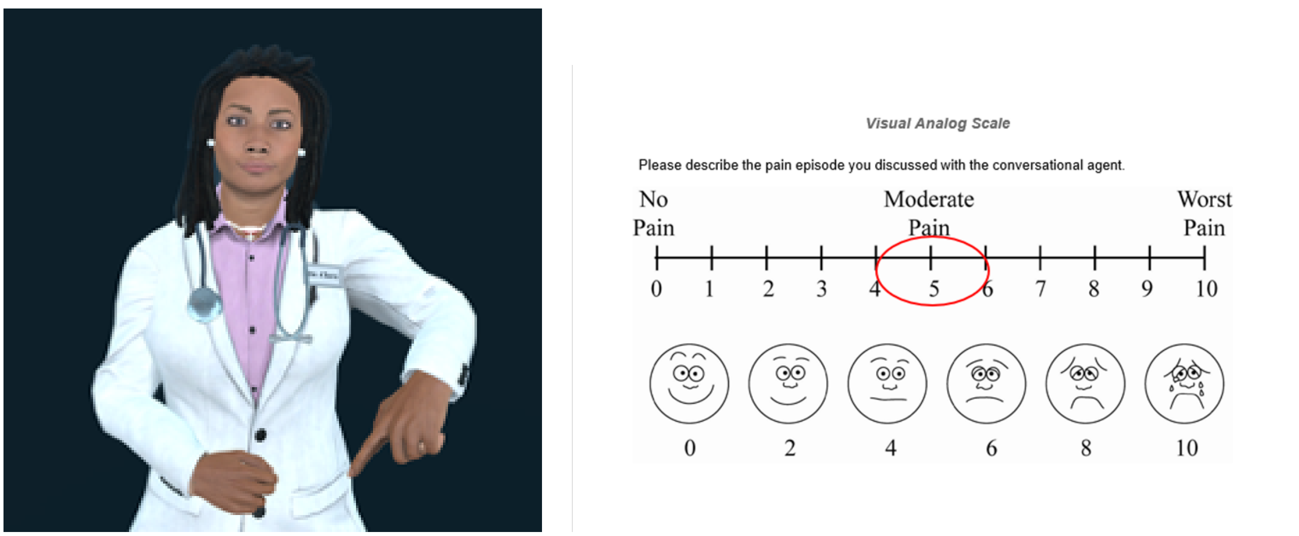

Figure 1: Left: ECA providing empathic feedback, demonstrating its understanding of a user's bodyache. Right: Conventional self-report on intensity of pain (from [37]).

ABSTRACT

Narrative accounts are the ultimate authoritative source for pain assessment, and face-to-face encounters provide a rich context in which nonverbal conversational behavior can be used to enrich the detail in these descriptions. Embodied Conversational Agents—animated characters that simulate face-to-face conversation—can provide a medium for automated pain assessment in which multimodal pain narratives are elicited, clarified, and grounded. These agents can also use facial displays and nonverbal behavior to provide simulated empathic responses to help patients feel understood. We describe work towards a conversational agent that elicits various aspects of a pain experience, followed by an empathic summary. Our preliminary findings indicate that individuals are comfortable reporting their pain experiences to the agent, and are largely satisfied with this assessment methodology. We also find evidence that patients prefer the conversational empathic summary to standard self-report measures.

ICMI '23 Companion, October 09–13, 2023, Paris, France

ACM ISBN 979-8-4007-0321-8/23/10…$15.00
https://doi.org/10.1145/3610661.3617132

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**.

KEYWORDS

Pain Assessment, Embodied Conversational Agent, Virtual Agent, Empathic Feedback

ACM Reference Format:
Prasanth Murali, Mehdi Arjmand, Matias Volonte, Zixi Li, James Griffith, Michael Paasche-Orlow, and Timothy Bickmore. 2023. Towards Automated Pain Assessment using Embodied Conversational Agents. In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '23 Companion), October 09–13, 2023, Paris, France.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3610661.3617132

1 INTRODUCTION

Pain is a multi-dimensional, subjective phenomenon typically assessed through self-report. However, the pain experience can be very difficult to describe [38]. Even complex pain assessment protocols, such as the McGill Pain Questionnaire [26], can miss many subtleties. This can lead to an incomplete picture of a patient's experience for consideration in clinical decision-making. Moreover, an incomplete assessment can leave patients feeling unheard and uncared for.

Although self-report measures are the clinical standard for pain assessment, pain narratives have been considered the "gold standard" for describing pain experience [43]. Narratives have been aligned with the International Association for the Study of Pain (IASP) definition of pain as a personal, subjective, qualitative experience. They are also aligned with ethical principles of autonomy. Relying exclusively on pain behavior, standardized self-report measures, or automated assessments can challenge patient authority on their pain experience.

Studies of pain narratives delivered in face-to-face encounters demonstrate that nonverbal behavior in general, and hand gestures in particular, can play an essential role in descriptions of pain experiences [20, 36–38]. One study found that 78% of the utterances describing pain sensations were accompanied by representational hand gestures and that 43% of these gestures contained complementary information that was not represented in the verbal channel [37]. Thus, hand gestures can be used to describe aspects of pain sensation that provide a fuller description of pain experience than speech alone [36, 37]. Hand gestures can indicate the location of pain, demonstrate the actions that cause pain, or describe the pain sensation (e.g., banging, tickling, compression, jabbing) [37].

