# Breaking Barriers in Remote Client-Therapist Interaction: Exploring Design Spaces of Sensing and Sharing Non-Verbal Cues in Remote Psychotherapy

LAN GAO*, University of Chicago, USA
MUNMUN DE CHOUDHURY, Georgia Institute of Technology, USA
JENNIFER G KIM, Georgia Institute of Technology, USA

In remote psychotherapy, challenges arising from remote client-therapist interactions can impact the therapeutic alliance and overall outcomes. HCI research has focused on leveraging sensing technology to bridge gaps in remote interactions. In this work, we investigate the values and risks of integrating sensing technology in remote psychotherapy, specifically to capture and interpret non-verbal cues, by conducting a speculative design study with both clients and therapists. Our findings reveal that sensing technology has the potential to facilitate self-reflection in therapy. The sharing of tracked non-verbal cues could also possibly foster mutual disclosure, supporting therapists' judgments and balancing power dynamics between clients and therapists. However, clients and therapists were concerned about the accuracy of sensing systems, potential privacy threats, and additional cognition burden. Our insights into system values imply how sensing technology could potentially balance power dynamics in client-therapist relationships as well as general interpersonal relationships. We also emphasize the increased considerations in sensing-technology-empowered communication for remote psychotherapy than in non-vulnerable settings.

CCS Concepts: • **Human-centered computing → Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: Remote Psychotherapy, Internet-Based Mental Health Intervention, Computer-Mediated Communication, Sensing Technology, Remote Communication, Non-Verbal Cues

**ACM Reference Format:**
Lan Gao, Munmun De Choudhury, and Jennifer G Kim. 2025. Breaking Barriers in Remote Client-Therapist Interaction: Exploring Design Spaces of Sensing and Sharing Non-Verbal Cues in Remote Psychotherapy. *Proc. ACM Hum.-Comput. Interact.* 1, 1, Article CSCW (January 2025), 36 pages. https://doi.org/XXXXXXX.XXXXXXX

**Note:** This is a preprint that has been accepted by CSCW'25. This is NOT the final version of the formal publication.

*Performed research when studying at Georgia Institute of Technology.

Authors' addresses: Lan Gao, langao@uchicago.edu, Department of Computer Science, University of Chicago, Chicago, Illinois, USA; Munmun De Choudhury, munmund@gatech.edu, School of Interactive Computing, Georgia Institute of Technology, Atlanta, Georgia, USA; Jennifer G Kim, jennifer.kim@cc.gatech.edu, School of Interactive Computing, Georgia Institute of Technology, Atlanta, Georgia, USA.


ACM 2573-0142/2025/1-ARTCSCW
https://doi.org/XXXXXXX.XXXXXXX

## 1 INTRODUCTION

Remote psychotherapy[1] indicates the practice of conducting synchronous mental therapy sessions through internet-supported video-conferencing or audio calls. This connects clients and therapists who are physically located in different places. With the evolution of Computer-Mediated Communication (CMC) technology, remote psychotherapy has been increasingly adopted by psychotherapy clients and psychotherapists, as it provides a more convenient venue for conducting therapy and shows effectiveness in dealing with various mental disorders [25, 26, 32, 37, 69, 107].

However, researchers found numerous clients and therapists question the reliability of remote psychotherapy due to concerns about reduced engagement and information loss in a remote setting [82, 97]. Notably, in a virtual environment, people's ability to catch or infer non-verbal information could be compromised, attributing to the diminished and delayed information transmission through text, audio, and video [19]. This challenge in remote communication can limit therapist's understanding of their clients through interpreting non-verbal reactions and expressions, impeding a strong therapeutic outcome eventually [48, 91, 97, 112]. The diminished non-verbal information also threatens a solid therapeutic alliance, a deep and collaborative client-therapist relationship that positively correlates to the therapeutic outcome of clients [9, 33, 66]. Previous research revealed that in remote communication, trust among people is fragile and delayed due to ineffective non-verbal information exchanges [19]. Likewise, the diminished non-verbal signals often lead to a weak and slow feeling of feedback from both clients and therapists when interacting over a distance, rendering their dialogue indifferent and creating a feeling of disconnection [69, 86, 112]. It results in a less robust therapeutic relationship when compared to face-to-face therapy [27, 30, 52, 80, 82].

Given the challenges in remote client-therapist interaction, enhancing the remote psychotherapy experience may be achieved by directly incorporating and interpreting non-verbal information during therapy sessions. Human-Computer Interaction (HCI) researchers have developed sensing-technology-based applications to capture, enhance, explain, and exchange additional information to improve interaction among humans. Sensing technology can promote communication through aggregating **explicit non-verbal information** – human-perceptible non-verbal contents in communication, such as non-verbal expressions (e.g., eye movement, facial expression) [99] and environmental clues [23]. It can also reveal **internal vitals** that cannot be perceived by humans, which were considered as 'expressive' and utilized for supplementing non-verbal expressions in prior works [49, 77, 79, 85, 122]. In this paper, we use **non-verbal cues** to describe these two types of information incorporated in sensing-technology-empowered communication. Sensing technology has predominantly been applied in supporting remote communication, as it directly addresses the challenge of information loss in remote settings by compensating with the tracked non-verbal cues [48, 79]. Researchers have shown that sharing non-verbal cues can solidify interpersonal relationships by promoting understanding, intimacy, and empathy among people [49, 60, 68, 77–79, 99, 106, 122]. Therefore, our work aims to explore opportunities and challenges of applying sensing systems to detect and share non-verbal cues in remote client-therapist interactions.

Sensing technology has been widely used to empower communication in casual settings, such as interactions between friends and couples. However, considering the potential invasion sensing systems bring and the fragile nature of therapy, the values and risks of applying sensing systems in remote psychotherapy should be further investigated. As the tracked data is personal, private, and holds contextual-relied meanings, prior works address the importance of studying the value sensing

[1]There is no consensus on the terminology of the internet-based psychological therapy [108]. "Teletherapy" commonly refers to mental health therapy over the phone or online. However, to specify the modality as internet-based, we use the term "Remote Psychotherapy" to create a clearer distinction.

and sharing process brings before incorporating it in different contexts [49, 83, 85]. Thus, it is important to understand clients' and therapists' attitudes toward how non-verbal cues should be tracked, explained, and shared to support client-therapist interaction. Moreover, remote psychotherapy is rigorously regulated regarding personal privacy and ethics, given the client's vulnerability in therapy and the data sensitivity of digital health services [8, 50, 81]. Therefore, potential risks should be carefully investigated when incorporating additional technology support in remote psychotherapy.

In this study, we investigate clients' and therapists' thoughts on leveraging sensing systems to facilitate client-therapist interaction in remote psychotherapy. Specifically, we focus on the values of tracked non-verbal cues to support the therapeutic alliance and outcomes, as well as concerns about personal and social risks. To this end, we ask:

- **RQ1:** *(Values)* How can tracked non-verbal cues, including non-verbal communication and vitals, be utilized to support client-therapist interactions?
- **RQ2:** *(Risks)* What potential risks can be introduced by the sensing and sharing process?

We conducted a speculative design study through scenario-based interviews with psychotherapy clients and psychotherapists. In these interviews, we explored their thoughts on potential approaches, values, and risks of sensing and sharing non-verbal cues during challenging situations in remote psychotherapy. Scenario-based interviews have been frequently adopted in previous empirical HCI studies to understand user perceptions and design norms of technology systems in novel socio-technical settings [36, 92]. Through interviews with 9 clients and 10 therapists, we found that clients and therapists envision the possibility of using the monitoring of their personal non-verbal cues as a means of self-reflection during therapy. They also perceived the potential benefits of sharing non-verbal cues in enhancing mutual disclosure, which not only could offer supplementary information for therapists' assessments but also could foster a deeper mutual understanding and balance the power dynamics within their relationship. However, our findings also revealed clients' and therapists' considerations on potential risks associated with sensing and analysis accuracy, personal privacy and system security, and increased cognition burden. We further discuss the impacts of sensing and sharing non-verbal cues on the client-therapist relationship and other interpersonal relationships. Additionally, we highlight the concerns about deploying sensing systems in vulnerable contexts like remote psychotherapy, compared to those in non-vulnerable or causal contexts like close relationships.

Our work contributes to digital mental health research by exploring the potential and risks of applying technology to facilitate client-therapist interaction, with the ultimate goal of promoting remote psychotherapy experiences. Moreover, our work extends recent literature in HCI of sensing-technology-empowered communication by investigating its potential application in a vulnerable social setting: remote psychotherapy. Taking remote psychotherapy as a representative case, our work also provides insights into end-users' different perceptions of sensing technology for communication in vulnerable settings and non-vulnerable ones.

## 2 RELATED WORK

### 2.1 Research Context: Client-Therapist Interaction and Therapeutic Alliance in Remote Psychotherapy

Remote psychotherapy, known as one of the most prevalent methods of internet-based psychological interventions, has long existed since the emergence of remote technology. Multiple meta-reviews have analyzed the development of remote psychotherapy, showing how the proliferation of CMC incentivizes its promotion [25, 26, 37, 107]. Prior works show that remote psychotherapy can be effective in various types of mental disorders, including anxiety [3, 15, 73], depression [3, 15, 73], trauma [53, 54], and anorexia [2]. Potential benefits have been found in remote psychotherapy,

such as reducing commuting and related cost [73], mitigating the embarrassment of clients [30], and serving marginalized populations [105] or people from under-served regions [16, 32].

Regardless of its benefits, numerous previous works pointed out the challenges of remote psychotherapy, specifically in remote client-therapist interaction. The interaction barriers in remote client-therapist interaction are similar to well-known challenges in typical remote communication such as information loss, transition delay, and tendency to multitask [19, 24]. In terms of clients, it is common to be distracted by things happening around them, such as receiving an email and multi-tasking [82]. According to an empirical study conducted by Robledo Yamamoto et al., young clients can be easily distracted in online therapy owing to their developing cognitions [97]. Prior studies also show that clients sometimes resist expression in remote psychotherapy, partly led by their unfamiliarity and untrustworthiness with online communication [114] or a poor therapeutic alliance [82]. In terms of therapists, it is difficult for them to determine information beyond verbal cues via remote communication [97, 112]. Robledo Yamamoto et al. pointed out that the lack and loss of non-verbal cues can be caused by poor video quality or audio-only situations, making it harder for therapists to get a deeper understanding of their clients [97]. Lacking non-verbal cues also leads to difficulties in conducting special types of therapy in which therapists heavily rely on non-verbal cues, such as Sensorimotor Psychotherapy [91], within a remote environment.

There is a broad agreement that the outcome of psychotherapy can be significantly affected by therapeutic alliance [45, 65, 66, 121], making it an important role in ensuring the quality of psychotherapy. However, the based verbal cues and diminished non-verbal cues may compromise the solidity of trustworthiness in remote communication [19], which is also applicable in remote psychotherapy. Unsurprisingly, the robustness of the therapeutic alliance in remote psychotherapy has been called into question by researchers, psychotherapy clients, and psychotherapists. Prior research reveals that therapists especially have reservations about the strength of therapeutic alliance in remote psychotherapy [27, 52, 80]. Lopez et al. found that concerns about the weak therapeutic alliance could lead to a low willingness to take psychotherapy online [80]. Chen et al. perceived a paradoxical sense of clients that although online interaction makes them feel safer, they identify their therapists as colder and more indifferent [30]. In practice, several studies reveal that a feeling of disconnection has been detected from both therapist's side [86, 112] and the client's side [69, 112], which can lead to less intimacy in the client-therapist relationship than it is in face-to-face therapy [82]. By investigating multiple practitioners' personal experiences, Russell argued that a screen relationship can hardly ever substitute for a real in-person relationship in psychotherapy [101].

To overcome challenges that arise from remote client-therapist interaction, researchers have proposed suggestions on how to enhance communication and trust in remote psychotherapy, usually regarding what therapist could do [28, 82, 97]. However, we found limited research that explores the potential of using assistive tools to mitigate the interaction barriers in remote psychotherapy. Meanwhile, digital health approaches have already been applied in psychotherapy. Technology intervention in therapy usually targets therapy experience improvement, such as motivating client's engagement or evaluating the therapy quality [64, 74]. The adoption of remote technology, including videoconferencing and web-based calls, for the delivery of psychotherapy is also well-established [25, 26, 37, 107]. However, additional technological support to facilitate remote interactions between clients and therapists is relatively limited.

Within this background, we regard sensing systems, widely used to empower interpersonal communication by HCI researchers, as potential assistive tools to support client-therapist interactions. Specifically, our work investigated the potential system deployments through a speculative design study.

### 2.2 Sensing and Sharing Non-verbal Cues in Interpersonal Communication

Psychologists have investigated the indispensable role of non-verbal communication. According to Mehrabian's model of communication, body language can account for more than half of the expression cues when interpreting the complete meaning of a message [84]. Several works have revealed people's tendency to interpret non-verbal cues to understand the true attitude of the other person during communication, which is especially helpful if there are conflicts between verbal and non-verbal expressions [5, 6, 94]. Non-verbal communication has been shown to be important in various social scenarios, such as classroom and workplace [18, 88].

Regarding the importance of non-verbal information in interpersonal communication, researchers have developed applications based on pervasive sensing and computing systems, to track, analyze, and interpret non-verbal cues to enhance the communication experience. For example, eye-tracking systems, beyond their use as research tools, have been employed to supplement remote communication and collaboration [87, 117], especially for people with limited verbal abilities [119]. Affective computing techniques, such as facial expression analysis and sentiment analysis, have also been widely applied to enhance emotional understanding in communication [99]. The potential of tracking and sharing environmental information in remote communication was also explored by researchers [23]. Moreover, sensing systems can also interpret implicit non-verbal information that people are unlikely or unable to detect on their own, such as micro-movements and biosignals. Researchers also recognized the potential of revealing and sharing this type of tracked information as 'expressive social cues' in communication [49, 77, 79, 85].

Prior works have already proved that sharing non-verbal cues tracked by sensing systems can promote empathy building [60, 78, 99, 106, 122], intimacy building [68, 70, 77, 79], social interaction [39, 76, 106], and collaboration [34]. Most of the systems work in a co-presence-required or asynchronously remote communication context, while there are also emerging applications showing how real-time sensing and sharing benefits synchronous remote communications. For instance, Rojas et al. investigated the potential of videoconferencing to enhance empathy by creating a videoconferencing application featuring real-time emotion cues as feedback. This approach led to increased awareness of others' emotions and a greater willingness to self-disclose [99]. Faucett et al. introduced a system that senses and provides real-time feedback on clinicians' speaking attributes in video telehealth. It resulted in successful improvements in clinicians' behavior, concentration, and self-reflection during clinician-patient communication [48].

Researchers also raised concerns about using pervasive sensing to support communication. First, the sensing technology itself might not perfectly track or interpret non-verbal cues, which complicates the communication process. For example, the accuracy of emotion detection and analysis has been identified as one of the biggest challenges of applying affective computing to communication support [41]. Thinking of incorporating physiological information in communication, Feijt et al. argued that the meaning of such information is highly context-dependent, leading to an ambiguity of interpretation by machines and humans [49]. Moreover, information collected by sensing systems could be treated as private by people. Prior works have discussed people's resistance to sensing systems in social spaces because of the feeling of intrusion and surveillance [35, 36, 83]. In terms of the sharing process, Ethical concerns of sharing tracked information in communication, such as privacy concerns and unwilling reveals, were also identified by researchers [23, 49, 85]. Moge et al. advocated researchers pay attention to how social norms change in sensing-technology-empowered communication compared to typical communication [85].

We have identified wealth attempts to share non-verbal cues through technology supports in causal contexts, such as communications between friends or couples. However, there is scarce research work exploring the opportunities for deploying sensing technologies in vulnerable settings

like remote psychotherapy, in which the system's potential usages and risks need to be re-evaluated. Our work extends the exploration of sensing-technology-empowered communication in the context of remote psychotherapy, inquiring about the system's potential values and risks by investigating clients' and therapists' perspectives. The implications from our work reveal users' different attitudes to sensing-technology-empowered communication in vulnerable settings compared to non-vulnerable ones.

## 3 METHODOLOGY

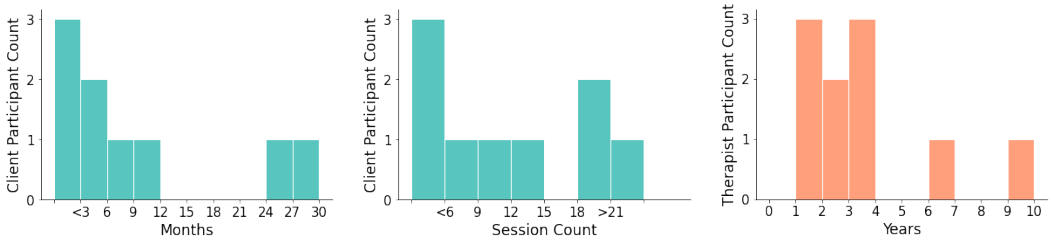
To explore clients' and therapists' expectations of sensing and sharing non-verbal cues during challenging situations in remote psychotherapy, we conducted a speculative design study with both psychotherapy clients and psychotherapists by taking the approach of scenario-based interviews. All sessions were conducted online via Zoom individually. During the interview, all participants completed the same procedure, with each session lasting no longer than 90 minutes. All the study procedures were approved by the Georgia Institute of Technology IRB (Protocol No. H22459).

We acknowledge that conducting a field study using actual sensing technology and data would provide the most accurate insights into the behaviors and perceptions of remote psychotherapy therapists and clients. However, we identified several challenges in deploying multiple sensing systems in real-world remote psychotherapy for our study. Psychotherapy has strict requirements in conducting conditions and client information protection to ensure clients' rights [8, 50]. Collecting and analyzing additional data in real therapy without preliminary compliance evaluation can have immeasurable and irretrievable impacts on clients, complicating the accountability of therapy eventually. This hazard can also be intensified by the technical difficulties of tracking and analyzing. Therefore, we opted for an alternative approach, conducting interviews with scenario-based activities in which participants speculated how sensing and sharing non-verbal cues could or could not address challenges presented in particular scenarios. Scenarios are one of the most common tools for design approaches, enabling experts and users to envision future socio-technical technology [100]. Leveraging scenarios that describe system use has been shown to uncover user behavior with novel technology systems [125], reveal possible use cases early on in system development [123], and highlight social values and ethical impacts for future designs [124]. Specifically, the in-scenario interview approach has been widely used for evaluating new socio-technical settings, providing insights into user-centered design requirements by revealing user perceptions of technology use [36, 92].

### 3.1 Participants

*3.1.1 Recruitment.* We recruited 9 psychotherapy clients and 10 psychotherapists from China Mainland. Our inclusion criteria for client participants required a minimum of one course of treatment through remote or hybrid (conducted partly in-person and partly remote) psychotherapy, lasting more than 2 months with at least 4 online sessions attended. Our inclusion criteria for therapist participants required a valid license conducting psychotherapy, a minimum of 1 year of both professional experience and previous experience conducting remote psychotherapy. We only recruited participants who were over 18 years old.

All client participants were recruited by advertising in verified psychological communities (N=6) or snowball sampling starting from authors' personal and professional networks (N=3). We employed this strategy for client recruitment because verifying the authenticity of the client's therapy-receiving experience is challenging. To reduce the risk of fake self-reports, the first author initiated the recruitment by reaching out to individuals known personally, whose experiences could be reliably confirmed. All therapist participants were recruited by contacting the licensed psychotherapist contact lists (N=1) and advertising in verified psychological communities (N=6), or




(a) Time Distribution of Client Participants Receiving Remote or Hybrid Psychotherapy.

(b) Session Count Distribution of Client Participants Receiving Remote Psychotherapy.

(c) Time Distribution of Therapist Participants Providing Remote Psychotherapy.

Fig. 1. Participant's experience in remote psychotherapy.

directly contacting via authors' personal and professional networks (N=3). To minimize the sample bias of snowball sampling, safeguard clients from the mental health stigma of familiars, and avoid ethical issues in therapy relationships, we ask participants to avoid soliciting potential participants from families, close friends, or clients and therapists with whom they have collaborated.