Another important factor in pain assessment is the social context in which it occurs. There is substantial evidence that people modulate their expression of pain depending on the social context, in particular the receptivity and empathy of their interlocutor [12]. This is not just a communicative adaptation; the subjective experience of pain can change depending on the social context. Enabling patients to describe their pain to a nonjudgmental and empathetic listener–ideally, someone they have established trust and working alliance with—can be crucial in obtaining detailed, unbiased reports.

Embodied Conversational Agents (ECAs) are animated characters that simulate face-to-face conversation using nonverbal conversational behaviors, such as hand gestures, gaze, posture shifts, and facial display, in synchrony with speech. ECAs can provide a rich, intuitive interface that captures many of the affordances and subtleties of human face-to-face interaction [10]. They have been shown to be particularly effective for health communication with patients with low health, reading, or computer literacy [7]. Given their use of narrative and nonverbal conversational behavior, ECAs could provide an ideal medium for eliciting and grounding pain narratives from patients. They can leverage their nonverbal behavior, such as hand gestures, to mirror patient behavior to demonstrate understanding of the patient's narrative. They can use facial displays of concern to demonstrate empathy. They can also use social and relational behavior (e.g., immediacy behavior, social chat, rapport building) to establish trust and working alliance with patients [5] to provide the optimal social context for pain assessment.

In this paper, we describe preliminary work on an ECA that elicits individual pain descriptions, along with a pilot feasibility study. The agent first builds trust with a user, then interviews them about different aspects of a past pain experience, and finally provides a multi-modal report of its understanding along with simulated empathy, paying particular attention to the use of hand gestures by participants. We compare the empathic ECA feedback to a control condition in which standard pain assessment forms are filled out for the patient to assess impacts on satisfaction and perceived empathy.

2 RELATED WORK

In this section, we review relevant literature on two main topics of our work: automated pain assessment and ECAs.

2.1 Automated Pain Assessment

Clinical research has enabled physicians to employ a diverse range of methods, encompassing self-reports, qualitative experiences, and non-verbal cues, to gain insights into patients' pain levels [43]. Furthermore, leveraging recent advancements in deep learning methods, researchers have made significant strides in automating pain assessments. Gkikas et al.'s systematic review underscores the various techniques used in automated pain detection, spanning both unimodal and multimodal approaches [18]. Notably, multimodal pain assessment, which integrates different signals like physiological and behavioral expressions, has proven particularly advantageous, especially in challenging scenarios where verbal descriptions are limited, such as in the case of infants [46]. A follow-up review by Al-Eiden et al. highlights the diverse sensors employed in automated pain sensing, including vision-based systems (e.g., RGB cameras), contact-based devices capturing physiological and respiratory data, and microphone-based systems for audio-recording [25]. Moreover, researchers have also successfully approximated standard self-reports using multimodal approaches, as demonstrated in [24, 45]. Their model comprises a two-part framework: a multitask learning neural network for predicting multidimensional pain scores, and an ensemble learning model for accurately estimating Visual Analog Scale (VAS) ratings. These advancements collectively contribute to automated pain assessment, with the potential to augment clinical practices and improving patient care.

2.2 Embodied Conversational Agents in Health Care

Embodied Conversational Agents ("ECAs") have proven to be particularly advantageous for health ECAs can increase users engagement [44], trust [34], even in longitudinal settings [16]. These constructs are vital for the medical field. Thus, ECAs are actively used in the health care field for training medical personnel [42], for providing mental health therapies [35], or for providing healthy lifestyle advice [17]. Additionally, researchers have also explored the development of ECAs for preconception care counseling in young women [4], medication adherence counseling for patients with atrial fibrillation [21], anxiety management [31, 32], and exercise promotion for geriatrics patients [48]. Moreover, ECAs that display relational skills such as empathy, have also been shown to increase working alliance in automated health behavior change interventions. For instance, Zhang et al., demonstrated that ECAs can effectively play the role of a decision coach and facilitate shared decision making between patients and clinicians, regarding prenatal testing for Down syndrome [47]. These systems show that ECAs have been effective for the delivery of health care information and decision support. Furthermore, ECAs have been found useful in engaging low-literacy populations [6]. Given their use in self-disclosure [30, 40], especially in sensitive topics, we are interested in exploring the role of ECAs in the context of pain assessment.

3 PAIN ASSESSMENT AGENT PROTOTYPE

We developed a prototype ECA to interview individuals about a past pain experience and provide an empathic summary in response. The ECA system used in this study was developed in the Unity 3D game engine. The 3D model of the agent was designed in Adobe Fuse. We also developed the verbal and non-verbal behaviors of the character in Autodesk Maya and Mixamo. These behaviors included facial expressions, hand gestures, and different body posture animations. The facial expressions included visemes and facial action unit blendshapes. Additionally, we used keyframe animation to create different arm, hand, head, and body behaviors. The verbal and non-verbal behaviors were created following existing literature [15, 28]. In order to conduct a feasibility and acceptance test of the ECA without fully developing a natural language understanding capability, we used a Wizard of Oz control system, where the agent's dialogue is controlled by a human confederate through a 'wizard' interface. The agent speaks using a speech synthesizer, and uses a range of nonverbal behavior, including facial display of affect, communicative eyebrow movements, directional gazes, and head nods, with most nonverbal behavior automatically generated using BEAT [11]. The conversation with the agent consists of two phases: the assessment phase and the reflection phase. In the assesment phase, the agent employs open-ended prompts to encourage users to describe their pain experiences. Based on our literature review, we identified key dimensions of pain, which the agent systematically queried for each participant (Table 1) [37]. The agent begins by prompting the user to describe their pain experience. Subsequently, the agent asks the user about the location, intensity, onset, frequency, and associated feelings of the pain. To elicit more open-ended responses, users are also asked to reflect on the emotional and physical sensations they experienced when being interviewed about each of those dimensions of pain. In the reflection phase, the agent reflects what it has learned from the user back to them to ensure accurate understanding.