*3.1.2 Participant Demographic.* There were 3 male and 6 female client participants. The gender distribution of our client participants corresponds with the typical gender patterns observed in therapy attendance [93]. Our client participants had varied experiences in engaging in remote psychotherapy (i.e. from 2 months, 5 times to 2.5 years, over 125 times). They reported a wide range of mental concerns and claims for the therapy including anxiety, depression, self-exploration, and so on. All client participants identified themselves as participating in Talk Psychotherapy. While all client participants possessed prior experience with remote psychotherapy, 7 out of 9 had experience in receiving both remote and in-person therapy. Conversely, 2 out of 9 had exclusively undergone remote therapy without prior in-person therapy experience.

There were 1 male and 9 female therapist participants. The predominance of females in our therapist participants matches a gendered skew in the psychological profession [109]. Our therapist participants represent a wide range of professional experiences in providing psychotherapy services (i.e. from 2 years to 14 years) and conducting remote psychotherapy (i.e. from 1 year to 9 years). Their expertise varied from general psychoanalysis to specific mental disorders such as anxiety, depression, and interpersonal relationship issues; and specific client types such as kids, teenagers, and family. In the remote condition, all of the therapist participants provided Talk Psychotherapy, spanning Cognitive Behavioral Therapy (CBT), Interpersonal Therapy (IPT), Psychodynamic Therapy, and Family Therapy. One therapist participant also provided Somatic Psychotherapy. Only one therapist participant provided remote psychotherapy exclusively, while the others (N=9) conducted psychotherapy both online and in person. All therapist participants had experience receiving supervision or personal therapy, in which they experienced therapy from the client's perspective. Note that during our study, two therapist participants (T5, T9) volunteered to share their perspectives on receiving psychotherapy as clients for their personal mental health needs[2].

We visualize the statistics of participants' experience in remote psychotherapy in Fig. 1. We also include participants' detailed demographic information in Table 3 (Appendix A.1).

[2]T5 and T9 were identified as therapist participants because most of their reports are from the therapist perspective. However, their insights from the client's perspective were also analyzed and added. In the subsequent Findings section, we distinguish and highlight their statements accordingly.

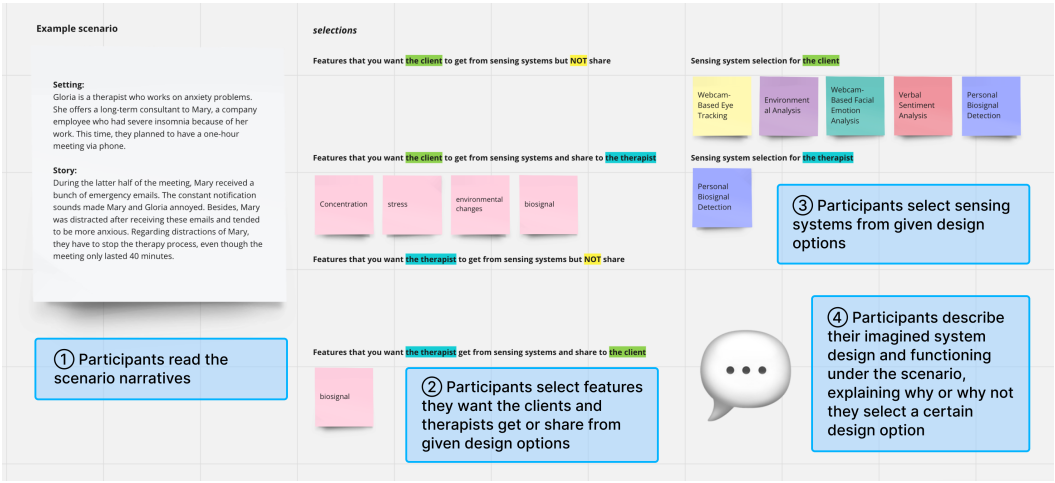



Fig. 2. In-scenario activity procedure (Taking *Example Scenario* as an example).

*3.1.3 Context Statement: Psychotherapy and Remote Psychotherapy in China.* With an increasing rate of mental disorders over 30 years [67], psychotherapy services are improving constantly in China [22]. Notably, the remote psychotherapy industry is also growing rapidly. A report shows that the number of online therapy platforms in China increased from 1 to 64 between 2014 and 2016 [22]. During the Covid-19 pandemic, therapists from China indicated more experience and preparation for remote psychotherapy, compared to therapists from other regions [31].

However, several challenges are identified in conducting psychotherapy in China. First, the regulations governing therapy have a short history, with China's first Mental Health Law established in 2013 [11, 22]. Given the short development time, the integrity and compliance of therapy-providing policies often remain uncertain [75]. The complex Chinese culture and lower awareness of therapy ethics also block the implementation of ethical standards, such as informed consent and confidentiality [1]. Also, there is a lack of regulatory supervision of remote psychotherapy [95]. Moreover, the relatively conservative Chinese society leads to a high rate of mental health stigma, making people reluctant to seek professional support [22, 126].

## 3.2 Materials Design for the In-Scenario Activities

An overview of the in-scenario activities in each scenario is shown in Fig. 2. For each **Scenario** (Fig. 2 ①), participants were asked to select sensing systems and their corresponding outputs from provided **Sensing System Design Options** (Fig. 2 ②③), then explained their imagined system design and how it could work in this situation (Fig. 2 ④). In the following section, we present the materials we used during the in-scenario activities. All the materials shown to the participants were originally in Simplified Chinese. Before the formal study, we conducted 3 pilot sessions with 2 clients and a therapist to ask about their opinion of the material design, after which we slightly revised the materials based on their feedback.

*3.2.1 Scenarios: Challenges in Remote Psychotherapy.* We designed five textual scenarios describing potential emergent use–cases of sensing systems. When brainstorming the scenario design, we considered problematic situations in therapy more representative and did not design a scenario representing a typical therapy session. As shown in Table 1, each scenario describes one fictional situation corresponding to one of four key challenges that could arise from remote client-therapist

interaction. The challenges in remote psychotherapy identified from existing studies and reviews around remote interactions [82, 97, 114]. The scenario narratives in *Scenario 1 (Distraction)*, *Scenario 2 (Lack of non-verbal cues)*, and *Scenario 3 (Self-overlooked therapist burnout)* were adapted from the excerpts of participant responses reported in Robledo Yamamoto et al.'s interview study about therapist's practices conducting therapy online [97]. All quotes we chose directly described real experiences in formal therapy. We then rephrased the quotes and blurred identifiable information. Also, *Example Scenario* and *Scenario 4 (Lack of trustworthy and safety)* were crafted by expanding on descriptions of real-world remote psychotherapy cases where challenges occurred, as reported in previous literature [82, 97]. Each scenario narrative contains the following two aspects. First, we included background settings about the client, the therapist, and the therapy. Second, a story describing the difficult situation was illustrated. The comprehensive narratives for each scenario can be found in Table 5 (Appendix A.2).

*3.2.2 Sensing System Design Options: Sensing systems and Corresponding Outputs.* We showed participants six types of sensing systems that can be potentially used in remote psychotherapy. We also presented five features – properties of tracked objects that the systems can interpret and output from tracked non-verbal cues. The six systems and five features served as options for participants to describe their envisioned socio-technical setting and system design in each scenario.

In this study, we narrowed the definition of "sensing system" to those used in daily life, only focusing on sensor-enabled technologies in common devices. Conversely, we opted for a broader definition of "sensors", including any receivers capable of capturing non-verbal cues, such as a webcam. We explored possible systems for client-therapist remote interaction support from previous literature and real applications. Eventually, we adopted *Sys 1 (Webcam-Based Eye Tracking)*, *Sys 2 (Webcam-Based Facial Expression Analysis)*, *Sys 4 (Verbal Sentiment Analysis)*, and *Sys 5 (Personal Biosignal Detection)* from prior works of real-time sensing and sharing for interaction support [48, 77, 99] and engagement evaluation [104]. We also opted for *Sys 3 (Accelerometer-Based Tremor Detection)* and *Sys 6 (Environmental Analysis)* based on the proposal by Robledo Yamamoto et al. on the potential application of sensing technology in remote psychotherapy [97], which also advocated the utilization of *Sys 4 (Verbal Sentiment Analysis)*. We found that sensing systems usually present interpreted features of tracked objects (i.e. emotion) to the end user rather than the raw data of non-verbal cues (i.e. facial expression). Therefore, we looked into input and output metrics that previous research or applications (cited above) used in their system design, identifying which features each system could interpret from the tracked non-verbal cues. We finally identified five different features that the sensing systems could interpret directly or infer indirectly by algorithm from non-verbal cues. The six systems and five corresponding features can be found in Table 2.

Additionally, we created an explanation for each system, which was shown to participants during the interview to provide them with an informative foundation of sensing technology. The textual introduction was written referring to works mentioned above as well as official handbooks of related smart devices, sensors, and real applications [4, 13, 96, 110]. Each introduction contained reading materials outlining the i) sensor(s) in the sensing system, ii) device(s) that support the system execution, iii) how the system works, including input non-verbal cues of the system, and iv) usages of the system, including output features the system computed from inputs. Complete reading materials for each sensing system can be found in Table 4 (Appendix A.2).

### 3.3 Study Procedure

Before the study, participants completed a screening survey that asked for demographic information and qualification confirmation. Afterward, participants were invited to sign the consent form and

Table 1. Scenario and Corresponding Challenges (Detailed narratives and reference are shown in Table 5)

| Scenario Code | Corresponding Challenge |
|---|---|
| *Example Scenario* | Distraction |
| *Scenario 1* | Distraction |
| *Scenario 2* | Lack of non-verbal cues |
| *Scenario 3* | Self-overlooked therapist burnout |
| *Scenario 4* | Lack of trustworthy and safety |

Table 2. Sensing Systems and Features Revealed from Tracked Non-verbal Cues (Detailed introductory materials and reference are shown in Table 4)

| | Code | Feature Name | |
|---|---|---|---|
| **Features Revealed From Non-Verbal Cues** | *Feature 1* | Concentration | |
| | *Feature 2* | Emotion and Sentiment | |
| | *Feature 3* | Stress | |
| | *Feature 4* | Biosignal Rate (i.e. Heartbeat, Blood Pressure) | |
| | *Feature 5* | Environmental Change | |
| **Sensing System and Corresponding Features** | **Code** | **Sensing System Name** | **Detectable Feature(s)** |
| | *Sys 1* | Webcam-Based Eye Tracking | Direct: N/A<br>Indirect: *Feature 1* |
| | *Sys 2* | Webcam-Based Facial Expression Analysis | Direct: *Feature 2*<br>Indirect: *Feature 1, Feature 3* |
| | *Sys 3* | Accelerometer-Based Tremor Detection | Direct: N/A<br>Indirect: *Feature 3* |
| | *Sys 4* | Verbal Sentiment Analysis | Direct: *Feature 2*<br>Indirect: *Feature 3* |
| | *Sys 5* | Personal Biosignal Detection | Direct: *Feature 4*<br>Indirect: *Feature 3* |
| | *Sys 6* | Environmental Analysis | Direct: *Feature 5*<br>Indirect: N/A |

then join in the interview session conducted by the first author. At the beginning of the session, we asked participants to give verbal consent of video and audio recordings during the session.

*3.3.1 Introductory Interview.* During the introductory interview phase, participants were encouraged to thoroughly contemplate the context of remote psychotherapy. This phase aimed to help them establish an initial understanding of how incorporating non-verbal cues sensing might address challenges in remote client-therapist interaction. It could preclude potential confusion in the following in-scenario activities. For both client and therapist participants, we first asked them about their experiences with remote psychotherapy. Then, we asked our client participants about their overall perceptions of therapeutic outcomes from remote psychotherapy and their experience interacting with therapists remotely. For therapist participants, we asked about their thoughts on delivering therapeutic intervention remotely and their experience interacting with clients. For the participants who have experience in both online and in-person therapy, we asked them to compare their perspectives on remote psychotherapy with in-person therapy. Considering the mental health stigma and to safeguard participants' privacy, we explicitly communicated to all participants that they could decline to answer any questions.

*3.3.2 In-Scenario Activities.* The goal of the in-scenario activities is to explore participants' perceptions of the idea of sensing and sharing non-verbal cues to support remote client-therapist

interaction in therapy sessions. During this phase, we presented all materials related to our study on Miro[3], an online collaborative whiteboard. First, we introduced participants to the sensors commonly found in everyday devices and the concept of sensing systems. We also introduced our design idea of deploying non-verbal cues sensing and sharing to support remote client-therapist interaction. After that, we asked them to read the explanation of each sensing system. During their review, participants were encouraged to ask questions related to sensing technology. They had the option to revisit the explanation about sensing systems or ask questions anytime during the following activities. Participants were encouraged to share their initial thoughts on the sensing systems and the idea of using those systems in remote psychotherapy.

Next, after presenting an example within the *Example Scenario*, participants were asked to complete the in-scenario activities under all four scenarios (*Scenario 1-4*) in Table 1. As shown in Fig. 2, the activity for each scenario was as follows: participants first read a narrative describing a situation in remote psychotherapy, which they were informed had occurred in actual therapy sessions. They were asked to imagine how to track and share non-verbal cues to support client-therapist interaction under the presented situations, given the goal of promoting therapy experience (Fig. 2 ①). We encouraged participants to start by assuming they were the role in the scenario, either as a client or a therapist, depending on their identity.

Then, participants were asked to think aloud about whose and which non-verbal cues could be tracked, interpreted, and shared. To describe their thoughts on the sensing and sharing pattern, they would select from the five feature options for four design decision categories: "Features you want the client to get from sensing systems and share", "Features you want the client get from sensing systems but NOT share", "Features you want the therapist to get from sensing systems and share", and "Features you want the therapist to get from sensing systems but NOT share" (Fig. 2 ②). After that, participants would think about sensing systems the client and the therapist could use within the scenario. They were asked to consider not only the ability of each system to capture and interpret non-verbal cues but also their satisfaction and concerns about the system mechanism. To reflect their thoughts on the system design, they would select six system options for two design decision categories: "Sensing system selection for the client" and "Sensing system selection for the therapist" (Fig. 2 ③). Participants dragged digital sticky notes to each design decision category in Miro to present their selections. The purpose of the design decision selection tasks was to provoke participants' deep thoughts on the values and risks of the sensing systems in context, rather than comparing which system or non-verbal cue was better.

When filling out design decisions, participants could use each design option one time, multiple times, or zero times. Moreover, participants were told that the use of sensing systems was optional. In other words, they could deny applying any tracking or sharing processes to the scenario, if they hold significant concerns or think that sensing aids were unnecessary here. For example, if they believe that no sensing system was needed for the therapist in the scenario, they did not need to choose anything for the categories "Features you want the therapist to get from sensing systems and share", "Features you want the therapist to get from sensing systems but NOT share", and "Sensing system selection for the therapist".

After finishing the activity in each scenario, participants were asked to describe their imagined system design and explain why or why not they made a certain selection before continuing to the next scenario (Fig. 2 ④).

After all four in-scenario activities, we asked participants about their overall feelings about sensing and sharing non-verbal cues in typical remote psychotherapy sessions. Specifically, participants

[3]https://miro.com/

were encouraged to share their opinions on which situations they found sensing systems most valuable, as well as which situations they didn't want to be tracked by the systems.

### 3.4 Data Analysis

Our data collection consists of two parts: study session records (qualitative data), and participants' design decision selections for four scenarios (quantitative data). Our main analysis reported in the present work is the qualitative analysis of the study session records. We include a response summary of every design decision inquiry for each scenario in the Appendix (Section A.3).

To perform qualitative data analysis, all 19 study sessions were conducted in Mandarin, the records of which were later transcribed into Simplified Chinese. All excerpts were translated into English by the first author, who is an English-Chinese bilingual. We then conducted an iterative analysis using open coding associated with the research questions. The analysis execution is as follows: First, the first author went through five transcriptions and developed codes for concepts in the text. They then categorized and merged these open codes and built an initial codebook using existing codes. After establishing the initial codebook, all members of the research team passed through the initial codes. With modification when disagreement occurs, the authors finally reach a consensus on the structure of the initial codebook. Then, the first author utilized the refined initial codebook to code the remaining transcriptions. During this process, all members of the research team regularly checked and discussed emerging themes. All members of the research team eventually ensured that they agreed on all the codes of the codebook. After the first round of coding, the first author iterated another round of open coding referring to the final codebook. All members of the research team passed the second round of coding results and discussed the codes until everyone agreed with the final analysis.

We eventually identified three key themes for **RQ1**, and three key themes for **RQ2**. We present the key themes, following our research questions, in Section 4.

## 4 FINDINGS

### 4.1 The Values of Tracked Non-verbal Cues: Enhanced Personal Reflections and Mutual Understandings

Overall, both client and therapist participants envisioned tracked non-verbal cues helpful in all four scenario we provided. They considered tracked non-verbal cues valuable for facilitating personal reflection and mutual understanding, especially when facing problems or challenges during remote psychotherapy. Sensing systems that can reveal nuanced changes were especially appreciated by participants, such as *Sys 3 (Accelerometer-Based Tremor Detection)* that were capable of detecting micro hand-shaking, and *Sys 5 (Personal Biosignal Detection)* that revealed vital signs that were not perceivable by the human eye. In the following section, we present participants' perceived value of tracked non-verbal cues in detail.

*4.1.1 Self-Monitoring Personal Cues for Self-Reflection.* Sensing systems could potentially facilitate self-reflection and self-understanding. Particularly, our participants envisioned self-tracking as an assistive tool for therapists to reflect on their work performance (4/9 Clients; 10/10 Therapists). To deliver high-quality mental health services, therapists must not only understand their client's needs but also be aware of their own states. This involved reflecting on their own performance during the session, which helped them hold up in their interaction with clients.

As psychological professionals, therapists should stay neutral in response to their client's statements and actions [115]. However, as T4 indicated, avoiding subjective expressions could be *"easy to overlook"*. Therefore, therapist participants found a potential value of self-reflection on non-verbal cues as helping them maintain a professional demeanor. For instance, T2 chose to track their *Feature*

*2 (Emotion and Sentiment)* if they had been the therapist in *Scenario 1 (Distraction)* since *"if I rarely encounter such cases, my emotional changes would be harder to control."* Considering the situation in *Scenario 4 (Lack of trust and safety)* that the therapist failed to comfort their client, T8 indicated that "*the therapist's emotions and stresses may undergo significant changes*". They explained that tracking such changes *"might not have any effect during this therapy session"*. However, they noted that it might help reflect in retrospect: *"It can assist the therapist in self-reflection after the session ends, providing a judgment like I tend to show lots of stress in these situations, indicating a prospect where they need to grow."*

Therapists could also leverage tracked cues as an alert for overall bad well-being and exhaustion, as our participants imagined. T8 explained that *"It is necessary for the therapist to reflect themselves like, am I in a good state? [...] However, if there is external data support, they may find it easier to convince themselves (that they are not in a good state)."* Client participants indicated that their therapists could potentially employ self-tracking methods to detect any decline in their mental condition, viewing this practice as a responsible approach to delivering professional mental health intervention. For example, C2 recommended the stressful therapist in *Scenario 3 (Self-overlooked therapist burnout)* to capture *Feature 3 (Stress)* and *Feature 4 (Biosignal)*, because *"He (the therapist) had a very busy schedule, so he could monitor his stress. I think he would react physiologically to his tiredness, so he could also monitor his biosignal. I think he can understand his situation by tracking this."*