We hypothesized that providing empathic feedback via the agent would enhance users' feelings of being heard and their acceptance of the agent system. To test this hypothesis, we designed two different conditions for the reflection phase:

a) EMPATHIC feedback: In this condition, the agent conveys empathy and understanding by summarizing the elicited pain information to the user using speech and non-verbal behavior, including gestures. For example, when the user describes a backache, the agent points to its back to indicate the location and says, "I understand that you are experiencing a sharp pain in your back. Is that correct?." This combination of verbal and non-verbal cues can enhance the empathic response of the agent, as shown in Figure 1 - Left.

b) REPORT feedback: In this condition (Figure 1 - Right), standard self-report questionnaires, such as the Visual Analog Scale, are displayed to the user with their pain information filled in.

By comparing these two reflection phase variants, we can evaluate the impact of empathic feedback on user experience and acceptance of the ECA.

4 PILOT STUDY

We conducted a pilot study in order to evaluate the acceptance and feasibility of an ECA to elicit multimodal pain descriptions from users with a between-subjects manipulation of agent feedback to compare the empathic ECA feedback to a control condition in which standard pain assessment forms are filled out for the patient.

4.1 Methods

4.1.1 Recruitment. Participants were recruited via email lists and online fliers. They were required to be 18 years or older, be a native speaker of English, and have experienced an episode of pain within the last two weeks. Our institution's Ethics Review Board approved the study. Participants were compensated for their time.

4.1.2 Study Protocol. Participants who confirmed their eligibility via email communication were scheduled for a one-hour in-person session at our lab for the study. They were randomly assigned to either the EMPATHIC Feedback or REPORT Feedback conditions. The session began with a brief study summary and obtaining verbal consent from the participants. They then completed the baseline measures described in section 4.1.3. Next, participants were informed that they would engage in a 15-minute interaction with an ECA, where they discussed their pain experiences through natural speech. The research assistant observed and controlled the interaction from an observation room using the Wizard of Oz protocol.

After the interaction, participants were asked to complete a final set of questionnaires outlined in section 4.1.3. Additionally, semi-structured interviews were conducted with each participant to gain deeper insights into their experiences during the conversation with the ECA. The interview transcripts were analyzed using thematic analysis techniques. The total interview time across all participants was approximately 40 minutes. Finally, participants received a debriefing about the system and the use of the wizard to control the interaction.

Dimension	Example pain questions asked by ECA
OPEN	Feel free to speak about the physical sensations, emotions, and thoughts connected to your pain.
WHEN	When did this happen to you?
WHERE	Can you describe in detail where in your body you experienced the pain?
SENSATION	How would you characterize the sensations you feel in that area?
INTENSITY	On a scale of 1 to 10, with 1 being mild and 10 being the most severe, how would you rate the intensity of the pain at its worst, at its average, and at the moment?
IMPACT	On a scale of 1 to 5, how has the pain affected your ability to perform certain activities or engage in specific tasks?
FREQUENCY	Within the last two weeks, how often did you experience pain?
PATTERN	In addition to the frequency, have you noticed any particular patterns or variations in the pain episodes?
TRIGGER	Are there certain times of day or specific situations that seem to trigger or exacerbate the pain?
FEEL	Is it sharp, dull, aching, burning, or do you feel any other sensation?
EMOTION	Could you elaborate on how the specific qualities of the pain you described affect your emotional state or overall well-being?
ONSET	Does the pain come on suddenly or gradually? Have you noticed any changes in the pain over time? For example, has it worsened, improved, or remained stable?
OPEN-2	As we conclude this session, take a moment to reflect on your pain experiences and how expressing them or speaking about it has influenced your understanding or emotional well-being.
STATUS	For example, has it worsened, improved, or remained stable?
REFLECT	If you have any more thoughts or if there is anything else you would like to share, please feel free to do so.
EMPATHY	I want to make sure I understood you correctly, so let me reflect on what you told me.

Table 1: Pain-related Dimensions and Sample Dialogs by the ECA

4.1.3 Measures. Participants completed several measures at various stages of the study. At baseline, a demographics questionnaire was administered along with an assessment of their baseline health literacy levels using the REALM measure [14]. During the interaction with the agent, we captured participants' speech and non-verbal behavior through a standard microphone and RGB camera. Additionally, a Microsoft Kinect was used to capture 3D depth data for motion analysis. Furthermore, their physiology (electrodermal activity and heart rate) was recorded using the Empatica E4 sensor. Following their interaction with the agent, participants completed the Working Alliance Inventory [19], the Active-Empathic Listening Scale (AELS) [8], the Responsiveness of Other Scale [41], the MSCEIT-Based Perceived EI Questionnaire [27], the Godspeed Agent Attitude Questionnaires [1], and a Satisfaction/Usability Rating Questionnaire [3]. Additionally, participants provided self-reports on their pain experiences using standard measures, including PROMIS [2], the visual analog scale [23], the McGill Pain Assessment [29], and the Patient Health Questionnaire-8 (PHQ-8) [22]. Finally, semi-structured interviews were conducted with participants to obtain an overall understanding of their experience. Interviews were recorded, transcribed, and analyzed using thematic analysis techniques [9].

4.2 Results

4.2.1 Participants. A total of 9 participants (5 males, 4 females) were recruited for this study. The participants had a mean age of 44 years, and 3 was categorized as having low literacy based on the REALM measure. The participants in our study also reported

Figure 2: An example of a user interacting with the ECA system

that they have an average pain level of 6.0 out of 10 on the Visual Analog Scale.

4.2.2 Quantitative Self-Report. Given the small sample size and non-normality of distributions, we used non-parametric Mann-Whitney U Tests to analyze the differences in perception of the agent between participants in the two conditions. The descriptive statistics are presented in Table 2.

The results revealed a significant effect, with participants in the EMPATHIC Feedback condition reporting higher empathy ratings of the agent on the Active Empathic Listening Scale compared to participants in the REPORT Feedback condition ($Z = -2.26$, $p = .02$). Similarly, participants in the EMPATHIC Feedback condition reported significantly higher scores for the agent on the Likability ($Z = -2.37$, $p = .02$) and Perceived Intelligence dimensions ($Z = -2.20$, $p = .03$) of the Godspeed Agent Attitude Questionnaire compared to those in the REPORT Feedback condition.

4.2.3 Qualitative Findings from Interviews. High-level themes emerged from our interviews with the participant, which both corroborated and contextualized our quantitative findings, and gave design directions for future systems in this space. The participants in our study perceived the agent as displaying empathy and sympathy toward their feelings. They mentioned that the agent *"seemed pretty caring for a robot[agent]"* [P1, EMPATHIC condition] and that *she seemed sympathetic* [P3, EMPATHIC]. They described the naturalness of interaction as *"like talking to a robot [agent], but a robot[an agent] that was sort of modernized to be more human-like"* [P3, EMPATHIC]. Participants especially highlighted the agent's attentive listening behavior and mentioned that *"I thought she listened to me, and, you know, she knew my name and everything"* [P6, EMPATHIC]. In addition, participants also expressed mixed views on the agent's understanding of their pain. The participants in the REPORT condition expressed uncertainty about whether the agent truly understood the pain experience or not. They expressed a preference for the agent to explicitly reiterate and acknowledge their pain, whereas participants in the EMPATHIC Feedback Condition felt that the agent understood their pain. For example, participants in the REPORT Feedback condition reported that *"She understood it, but she didn't really repeat it back to me."* [P6, FORM] and that *"I didn't get the responsive answers relevant to my particular scenario. Those are so generic."*[P5, EMPATHIC]. They also mentioned that the "*The way it was getting the points of me and summarizing it at the end"*[P9] made them like the agent. On the other hand, participants in the EMPATHIC Feedback condition mentioned that the agent understood their pain *"very well"* [P5, EMPATHIC] and that "She understood it. Because she said, on a scale of 1 to 10 what is your pain level, and I said it was a 12, And she kept repeating that number to me. So, she said it was a 12 for you." [P1, EMPATHIC]. This suggests that the agent's responses' reflective nature could offer more utility than the conventional standard of care reflection. Furthermore, participants also perceived the agent as a potentially useful tool for assessing and addressing their medical issues remotely, without *"getting up and going to make an appointment and everything"* [P5, EMPATHIC]. They appreciated the convenience of avoiding physical appointments and the ability to communicate their concerns effectively. Participants also expressed critiques, which could serve as design guidelines for developing future systems. In particular, participants expressed a desire for more *tailored questions* [P2, REPORT], *responses* [P6, EMPATHIC], and *follow-up queries* [P7, CONTROL] that demonstrated a desire for deeper understanding of their specific pain situations. This led to some participants feeling that *"it felt a bit artificial in terms of the interaction."* [P1, EMPATHIC].

4.3 Analysis of Nonverbal Behavior

The video recordings of participant verbal and nonverbal responses to agent questions were manually analyzed and annotated for hand gesture, following McNeill [28]. Overall interactions ranged from 7.0 minutes to 14.7 minutes, with an average time of 10.6 (SD 2.2) minutes. Interactions were significantly longer for the six participants in EMPATHIC condition, Mann-Whitney $U = 1.5$, $p < .05$, with interactions running an average of 11.7 (SD 2.0) minutes compared to 9.2 (SD 1.5) minutes for those in the REPORT condition (Table TBD). This is due to the added time of the empathic feedback from the agent, which averaged 30.2 (SD 7.3) seconds. McNeill's typology classifies gestures into one of four categories [1]: 1) Beat gestures are bi-phasic flicks of the hand in time with the prosodic peaks of speech, used primarily to add emphasis to what is being said. 2) Deictic gestures point to place either in physical space (e.g., a body part that has experienced pain) or to a conceptual entity in gesture

Measures	EMPATHY Mean	EMPATHY SD	REPORT Mean	REPORT SD
Active Empathic Listening Scale (*)	5.20	0.84	3.25	0.96
Agent Satisfaction	5.80	0.84	5.75	0.96
Working Alliance	59.80	20.55	48.50	8.96
Responsiveness Of Other	5.40	1.67	3.75	1.89
Godspeed Agent (anthropomorphism)	2.25	1.71	2.00	1.41
Godspeed Agent (animacy)	3.00	0.71	1.75	1.26
Godspeed Agent (likeability) (*)	4.80	0.45	2.75	1.90
Godspeed Agent (perceived intelligence) (*)	4.00	0.00	2.50	1.73

Table 2: Descriptive Statistics of Quantitative Self-Report Measures

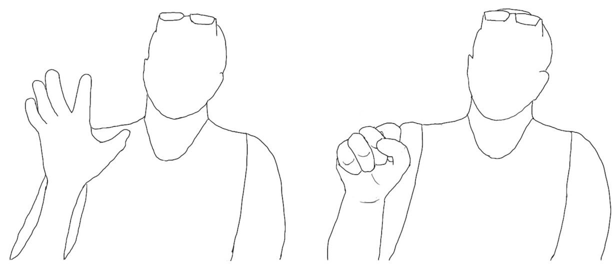

Figure 3: Sequence of gestures, where the participant describes their pain as *constricting*