Moreover, our participants speculated that clients could also track themselves to support their interaction with therapists as well as self-improve simultaneously, potentially aligning with the therapy goal (4/9 Clients; 1/10 Therapists). Imagining themselves as the nervous client in *Scenario 4 (Lack of trustworthy and safety)*, C8 was willing to monitor *Feature 1 (concentration)* to prevent themselves from becoming overly self-centered in client-therapist interaction: *"When I find it hard to control my emotions, I want to remind myself, for instance, not to concentrate on past events that are traumatic for me, but perhaps to focus on the present or the other person's reactions. So monitoring concentration could be a reminder."* C4, imagining themselves as the client in *Scenario 1 (Distraction)*, chose to be tracked by a combination of *Sys 1 (Webcam-Based Eye Tracking)*, *Sys 2 (Webcam-Based Facial Expression Analysis)*, *Sys 4 (Verbal Sentiment Analysis)* and *Sys 6 (Environmental Analysis)*, indicating a need for a self-reflection on personal reaction during communication with the therapist: *"[...] Tracking where my eyes are looking when I'm distracted. I need to know what factors in my environment might distract me. [...] It is necessary to understand how my surroundings impact my behavior and emotions."*

#### 4.1.2 *Clients Sharing their Cues with Therapists to Support Therapists' Judgement.*

> *When conducting therapy sessions in person, I can see the clients with my own eyes, hear their voices with my ears, and even feel their feelings. However, I can't genuinely perceive all these things in virtual. Facial expression analysis and biosignal detection are like placing my eyes, nose, and ears in front of the clients. They help me collect information and assist me in making a judgment.* (T8)

One of the most significant potentials of incorporating sensing systems, as agreed upon by all client and therapist participants, was improving the capacity for collecting valuable information. Therapist participants talked about the possibility of gaining a deeper understanding of clients with the support of sensing systems. As T8 noted, receiving non-verbal cues from clients might simulate the experience of being physically present with them, allowing him to assess their mental state by referencing their immediate expressions. Additional information provided by sensing systems might also assist therapists in offering more effective and personalized interventions. As C4 suggested, *"To help someone alleviate their tension, you need to know how intense their tension is."*

During therapy sessions, therapists often analyze their clients continually, drawing inferences from their every action to guide the therapeutic process. In this continuous analysis, all of our participants reported that the tracked cues from clients could potentially offer supplementary information for therapists, helping them make a holistic assessment. Therapists noted that micro changes of clients, such as *micro-sentiment changes apart from normal sentiment* (T4), *a moment of eye closing* (T4), and *performance of nervous* (T8), were easily being overlooked. Some participants also stated that certain cues, such as biosignals (C3, T10, T11), were not identifiable without technological assistance. These cues could play a crucial role in therapists' evaluations of their clients, especially when therapists were uncertain about their judgment. For example, clients might act differently from their thoughts and feelings, making the therapists' judgment process more challenging. As C8 described their experience of noticing their therapist distracted in their first meeting, *"I didn't express my dissatisfaction (verbally) because it was the first meeting."* They thought *Sys 2 (Webcam-Based Facial Expression Analysis)* possibly worked in this situation to help them signal their feelings without verbal expression, *"I hope he (therapist) realizes I'm quite unhappy right now."* T5 also envisioned that seeing clients' cues could help reduce therapists' stress and mental load, *"(Clients and therapists) might have some misunderstandings. The therapist may feel like the client does not trust them. However, if the system shows that the client is just too nervous, it will alleviate therapists' pressure."*

Additionally, both client and therapist participants conjectured that therapists could store clients' tracked cues after every session and aggregate them for long-term analysis (4/9 Clients; 4/10 Therapists). T3 suggested these tracked cues could *"be a part of the process note"*, from which therapists might *"use quantitative data to observe whether the client has become more stable"*. C7 also pointed out that *"there would be a large quantity of information exchange in a session. Therapists may have limited ability to handle it"*. They expected the therapists *"could find out some breakthrough points by retrospecting (tracking records) when encountering a bottleneck in therapy"*.

Participants also recognized the potential benefits of this sharing pattern in spontaneous situations, such as when therapists experience fatigue, leading to reduced responsiveness (C6, C7, C8), and when therapists need to make quick judgments (C1, T2). Specifically, in extreme situations where direct information gathering is difficult, therapists might still obtain valuable insights through the tracked cues from their clients. For instance, T2 noticed that the therapist in *Scenario 2 (Lack of non-verbal cues)* might not be able to get useful information from the video. They then suggested the client opt for *Sys 4 (Verbal Sentiment Analysis)* since *"client's emotional changes are still identifiable through their voice"*.

In psychotherapy, silence – where there is no verbal communication between clients and therapists – is a common situation. Therapists often perceive silence as an opportunity, and even a tool, to explore clients' mental models by observing their non-verbal reactions like facial expressions and micro-movements, which may differ from those in regular conversations and convey additional information [63, 102]. Using sensing systems, such as *Sys 2 (Webcam-Based Facial Expression Analysis)*, to analyze those non-verbal expressions could potentially assist therapists in this process. For example, T7 explained a case where extended silence persisted across 17 therapy sessions, suggesting that sensing systems could uncover underlying changes in non-verbal cues: *"I know there was a therapist and a client that had maintained silence for over 17 sessions. They hadn't spoken even a single word. [...] Information changes are happening in silence. I think it might be interesting to incorporate sensing systems and provide analysis in this case."*

#### 4.1.3 Therapists Sharing their Cues with Clients for an Equal Mutual Disclosure and Power Dynamic.

> *Since the client is willing to provide these (tracked information) to the therapist, then to be equal, the therapist should also, as reassurance, display the same information to the*

*client. [...] This approach maintains an equal power dynamic between the client and the therapist.* (C3)

To keep themselves professional, therapists generally maintain a good boundary with their clients, merely disclosing themselves in therapy sessions [62, 90]. However, appropriate therapist self-disclosure has been shown to be valuable for promoting therapeutic alliance [55, 98, 127]. Sharing therapists' non-verbal cues with clients might serve as an additional channel for their self-disclosure to balance the disclosure inequality. All of our client participants indicated that they would like to know more about their therapist through the sharing process. C3 articulated the idea that clients should possess an equitable level of power within their therapeutic relationships. They imagined that therapists could actively engage in sharing the same sensing information as their clients. C9 believed that many mental concerns stemmed from unequal disclosure in interpersonal communication, which was also applicable in therapeutic relationships. With a high willingness to avoid *"treating therapists as emotional trash bins"*, they especially appreciated it if therapists could share their non-verbal cues with the support of sensing systems.

As mental service receivers, clients cared about their therapists' overall well-being and behavior, to ensure the delivery of high-quality therapeutic care. Regarding *Scenario 3 (Self-overlooked therapist burnout)* where the therapist was exhausted, all client participants suggested that the client should see the therapist's tracked cues if there were sensing systems deployed. T1, even from a therapist's perspective, indicated that *"They (clients) also have the right to know whether the therapist, as a service provider, is in a good mood."* Clients also wished to know if their therapists genuinely cared about them, particularly when therapists exhibited limited responsiveness (3/9 Clients). For example, C1 mentioned that their therapist usually remained silent when they were talking. Therefore, they regarded seeing therapists' *Feature 2 (Emotion and Sentiment)* provided by *Sys 2 (Webcam-Based Facial Expression Analysis)* might *"help clients judge whether their therapist is following their narratives"*. C8, similarly, wanted the therapist in *Scenario 4 (Lack of trustworthy and safety)* to share their *Feature 1 (Concentration)* since *"Although he (the therapist) always says it's fine, it sounds a bit perfunctory. I hope to know that he is truly paying attention to the matter."* Interestingly, some clients also expressed concern about the potential impact of their actions on their therapists (3/9 Clients). For example, *"whether the therapist feels angry [...] because of my behavior"* (C6). In these cases, clients wanted to see therapists' tracked cues to keep them sympathetic.

Notably, participants suggested that accessing therapists' tracked cues could also offer comfort to clients, as these cues can foster a sense of empathy towards their therapists (2/9 Clients; 2/9 Therapists). Regarding the case of *Scenario 4 (Lack of trustworthy and safety)*, T9 noted that *"therapists can show concrete feelings of themselves to help clients reflect on their feelings (through sensing systems)."* T10 believed that clients might find reassurance in seeing therapists' bad moods and exhaustion displayed by the system, since *"they might realize that the therapist isn't so untouchable"* through this process. C7 indicated a similar statement from a client's perspective: *"I might feel awful if I had a bad experience in therapy. However, if I know that the therapist also feels terrible, I might have a sense of resonance. This is not my fault nor yours; rather, the remote approach limits our performance."*

### 4.2 The Risks of Sensing and Sharing: Compromised Mental Interventions, Personal Rights, and Social Dynamics

Although most participants found it acceptable to track and share non-verbal cues in all four scenarios we provided, many of them questioned whether sensing systems and sharing pipelines were reliable and secure enough. Participants also raised concerns about whether tracking and sharing non-verbal cues could undermine the therapeutic alliance. Therefore, some participants expressed hesitance to use these systems when clients discussed sensitive topics, or when either

clients or therapists felt mentally overwhelmed. In the following section, we present participants' main concerns and perceived risks related to sensing and sharing non-verbal cues in remote psychotherapy.