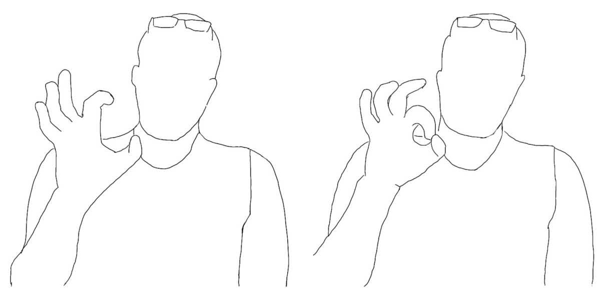

Figure 4: Sequence of gestures, where the participant describes their pain as going from *2 inches in diameter to 1 inch*

space. 3) Iconic gestures represent images of concrete entities or actions (e.g., moving the hand up an imaginary staircase step-by-step when describing walking upstairs). 4) Metaphoric gestures map an abstract concept to a physical shape or trajectory using a spatial metaphor [39], then represent the spatial aspects with gesture (e.g., raising a hand to show increased pain). Iconics and Metaphorics together are referred to as Representational gestures. In addition to these four core types, emblems (or emblematic gestures) are those that have conventional meaning (e.g., the "OK" gesture) and contrast gestures are pairs of beats that are used to mark antonyms (e.g., 'good or bad').

In our corpus, hand gestures were observed for eight of the nine participants, with the ninth never lifting his hands into camera view. The following analyses exclude this participant from the count statistics. Non-communicative self-adapters (e.g., scratching the face) are also excluded. Participants exhibited 6 to 48 communicative hand gestures during their overall interaction with the agent, average 23.9 (SD 14.2). As found in many other studies of face-to-face conversation, the majority of these (56.5%) were beat gestures (range 0 to 43, mean 13.5, SD=13.8), used for emphasis, occasionally co-articulated on top of a representational gesture. The next most common type of gesture observed was deictic gesture, averaging 4.5 (SD 4.8) per participant. The majority (86.1%) of these were to body parts that participants were describing related to their pain experience. Representational gestures (iconics and metaphorics [1]) were observed but at a lower frequency, averaging 4.1 (SD 3.1) per interaction. Metaphoric gestures were used to describe concepts including 'repetitive', 'alone', 'group of people', 'offer', 'higher level of pain', 'higher difficulty', 'active (as in sports)', 'jumping around', 'constricting (pain feeling)', 'pain event (as conceptual object)', and 'uncertainty'. Contrastive gestures (e.g.,

Pain Dimension	Duration (seconds)		Hand Gestures per Response (count) **			
	Mean	**Stdev**	**Beats**	**Deictics**	**Representational**	**TOTAL*** **
EMOTION	20.10	14.70	1.50	0.25	0.75	2.60
FEEL	15.70	11.50	1.63	0.75	0.75	3.13
FREQUENCY	15.00	10.30	0.88	0.50	0.63	2.00
IMPACT	12.70	8.70	0.25	0.00	0.00	0.38
INTENSITY	12.90	7.10	0.50	0.00	0.00	0.50
ONSET	17.00	18.30	2.50	0.38	0.38	3.63
OPEN	13.60	10.60	0.38	0.13	0.00	0.63
OPEN-2	19.00	14.40	0.00	0.00	0.00	0.00
PATTERN	13.90	8.30	0.75	0.13	0.50	1.38
REFLECT	14.40	11.40	0.00	1.60	0.20	1.80
SENSATION	20.25	17.50	1.13	0.38	0.25	1.75
STATUS	18.30	13.90	1.50	0.00	0.38	2.25
TRIGGER	12.00	11.50	0.38	0.25	0.50	1.25
WHEN	16.40	10.60	2.00	0.00	0.00	2.00
WHERE	17.10	16.50	0.00	0.75	0.13	0.88
EMPATHY*	30.20	7.30	0.00	0.00	0.00	0.00
TOTAL PER SUBJECT	635.78	128.76	13.50	4.50	4.10	23.90

Table 3: Participant Response Durations and Gesture Counts. * For the 5 subjects in EMPATHY condition. ** Excluding one subject who apparently did not gesture at all. * Includes all other communicative gestures**

'whether I'm lonely or not lonely') and emblematics (e.g., 'two' with two fingers raised) were also observed, but were relatively rare. When broken down by pain question type, most gestures occur when asking about pain onset, sensation, and impact on emotional state. Deictics are most commonly found in response to questions about pain location, sensation, and at the end of the interview when participants are asked to reflect on their pain experience (some participants waited until this prompt to give a full description of their pain experience). Representational gestures were most often used in response to questions about pain sensation, emotional impact, and frequency.

4.3.1 Example of Complementary Meaning from Metaphoric Gesture. In order to demonstrate the value of multi-modal analysis of narrative descriptions of pain experience, we present a detailed analysis of one patient response that includes a range of hand gestures. In response to the PATTERN prompt in Table 3, one participant described a change in pattern in their pain experience. As shown in Figure 5, the participant (P) describes a change in their pain over time from a two-inch diameter region in their lower back to a one-inch diameter region. They use a deictic gesture to indicate where in their back they experience the pain, and three iconic gestures to reinforce the concept that it decreased in size. They use one gesture to describe their pain as "constricting" (Figure 3). They brought their fingers together in a grabbing-like motion, potentially illustrating the sensation of reduction. Additionally, to depict the reduction in pain from a "two-inch" to a "one-inch" diameter, they used another iconic gesture (Figure 4). They opened their fingers wide to indicate a larger circle and closed their thumb and index finger to represent a smaller circle, symbolizing the reduction in pain size. Together these gestures convey both the location, size, and change in physical area affected by pain, conveying additional information and subtlety not present in speech.