*4.2.1 Potential Inaccuracy in Detection and Analysis.* Participants described their hesitation to use the system because of the potential inaccuracy of sensing and analysis. Some participants raised the doubt of accuracy regarding the technology limitation of the tracking and sharing process (3/9 Clients; 2/10 Therapists). For example, T5 expressed that *the error rate of detection could be a bit high.* T10 doubted the capacity of internet transition: *"As sharing so much tracking information, can the internet speed handle it?"*

Participants were also concerned that sensing systems might not possess the capability to interpret the diverse modes of expression exhibited by different individuals in different situations. They thus saw a need for a contextual framework for the system analysis (4/9 Clients; 4/10 Therapists). For example, T5 noted that the tracking of eye movement by *Sys 1 (Webcam-Based Eye Tracking)* was not a reliable indicator of concentration, referring to their personal experience as a client: *"When I was in a therapy session, I could feel my eyes wandering. [...] This could also be because I don't want to stare at the therapist all the time [...] so I might look elsewhere. But this doesn't mean I'm distracted."* C5 also questioned the accuracy of *Sys 3 (Accelerometer-Based Tremor Detection)*, which uses hand trembling as an indicator of pressure: *"Sometimes hands may shake when people are very emotional. But for many people and in different situations, the presence of significant pressure won't manifest as hand trembling."* Moreover, the analysis and interpretation could be particularly difficult for specific cues. For example, many participants regarded *Feature 2 (Emotion and Sentiment)* as complex. C8 commented that it *"can not be simply quantified"* and *"is not just a label (describing a person)"*. C2 also indicated that people can pretend their expressive emotion, making it hard to interpret referring to just facial expressions: *"A person may feel upset, but they turn their mouth corners to lift into a forced smile. How can the system distinguish whether the person is truly happy or sad? It just observes how the muscles are moving and then determines whether the facial expression represents a smile or not."*

The potential of inaccurate sensing and analysis could directly compromise the therapy experience, as many participants speculated. C2 imagined their feeling of being misinterpreted by the system as *"the machine doesn't understand me"*, which might intensify their discomfort in therapy. T5 raised the concern that inaccurate analysis could lead to confusion when therapists were making a judgment. They indicated that *"Maybe it's enough for the therapist to rely merely on their intuition to understand the client's condition, however, adding wrong information (provided by sensors) [...] can be a disturbance on the contrary."*

*4.2.2 Personal Information Privacy Concerns.* Participants expressed their concerns about their personal information privacy if they were tracked by a system. Some participants were especially concerned about the collection and analysis processes of specific non-verbal cues (2/9 Clients; 1/10 Therapists). For example, C2 identified *Feature 4 (Biosignal Rate)* as *"an extremely private and personal signal"*. They were cautious about their heartbeat being collected and resisted being tracked in general. C1, similarly, also thought *Feature 2 (Emotion and Sentiment)* private, and therefore refused to share it with their therapist. Additionally, many clients commented that they didn't want to be tracked during sensitive conversations (6/9 Clients; T9 from a client perspective), reluctant to any information being recorded as they might talk about content never shared with others before. For example, C2 elaborated on their concerns: *"(I don't want to be tracked) When I talk about things that shame me, such as topics [...] like sex, I want these things to be kept as confidential as possible. Even if you tell me these data will be well managed or whatever, I might still have doubts because, after all, this stuff exists."*

Some client participants were especially uncomfortable regarding the passive disclosure through the sensing and sharing process. As C5 addressed, they would feel as if they were *"naked, exposed body completely"*. Client participants, therefore, wanted the right to selectively disclose themselves when using the systems (3/9 Clients). C5, for example, indicated that they preferred *"the therapists guess (my thoughts) rather than providing a clear-cut analysis."* C1 mentioned that they didn't want to be tracked because they sometimes lied to their therapist: *"Sometimes I don't want to tell the truth to my therapist. Maybe because I did something bad and I want to beautify my behavior [...] I hope to turn them (sensing systems) off because I don't want my attitude to be exposed. I would also have an exaggerated feeling of lying (if turned them on)."*

Participants also reported their concerns about potential data breaches during the transition and storage process (1/9 Clients; 4/10 Therapists). For example, T5 noticed that sharing the tracked cues was *"Relying on internet"*, therefore stating that *"There are risks here [...] possibly be spied on by hackers."* T7 also mentioned their doubt about the database storing tracked cues *"not absolutely secure"*. Additionally, both clients and therapists identified the risk that stored data could potentially be accessed and utilized without their explicit consent (2/9 Clients; 1/10 Therapists).

#### 4.2.3 Extra Cognition Burden.

> *I feel that sensing systems themselves might interfere with our communication. These sensors don't ever exist in general situations. It's something additional. If this extra factor interferes with our communication, how should we assess accountability?* (C3)

During client-therapist interaction, a wealth of information was exchanged both explicitly and implicitly. Although seeing tracked cues could promote understanding of themselves and each other (Section 4.1), as a trade-off, this process could potentially introduce a cognitive burden for both clients and therapists. C3 perceived sensing systems as extra factors in communication, addressing potential interference the sensing and sharing process might bring.

Therapists require strong skills for improvisation and analysis during therapy. Therefore, some therapist participants expressed concern that the introduction of extraneous tracked data, while providing more details about the therapy and their clients (Section 4.1.1, Section 4.1.2), might interfere with the therapy conduction process (6/10 Therapists). On the one hand, viewing their personal tracked cues might be a distraction. As T5 commented on the situation in *Scenario 3 (Self-overlooked therapist burnout)*, *"Therapists usually have the self-awareness to adjust their state accordingly. However, if external sensors provide them with some data, it might distract or confuse them."* On the other hand, getting additional clients' information might adversely interfere with therapists' judgment. T6 noted that the presence of clients' tracked cues required additional efforts for evaluation: *"If I see a sudden increase in their stress level, I have to ask them, 'Are you feeling anxious right now? Why are you anxious?' I have to think about it. This might be a disturbance for me."*

Compared to therapists, clients could potentially face greater mental threats when interacting with the sensing systems. Some client participants raised concerns that the feeling of being under surveillance could make them uncomfortable (4/9 Clients). It aligned with the concept of the "white coat effect", *"a kind of psychological intrusion, as if you feel you have a tube inserted into you, or something else has entered your body"* (C3). C6, for example, was afraid of being tracked, indicating a possible tendency to express less because of fear: *"I become the one being experimented on [...] (being tracked) reminds me to be rational, to hide my emotions, and to act like a person without mental issues."*

Seeing an excessive number of tracked cues might pose a burden on clients' concentration and mood (5/9 Clients; 3/10 Therapists). Both T3 and T10 suggested presenting the analysis of non-verbal cues to their clients as little as possible, concerning that clients could not process an overload

of information. T4 also imagined clients viewing their personal tracked cues *similar as looking into the mirror constantly*, which could *negatively impact the efficiency of therapy*. Furthermore, awareness of poor personal state, which could become easier supported by the tracked data, might also intensify the depression of clients. C8 gave an example regarding this: *"if I'm already very nervous [...] recalling a painful experience about failing an exam, and then the system tells me 'you don't seem to be doing well', or 'you seem to be under a lot of pressure', of course, I already know that. But what can I do about it?"*

As proposed by our participants, clients might feel comforted if seeing therapists' tracked cues, which promote equity and empathy between them (Section 4.1.3). However, they raised a trade-off that sharing therapists' tracked cues might sometimes burden clients (3/9 Clients; 8/10 Therapists). Many therapist participants, therefore, hesitated to share their cues with clients. Participants attributed this concern to two aspects: the already heightened mental burden on clients and the potential negative implications of therapists' tracked cues. First, in many situations, clients were already struggling to manage themselves and had little capacity to process additional information. Therefore, they indicated a lack of interest in their therapists' actions at that time. C6 confessed that *"Whether (therapist) share it or not makes no big difference [...] I'm self-conscious. I don't want to consider additional information."* T10 was afraid that the clients *"start guessing their therapists after they receive additional information."* During therapy sessions, therapists' negative moods and unprofessional behaviors, such as those in *Scenario 3 (Self-overlooked therapist burnout)*, could hurt clients. Therapists were particularly concerned about their negative cues being aggregated and augmented by the sharing and visualizing process, therefore being reluctant to share their cues with clients. For example, perceiving a conflict between professional and personal perspectives, T6 refused to be tracked if they had a negative tendency towards clients' statements. They indicated that *"[...] if my (negative) tendency was tracked, it might hurt the clients, or even led to their distrust and resistance"*. T5, similarly, refused to reveal their tracked cues to clients since they were *"not certain what interpretation the client might make"*, which *"would make the client-therapist relationship more complex"*.

## 5 DISCUSSION

We investigated perceived values and risks of sensing and sharing non-verbal cues in remote psychotherapy, by inquiring clients and therapists about speculating sensing technology for communication in challenging therapy scenarios. Our findings provide initial insights into designing sensing systems for remote client-therapist interaction, highlighting the possibility of extending the application of sensing-technology-empowered communication to vulnerable settings like therapy. In the following sections, we highlight the implications of our findings. We discuss how the utilization of sensing systems could influence power dynamics in therapeutic alliances and other interpersonal relationships. We also assess the considerations of incorporating sensing-technology-empowered communication for remote psychotherapy in real practice, comparing it to other non-vulnerable settings.

### 5.1 Reciprocal Disclosure of Non-verbal Cues: Balancing Power Dynamic Within Therapeutic Alliance and Other Interpersonal Relationships

Sensing and sharing non-verbal cues has shown success in enhancing understanding, intimacy, and empathy in task-oriented collaboration and day-to-day social interaction [49, 60, 68, 77–79, 99, 106, 122]. Aligning with previous reports, our findings highlight the potential benefits of mutual sharing of non-verbal cues in fostering therapeutic alliance. From the therapist participants' standpoint, such integration could deepen their understanding of clients and improve their evaluative capabilities, enabling more targeted interventions (Section 4.1.2). From the client participants' perspective, such

integration could offer them insights into their therapists, giving them a sense of empathy (Section 4.1.3).

Notably, the reciprocal sharing of non-verbal cues could potentially balance the asymmetric client-therapist power dynamic by encouraging more self-disclosure from therapists (Section 4.1.3). Similar to the tension between health service providers and customers [103], clients usually have less power than their therapists. Aiming to seek mental support, clients invest authority in their therapists, enabling therapists to use their specialized expertise to improve their mental well-being. This mental model creates an inherent power imbalance in the therapeutic relationship [21, 51, 116]. Moreover, clients often disclose their personal experiences much more than their therapists during the therapy session [72], which could exacerbate the power asymmetry in their relationship. As more information provides more power in the therapeutic relationship [21], clients can be empowered through the sharing process. The act of therapists sharing their tracked cues with clients can be considered a form of therapist self-disclosure, the openness of which contributes to a feeling of equality by exposing therapists' humanness to clients [55, 98, 127]. Our client participants perceived a potential to see therapists' tracked cues to ensure their interests as service receivers (Section 4.1.3). Therefore, the asymmetric power between therapists and clients could be mitigated through the reciprocal disclosure of non-verbal cues, leading to a more equitable and intimate therapeutic relationship.