5 DISCUSSION

Our results provide evidence for the acceptance of a multi-modal ECA as a pain assessment tool. All participants readily described various aspects of their pain experience to the agent, and rated their satisfaction with the experience 5.8 on a 1-7 scale. The participants in the EMPATHIC Feedback condition of our study indicated that the agent demonstrated sympathy, empathy, and an understanding of their painful experiences. This aligns with previous virtual human interviewers that facilitate sharing psychological distress information from users [13]. Next, our study's findings demonstrate the importance of the reflection component in our system. Participants preferred the agent to showcase understanding and provide feedback explicitly during the conversation. Our findings contribute to that by demonstrating the effectiveness of empathic reflection through the agent's verbal and non-verbal behavior in pain assessment. Finally, our qualitative analysis suggests that future designs should adopt user-centric approaches that adapt to individual needs, preferences, and pain contexts. This can ensure a more personalized and tailored user experience. Through our study, we also contribute to the understanding the relationship between language and non-verbal behavior, in the context of pain expression. Our work also aligns with previous research on eliciting unpleasant and felt experiences from users. A related study explored the use of gestures by individuals while describing their toothaches and backaches [33]. For instance, when discussing backaches, participants used gestures to visually illustrate the pain as "the crest of it." They raised their hand and pointed to the highest point in the gesture space, effectively conveying the intensity of the pain through this

```
A: In addition to the frequency, have you noticed any particular patterns or variations in the pain episodes?

P: Yah, it seems [to] / constrict / in the pain.
  Gesture type: iconic, right hand, shape S, upper right center, motion: closing to fist, hand meaning: constricting

P: As the days went on / [it] [wasn't] [as] / oh, inflammatory. /
                         (1)     (2)   (3)
   (1) Gesture type: beat, right hand, shape B (open flat)
   (2) Gesture type: beat, right hand, shape B (open flat)
  (3) Gesture type: iconic, right hand, shape S, upper right center, motion: closing to fist, hand meaning: constricting

P: Sooo if it was [like two inches I'd say] it went down to like an inch of [pain], in one area, when it first happened.
                             (1)                                                (2)
   (1) Gesture type: iconic, right hand, shape F (finger thumb  in circle), upper right center,
       motion: circle decreasing in diameter, hand meaning: reducing in size
   (2) Gesture type: deictic, right hand, point at lower back right side
```

Figure 5: Example user description of pain, following coding conventions in [28]

upward motion. We thus argue that understanding and accounting for the role of gestures in expressing narrative subjective experiences is essential. Thus, multi-modal systems should capture both verbal and non-verbal cues such as gestures, to better understand the nuances of user experiences related to their pain.

Despite the positive results, it is important to note that there are several limitations in the current experiment, beyond the small convenience sample used. First, using a wizard-of-oz setup to facilitate the dialog may provide results that are different from those using a fully-automated ECA. Additionally, we only asked participants to retrospectively describe a pain experience from the prior two weeks. Our results may not generalize to ECA assessment of acute, ongoing pain experience.

6 FUTURE WORK

There are several research paths that our exploratory work points to. We outline three key areas for future research based on the findings and implications of our study.

Multimodal Pain Assessment should include Non Verbal Expressions: Narrative descriptions of pain experiences are multidimensional, encompassing linguistic expression, use of specific facial expressions, complex hand gestures, and other behavioral cues. Our study collected data from multiple modalities, including audio, video, Kinect, and physiological measurements. This data collection effort provides a valuable resource for furthering multimodal assessment in pain research. Future research should explore understanding the non-verbal expressions of pain, helping us gain a deeper understanding of pain experiences. In understanding the gesture correlates of pain, we can develop specialized gesture recognition algorithms that specifically identify pain-related movements.

Automated Dialog Management: Another avenue for future work is the development of automated dialog generation and management for the ECA system. Although our current study used a wizard to facilitate the dialog, leveraging recent advancements in large language models presents an opportunity to automate the dialog process. By employing natural language understanding, researchers can explore the development of systems that can more effectively engage users, adapt to their specific pain contexts, and provide tailored empathetic responses.

At-home assessment of patient reported outcomes: The rise of remote work and telemedicine, since the pandemic, has emphasized the importance of at-home systems for assessing, diagnosing, and monitoring of patients; including a focus on patient-reported pain outcomes [2]. Current pain elicitation methods rely on repeated self-report, which may challenge the ability to engage patients and maintain compliance in the context of longitudinal at-home symptom surveillance. Using an ECA as a pain assessment tool offers several potential benefits in this context. Firstly, an ECA can create more engaging and interactive user experiences, and increase user willingness to participate in pain assessments over extended periods. Secondly, when combined with automated pain detection techniques, the ECA can contribute to an understanding of pain that is closer to the user's subjective experience, improving the reliability and accuracy of pain-related outcomes.

7 CONCLUSION

In this work, we demonstrate the potential use of ECAs for automated pain assessment. Participants acknowledged the empathetic and caring nature of the agent feedback, indicating that ECAs can effectively display empathy and sympathy towards individuals experiencing pain. Additionally, participants perceived the agent as a potentially useful tool for assessing and addressing their medical issues remotely, emphasizing the convenience and effective communication offered by ECAs. This work contributes to the fields of automated assessment of pain, healthcare, and ECAs, and we identify several directions for future research in this space.