The possibility of sharing non-verbal cues to balance power dynamics is worth exploring. Beyond remote psychotherapy, the challenges in power imbalances are prevalent across a diverse range of interpersonal relationships. For example, power asymmetric among colleagues, especially between workers and their managers, is commonly observed in workplace settings, leading to obstacles in their collaborations [43]. Within the context of family, the power asymmetry between parent and child, often exacerbated by parental over-control, has been identified as a primary source of tension between them [58]. Power imbalances are prevalent even within intimate relationships. Usually stemming from gender nature, controlling behaviors, and imbalanced responsibility [59, 111], this asymmetry in power poses a significant threat to the stability of close relationships. Although the sensing and sharing of non-verbal cues has been employed to strengthen understanding, intimacy, and empathy, how such reciprocal disclosure could influence power imbalances remains largely unexplored. We recommend future works investigate this possibility in multiple contexts.

## 5.2 Deploying Sensing Systems in Remote Psychotherapy: Increased Cautions and Considerations

Unlike casual interactions, psychotherapy maintains a professional, service-oriented endeavor. Clients seeking mental support through therapy are often fragile and sensitive about their personal rights. Considering the vulnerable nature of remote psychotherapy, the risks introduced by sensing systems could be highly detrimental to clients and therapists. Our findings revealed that both clients and therapists have many concerns about the use of sensing systems. They raised questions regarding potential violations of not only client rights but also therapy professions and ethics (Section 4.2), which go beyond the known concerns about sensing-technology-empowered communications (Section 2.2). Based on our findings, we discuss how sensing systems should be designed to empower remote client-therapist interactions, highlighting how these considerations exceed those for system deployment in non-vulnerable interactions.

*5.2.1 Demand for System Accuracy and Contextualized Analysis.* While system inaccuracy has been a concern in applying sensing technology to social activities [41, 49, 79], researchers have found that people maintain a high tolerance for system errors in casual interaction contexts [71, 79]. In contrast, when it comes to client-therapist interactions, both of our client and therapist participants

expressed significant concerns about system accuracy, stating that poor quality analysis could harm the quality of therapy (Section 4.2.1). Therapists are especially vulnerable to inaccurate analysis, as it could lead to biased judgments and ineffective mental interventions (Section 4.2.3). Inaccurate analysis could also introduce complexities in accountability, similar to the problems faced in technology-mediated medical practices [38, 42].

Highly accurate sensors and robust predictive algorithms are not enough to meet the needs of client-therapist interaction. Both of our clients and therapists argued that expressions can vary significantly between individuals and contexts (Section 4.2.1), which poses a risk of erroneous algorithmic estimates. People's behavior can also vary because of their personality and contextual factors, which has already been identified as a challenge for interpreting sensing data for both machines and humans [49, 79]. Moreover, as therapy continues, clients' mental transformations can also result in a changed behavior [47]. Therapists, therefore, must continuously learn about their clients through observation and reflection [40]. Admittedly, internal states like biosignals are more reliable for analysis because they cannot be disguised (Section 4.1.2), reducing the need for complex contextual analysis. However, the implications of internal states can be uncontrolled and significantly affected by personal factors [49, 76, 106], making them harder to interpret than external expressive cues.

Therefore, our findings suggest that remote psychotherapy, due to its professional nature, requires higher system accuracy supported by contextualized analysis compared to casual conversations to effectively leverage sensing systems.

*5.2.2 Rigorous Requirements for System Privacy and Security.* As prior works report, privacy and security risks associated with sensing-technology-empowered communication have long been questioned by end-users [29, 49, 79, 85]. Biological, vital, and psychological signals are more widely used beyond communication support for physical health and mental well-being inference. Many people, therefore, regard this type of tracked signal as identifiable, treating it as private, sensitive information that needs extra protections [83].

Unsurprisingly, we found the client and therapist participants concerned about the system's security risks and invasion of personal information privacy in both the sensing and sharing process (Section 4.2.2), aligning with the general concerns above. Furthermore, we consider system privacy and security in remote psychotherapy as more critical than those in causal communications. First, successful psychotherapy hinges on conducting therapy sessions in a private, comfortable setting while maintaining confidentiality [10, 113], where clients are more likely to open up and build intimacy with their therapists [47, 113]. Moreover, in therapy sessions, clients may share experiences they never disclosed to anyone else (Section 4.2.2, Section 4.2.3). The disclosure of private content could make the invasion and leakage of personal privacy intolerable [44]. Therefore, confidentiality and privacy protection are highly respected in psychology's code of ethics [8] and emphasized in health laws, such as privacy requirements outlined in *The Health Insurance Portability and Accountability Act* (HIPAA) [50]. Furthermore, there is a broad agreement that digital health carries a higher safety risk compared to standard digital services, especially in the transition, storage, and processing of sensitive health data [81]. Similarly, sensing systems introduced in remote psychotherapy, compared to communications in health-irrelevant contexts, could be more vulnerable to privacy and security risks.

Therefore, ensuring the security of these systems is crucial for supporting effective remote client-therapist interactions, more so than in non-private or less sensitive contexts. To further reduce the risk of potential security incidents, we suggest limiting the duration of sensing system use. For example, instead of tracking throughout every therapy session, clients and therapists could

consider turning off the sensing systems when the session is going smoothly, as tracking may not always be necessary.

*5.2.3 Complex Norms for the Sharing Process.* Resistance to sharing tracked cues with others is frequently reported in sensing-technology-empowered communication. People often have a feeling of invasion and involuntary disclosures on the sharing process [49, 76, 79]. When sharing biological and vital signals, this feeling could be exaggerated their personal health-related nature. Moreover, the willingness to share highly depends on the relationship solidity and trustworthiness between the communication entities. And, unsurprisingly, people expressed the highest comfort in sharing the information in close relationships [23, 49, 60, 76, 77]. In the context of remote psychotherapy, both clients and therapists expressed heightened considerations about the sharing process, which can be more complicated than in close relationships, intruding on therapeutic relationship boundaries (Section 4.2.2, Section 4.2.3).

According to psychology's code of ethics, clients have the right to selectively choose what and how much information to disclose [10]. Furthermore, disguising is a common social manner in interpersonal communication. For example, "butler lies" have been frequently identified in messenger-mediated communication, indicating a tendency of communication avoidance while maintaining politeness [57]. Clients, similarly, may withhold information or deceive their therapists. Such behavior can manifest as an expression of underlying psychological conflicts, such as a feeling of shame [17, 46]. However, sensing systems usually collect and process information passively and continuously. This could undermine clients' right to selective disclosure, making them excessively uncomfortable or reluctant to self-disclosure in therapy. Further, sharing tracked data and analysis without constraints triggers users' anxiety especially when they feel a lack of control over their physiology [85], which could be particularly harmful to clients considering their elevated cognitive and mental burden [28].

Although moderate therapist self-disclosure is valuable for therapeutic alliance promotion [55, 98, 127], psychologists have constantly argued the risks of excessive disclosure on clients intruding, boundary violations, and role reversal [12, 14, 90], eventually resulting in the motivations departure from the original therapeutic stance [14, 56]. Therapists are required to be client-oriented in self-disclosure, disclosing themselves less frequently at the proper time, in the proper way, and with the proper reasons [61, 72]. Moreover, boundary violation has long been a consideration in remote psychotherapy, as clients can easily get in touch with therapists through the internet, where therapists' unnecessary sharing occurs [128]. Sensing systems without sharing restrictions could intensify the risk of oversharing from therapists, which may not only disrupt the clients but also intensify privacy and ethical concerns for the client-therapist relationship.

Given the nature of therapy and the importance of the therapeutic alliance, we recognize complex norms needed to regulate the sharing of non-verbal cues in remote psychotherapy, compared to what happens between intimate relationships. Below, we further discussed the implications of setting proper norms in sensing-system-empowered client-therapist interactions.

**Prohibited Sharing Cases.** According to Nissenbaum's contextual integrity framework [89], people's perception of privacy in data sharing highly depends on context, including 'when' and 'where' the data transmission happens. Based on this framework, we suggest that sharing tracked cues should be strictly prohibited in many situations to maintain personal and therapeutic boundaries. Any sharing of tracked cues and data must only happen with the explicit consent of both clients and therapists. To protect clients' privacy and rights, they should have the option to withhold sharing whenever especially when they are discussing sensitive content. Therapists should be conscious when sharing their tracked cues with clients, refraining from sharing any information that does not benefit the therapeutic process.

**Informing Sharing Norms through End-User Participation.** Beyond incorporating system features and principles to prevent unrestricted sharing, we see an opportunity to better maintain the boundary of sharing by involving end-users – both clients and therapists – in the design and adjustment of sensing systems. This approach is inspired by the concept of users' participation in the design process to promote control and agency over the system [120]. First, when developing the sensing system, the system manufacturer could invite clients and therapists to participate from the early stages, including design ideation to system testing. They could gather feedback on preferences for system usage (i.e., what sensing system and tracked cues should or should not be engaged in certain therapeutic contexts), and what and how extra features could be introduced to protect the boundary when sharing non-verbal cues (e.g., auto-shutdown of a sharing process). Then, to ensure the sharing process fits the personalized boundary in a specific therapeutic relationship, the client and therapist could actively and periodically discuss their system use preference, adjusting the sharing pattern as the therapy series continues. Their system use logs could also be shared with the system manufacturer to inform further adjustments on boundary protection settings in general.

## 6 LIMITATIONS AND FUTURE WORK

As a speculative study, we conducted scenario-based interviews with participants. Although we provided a comprehensive introduction to every sensing system, participants' understanding of sensing and sharing technology relies mostly on their knowledge and experience, which could bias their judgment. Moreover, all scenarios used in the study were developed based on prior research [82, 97, 114] that represents problematic situations in therapy, whereas no scenario represented a typical or control therapy session. Although we received feedback on the scenario design from 2 clients and 1 therapist to ensure the scenarios could reflect real-world therapy situations, it is possible that they may not accurately represent real therapy sessions, which could be more complex.

The participant demographic also introduces bias to this work. Participants were all from China Mainland and identified themselves as Chinese. They all collaborated with Chinese therapists or clients as well. The practice of psychological intervention is varied due to cultural differences [118]. Therefore, our participants can only represent limited diverse cultural backgrounds in psychotherapy. Moreover, our client participants represented limited client types in psychotherapy. The mental concerns of all client participants were self-identified or diagnosed as mild or moderate. People with severe mental disorders usually have physical and cognitive impairment besides mental pains [20], compromising their communication skills drastically. They may even hold a different view on therapy goals and client-therapist relationships. This group of clients may have a different perspective on sensing systems, which future works can address.