ACKNOWLEDGMENTS

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award number R01CA271145.

REFERENCES

[1] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International journal of social robotics* 1 (2009), 71–81.

[2] Margaret Bevans, Alyson Ross, and David Cella. 2014. Patient-Reported Outcomes Measurement Information System (PROMIS): efficient, standardized tools to measure self-reported health and quality of life. *Nursing outlook* 62, 5 (2014), 339–345.

[3] Timothy Bickmore, Everlyne Kimani, Ameneh Shamekhi, Prasanth Murali, Dhaval Parmar, and Ha Trinh. 2021. Virtual agents as supporting media for scientific presentations. *Journal on Multimodal User Interfaces* 15 (2021), 131–146.

[4] Timothy Bickmore, Zhe Zhang, Matthew Reichert, Clevanne Julce, and Brian Jack. 2020. Promotion of preconception care among adolescents and young adults by conversational agent. *Journal of Adolescent Health* 67, 2 (2020), S45–S51.

[5] Timothy W Bickmore, Laura M Pfeifer, Donna Byron, Shaula Forsythe, Lori E Henault, Brian W Jack, Rebecca Silliman, and Michael K Paasche-Orlow. 2010. Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. *Journal of health communication* 15, S2 (2010), 197–210.

[6] Timothy W Bickmore, Laura M Pfeifer, and Brian W Jack. 2009. Taking the time to care: empowering low health literacy hospital patients with virtual nurse agents. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1265–1274.

[7] Timothy W Bickmore and Rosalind W Picard. 2005. Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction (TOCHI)* 12, 2 (2005), 293–327.

[8] Graham D Bodie. 2011. The Active-Empathic Listening Scale (AELS): Conceptualization and evidence of validity within the interpersonal domain. *Communication Quarterly* 59, 3 (2011), 277–295.

[9] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[10] Justine Cassell. 2001. Embodied conversational agents: representation and intelligence in user interfaces. *AI magazine* 22, 4 (2001), 67–67.

[11] Justine Cassell, Hannes Högni Vilhjálmsson, and Timothy Bickmore. 2001. Beat: the behavior expression animation toolkit. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*. 477–486.

[12] Kenneth D Craig. 2009. The social communication model of pain. *Canadian Psychology/Psychologie canadienne* 50, 1 (2009), 22.

[13] David DeVault, Ron Artstein, Grace Benn, Teresa Dey, Ed Fast, Alesia Gainer, Kallirroi Georgila, Jon Gratch, Arno Hartholt, Margaux Lhommet, et al. 2014. SimSensei Kiosk: A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. 1061–1068.

[14] Levent Dumenci, Robin K Matsuyama, Laura Kuhn, Robert A Perera, and Laura A Siminoff. 2013. On the validity of the shortened rapid estimate of adult literacy in medicine (REALM) scale as a measure of health literacy. *Communication methods and measures* 7, 2 (2013), 134–143.

[15] Paul Ekman and Wallace V Friesen. 1978. Facial action coding system. *Environmental Psychology & Nonverbal Behavior* (1978).

[16] Jasper Feine, Ulrich Gnewuch, Stefan Morana, and Alexander Maedche. 2019. A taxonomy of social cues for conversational agents. *International Journal of Human-Computer Studies* 132 (2019), 138–161.

[17] Paula M Gardiner, Kelly D McCue, Lily M Negash, Teresa Cheng, Laura F White, Leanne Yinusa-Nyahkoon, Brian W Jack, and Timothy W Bickmore. 2017. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. *Patient education and counseling* 100, 9 (2017), 1720–1729.

[18] Stefanos Gkikas and Manolis Tsiknakis. 2023. Automatic assessment of pain based on deep learning methods: A systematic review. *Computer methods and programs in biomedicine* 231 (2023), 107365.

[19] Adam O Horvath and Leslie S Greenberg. 1989. Development and validation of the Working Alliance Inventory. *Journal of counseling psychology* 36, 2 (1989), 223.

[20] Lars-Christer Hydén and Michael Peolsson. 2002. Pain gestures: The orchestration of speech and body gestures. *Health:* 6, 3 (2002), 325–345.

[21] Everlyne Kimani, Timothy Bickmore, Ha Trinh, Lazlo Ring, Michael K Paasche-Orlow, and Jared W Magnani. 2016. A smartphone-based virtual agent for atrial fibrillation education and counseling. In *Intelligent Virtual Agents: 16th International Conference, IVA 2016, Los Angeles, CA, USA, September 20–23, 2016, Proceedings 16*. Springer, 120–127.

[22] Kurt Kroenke, Tara W Strine, Robert L Spitzer, Janet BW Williams, Joyce T Berry, and Ali H Mokdad. 2009. The PHQ-8 as a measure of current depression in the general population. *Journal of affective disorders* 114, 1-3 (2009), 163–173.

[23] GB Langley and H Sheppeard. 1985. The visual analogue scale: its use in pain measurement. *Rheumatology international* 5, 4 (1985), 145–148.

[24] Daniel Lopez-Martinez, Ognjen Rudovic, and Rosalind Picard. 2017. Physiological and behavioral profiling for nociceptive pain estimation using personalized multitask learning. *arXiv preprint arXiv:1711.04036* (2017).

[25] Rasha M. Al-Eidan, Hend Al-Khalifa, and AbdulMalik Al-Salman. 2020. Deep-learning-based models for pain recognition: A systematic review. *Applied Sciences* 10, 17 (2020), 5984.

[26] Chris J Main. 2016. Pain assessment in context: a state of the science review of the McGill pain questionnaire 40 years on. *Pain* 157, 7 (2016), 1387–1399.

[27] John D Mayer, Peter Salovey, David R Caruso, and Gill Sitarenios. 2003. Measuring emotional intelligence with the MSCEIT V2. 0. *Emotion* 3, 1 (2003), 97.

[28] David McNeill. 1992. *Hand and Mind: What Gestures Reveal about Thought*. Cambridge University Press, Cambridge.

[29] Ronald Melzack and Srinivasa N Raja. 2005. The McGill pain questionnaire: from description to measurement. *The Journal of the American Society of Anesthesiologists* 103, 1 (2005), 199–202.

[30] Prasanth Murali, Teresa O'Leary, Ameneh Shamekhi, and Timothy Bickmore. 2019. Health counseling by robots: Modalities for breastfeeding promotion. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–6.

[31] Prasanth Murali, Lazlo Ring, Ha Trinh, Reza Asadi, and Timothy Bickmore. 2018. Speaker hand-offs in collaborative human-agent oral presentations. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. 153–158.

[32] Prasanth Murali, Ha Trinh, Lazlo Ring, and Timothy Bickmore. 2021. A friendly face in the crowd: Reducing public speaking anxiety with an emotional support agent in the audience. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 156–163.

[33] Nicola Plant and Patrick GT Healey. 2012. The use of gesture to communicate about felt experiences. In *Proceedings of SemDial 2012 (SeineDial): The 16th Workshop on the Semantics and Pragmatics of Dialogue*, Vol. 80.

[34] Manuel Portela and Carlos Granell-Canut. 2017. A new friend in our smartphone? Observing interactions with chatbots in the search of emotional engagement. In *Proceedings of the XVIII International Conference on Human Computer Interaction*. 1–7.

[35] Erin D Reilly, Ummul-Kiram Kathawalla, Hannah E Robins, Alicia A Heapy, Timothy P Hogan, Molly E Waring, Karen S Quigley, Charles E Drebing, Timothy Bickmore, Matias Volonte, et al. 2023. An Online Acceptance and Mindfulness Intervention for Chronic Pain in Veterans: Development and Protocol for a Pilot Feasibility Randomized Controlled Trial. *JMIR Research Protocols* 12, 1 (2023), e45887.

[36] Samantha Rowbotham, Judith Holler, Donna Lloyd, and Alison Wearden. 2012. How do we communicate about pain? A systematic analysis of the semantic contribution of co-speech gestures in pain-focused conversations. *Journal of nonverbal behavior* 36 (2012), 1–21.

[37] Samantha Rowbotham, Judith Holler, Donna Lloyd, and Alison Wearden. 2014. Handling pain: The semantic interplay of speech and co-speech hand gestures in the description of pain sensations. *Speech Communication* 57 (2014), 244–256.

[38] Samantha J Rowbotham, Judith Holler, Alison Wearden, and Donna M Lloyd. 2016. I see how you feel: Recipients obtain additional information from speakers' gestures about pain. *Patient Education and Counseling* 99, 8 (2016), 1333–1342.

[39] Carolyn Saund, Marion Roth, Mathieu Chollet, and Stacy Marsella. 2019. Multiple metaphors in metaphoric gesturing. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 524–530.

[40] Lin Shi, Timothy Bickmore, and Roger Edwards. 2015. A feminist virtual agent for breastfeeding promotion. In *Intelligent Virtual Agents: 15th International Conference, IVA 2015, Delft, The Netherlands, August 26-28, 2015, Proceedings 15*. Springer, 461–470.

[41] Dina Utami and Timothy Bickmore. 2019. Collaborative user responses in multiparty interaction with a couples counselor robot. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 294–303.

[42] Matias Volonte, Andrew Robb, Andrew T Duchowski, and Sabarish V Babu. 2018. Empirical evaluation of virtual human conversational and affective animations on visual attention in inter-personal simulations. In *2018 IEEE conference on virtual reality and 3D user interfaces (VR)*. IEEE, 25–32.

[43] Timothy H Wideman, Robert R Edwards, David M Walton, Marc O Martel, Anne Hudon, and David A Seminowicz. 2019. The multimodal assessment model of pain: a novel framework for further integrating the subjective pain experience within research and practice. *The Clinical journal of pain* 35, 3 (2019), 212.

[44] Ziang Xiao, Michelle X Zhou, Q Vera Liao, Gloria Mark, Changyan Chi, Wenxi Chen, and Huahai Yang. 2020. Tell me about yourself: Using an AI-powered chatbot to conduct conversational surveys with open-ended questions. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 3 (2020), 1–37.

[45] Xiaojing Xu and Virginia R de Sa. 2020. Exploring multidimensional measurements for pain evaluation using facial action units. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, 786–792.

[46] Ghada Zamzmi, Dmitry Goldgof, Rangachar Kasturi, Yu Sun, and Terri Ashmeade. 2016. Machine-based multimodal pain assessment tool for infants: a review. *arXiv preprint arXiv:1607.00331* (2016).

[47] Zhe Zhang and Timothy Bickmore. 2018. Medical shared decision making with a virtual agent. In *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. 113–118.

[48] Zhe Zhang, Ha Trinh, Qiong Chen, and Timothy Bickmore. 2015. Adapting a geriatrics health counseling virtual agent for the chinese culture. In *Intelligent Virtual Agents: 15th International Conference, IVA 2015, Delft, The Netherlands, August 26-28, 2015, Proceedings 15*. Springer, 275–278.