Additionally, our present work emphasizes breadth over depth. We offer a broad overview of sensing systems and recommend that future research concentrate on a specific type of sensing technology or non-verbal cue. Similarly, we disregarded the distinction among different types of psychotherapy. Following different psychological theories, the requirements and implementation of psychotherapy can vary a lot [7]. We acknowledge that this consideration is beyond the scope of our work.

## 7 CONCLUSION

In the present work, we explore the design space of sensing and sharing non-verbal cues in remote psychotherapy through scenario-based interviews with both clients and therapists. By asking clients and therapists to envision sensing technology for challenging situations in remote psychotherapy, our study reveals the potential value of tracked non-verbal cues as a facilitator for self-reflection and mutual disclosure in client-therapist interaction. Our findings also highlight clients' and therapists' concerns about the system accuracy, as well as privacy threats and cognitive load introduced by

the sensing and sharing process. We discuss how reciprocal disclosure of tracked non-verbal cues potentially enhances the client-therapist relationship and general interpersonal relationships. We also point out the cautions of applying sensing-technology-empowered communication in remote psychotherapy, compared to its use in non-vulnerable settings such as interactions between couples. Our work explores the values and risks of implementing sensing technology to overcome challenges in remote client-therapist interaction. It also adds to the growing HCI research that leverages sensing technology to promote interpersonal interactions by inquiring about such socio-technical systems served in a vulnerable setting.

## ACKNOWLEDGMENTS

De Choudhury was partly supported through NIH grants R01 MH117172 and P50 MH115838, and grants from Cisco, Microsoft, and the American Foundation for Suicide Prevention. We thank Marshini Chetty and members of the AIR Lab at the University of Chicago for providing valuable feedback during the paper revision. We appreciate the anonymous reviewers for their constructive feedback which significantly improved this paper. We thank all participants for their input to this study. Lastly, we appreciate our friends and colleagues who helped us connect with participants.

## REFERENCES

[1] Qin An, Xiubin Lin, Zhiqin Sang, and Mingyi Qian. 2023. Practice of common ethical standards in the field of counseling and psychotherapy in mainland China. *Ethics & Behavior* 33, 3 (2023), 183–192. https://doi.org/10.1080/10508422.2022.2118120

[2] Kristen E Anderson, Catherine E Byrne, Ross D Crosby, and Daniel Le Grange. 2017. Utilizing telehealth to deliver family-based treatment for adolescent anorexia nervosa. *International Journal of Eating Disorders* 50, 10 (2017), 1235–1238. https://doi.org/10.1002/eat.22759

[3] G Andrews, A Basu, P Cuijpers, MG Craske, Peter McEvoy, CL English, and JM Newby. 2018. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. *Journal of anxiety disorders* 55 (2018), 70–78. https://doi.org/10.1016/j.janxdis.2018.01.001

[4] Apple. 2024. Healthcare - Apple Watch. https://www.apple.com/healthcare/apple-watch/ Accessed: 2024-06.

[5] Dane Archer and Robin M Akert. 1977. Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of personality and social psychology* 35, 6 (1977), 443. https://doi.org/10.1037/0022-3514.35.6.443

[6] Michael Argyle. 1972. *Non-verbal communication in human social interaction*. Vol. 2. Cambridge.

[7] American Psychological Association. 2009. Different approaches to psychotherapy. https://www.apa.org/topics/psychotherapy/approaches Accessed: 2024-06.

[8] American Psychological Association. 2017. Ethical Principles of Psychologists and Code of Conduct. https://www.apa.org/ethics Accessed: 2024-06.

[9] American Psychological Association. 2018. Therapeutic Alliance - APA Dictionary of Psychology. https://dictionary.apa.org/therapeutic-alliance Accessed: 2024-06.

[10] American Psychological Association. 2019. Protecting your privacy: Understanding confidentiality in psychotherapy. https://www.apa.org/topics/psychotherapy/confidentiality Accessed: 2024-06.

[11] American Psychological Association. 2019. Psychotherapy in China. https://www.apa.org/monitor/2019/10/psychotherapy-china Accessed: 2024-06.

[12] Cristelle T Audet. 2011. Client perspectives of therapist self-disclosure: Violating boundaries or removing barriers? *Counselling Psychology Quarterly* 24, 2 (2011), 85–100. https://doi.org/10.1080/09515070.2011.589602

[13] Amazon AWS. 2024. What is Sentiment Analysis? https://aws.amazon.com/what-is/sentiment-analysis/ Accessed: 2024-06.

[14] Jeffrey E Barnett. 1998. Should psychotherapists self-disclose? Clinical and ethical considerations. (1998).

[15] Penny E Bee, Peter Bower, Karina Lovell, Simon Gilbody, David Richards, Linda Gask, and Pamela Roach. 2008. Psychotherapy mediated by remote communication technologies: a meta-analytic review. *BMC psychiatry* 8, 1 (2008), 1–13. https://doi.org/10.1186/1471-244X-8-60

[16] Sandra Benavides-Vaello, Anne Strode, and Beth C Sheeran. 2013. Using technology in the delivery of mental health and substance abuse treatment in rural communities: a review. *The journal of behavioral health services & research* 40 (2013), 111–120. https://doi.org/10.1007/s11414-012-9299-6

[17] Matt Blanchard and Barry A Farber. 2018. Lying in psychotherapy: Why and what clients don't tell their therapist about therapy and their relationship. In *Disclosure and Concealment in Psychotherapy*. Routledge, 90–112.

[18] Silvia Bonaccio, Jane O'Reilly, Sharon L O'Sullivan, and François Chiocchio. 2016. Nonverbal behavior and communication in the workplace: A review and an agenda for research. *Journal of Management* 42, 5 (2016), 1044–1074. https://doi.org/10.1177/0149206315621146

[19] Nathan Bos, Judy Olson, Darren Gergle, Gary Olson, and Zach Wright. 2002. Effects of four computer-mediated communications channels on trust development. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 135–140. https://doi.org/10.1145/503376.503401

[20] Christopher R Bowie. 2019. Cognitive remediation for severe mental illness: state of the field and future directions. *World Psychiatry* 18, 3 (2019), 274. https://doi.org/10.1002/wps.20660

[21] KK Boyd. 1996. Power imbalances and therapy. *Focus (San Francisco, Calif.)* 11, 9 (1996), 1–4.

[22] China Briefing. 2018. The Mental Healthcare Industry in China. https://www.china-briefing.com/news/mental-healthcare-industry-china Accessed: 2024-06.

[23] Daniel Buschek, Mariam Hassib, and Florian Alt. 2018. Personal mobile messaging in context: Chat augmentations for expressiveness and awareness. *ACM Transactions on Computer-Human Interaction (TOCHI)* 25, 4 (2018), 1–33. https://doi.org/10.1145/3201404

[24] Hancheng Cao, Chia-Jung Lee, Shamsi Iqbal, Mary Czerwinski, Priscilla NY Wong, Sean Rintel, Brent Hecht, Jaime Teevan, and Longqi Yang. 2021. Large scale analysis of multitasking behavior during remote meetings. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–13. https://doi.org/10.1145/3411764.3445243

[25] Gianluca Castelnuovo, Andrea Gaggioli, Fabrizia Mantovani, and Giuseppe Riva. 2003. From psychotherapy to e-therapy: the integration of traditional techniques and new communication tools in clinical settings. *CyberPsychology & Behavior* 6, 4 (2003), 375–382. https://doi.org/10.1089/109493103322278754

[26] Gianluca Castelnuovo, Andrea Gaggioli, Fabrizia Mantovani, and Giuseppe Riva. 2003. New and old tools in psychotherapy: The use of technology for the integration of the traditional clinical treatments. *Psychotherapy: theory, research, practice, training* 40, 1-2 (2003), 33. https://doi.org/10.1037/0033-3204.40.1-2.33

[27] Francesco Cataldo, Shanton Chang, Antonette Mendoza, George Buchanan, et al. 2021. A perspective on client-psychologist relationships in videoconferencing psychotherapy: Literature review. *JMIR mental health* 8, 2 (2021), e19004. https://doi.org/10.2196/19004

[28] Francesco Cataldo, Antonette Mendoza, Shanton Chang, George Buchanan, Nicholas T Van Dam, et al. 2023. Enhancing Therapeutic Processes in Videoconferencing Psychotherapy: Interview Study of Psychologists' Technological Perspective. *JMIR Formative Research* 7, 1 (2023), e40542. https://doi.org/10.2196/40542

[29] Guillaume Chanel and Christian Mühl. 2015. Connecting brains and bodies: applying physiological computing to support social interaction. *Interacting with Computers* 27, 5 (2015), 534–550. https://doi.org/10.1093/iwc/iwv013

[30] Cory K Chen, Nicole Nehrig, Lauren Wash, Jennifer A Schneider, Sagiv Ashkenazi, Elana Cairo, Angel F Guyton, and Amy Palfrey. 2021. When distance brings us closer: Leveraging tele-psychotherapy to build deeper connection. *Counselling Psychology Quarterly* 34, 3-4 (2021), 554–567. https://doi.org/10.1080/09515070.2020.1779031

[31] Zhaoyi Chen, Katie Aafjes-van Doorn, and Vera Békés. 2024. Therapists' teletherapy experiences during the pandemic in China and the United States. *Asia Pacific Journal of Counselling and Psychotherapy* (2024), 1–19. https://doi.org/10.1080/21507686.2024.2341336

[32] Sabrina Cipolletta and Damiano Mocellin. 2018. Online counseling: An exploratory survey of Italian psychologists' attitudes towards new ways of interaction. *Psychotherapy research* 28, 6 (2018), 909–924. https://doi.org/10.1080/10503307.2016.1259533

[33] Marylene Cloitre, K Chase Stovall-McClough, Regina Miranda, and Claude M Chemtob. 2004. Therapeutic alliance, negative mood regulation, and treatment outcome in child abuse-related posttraumatic stress disorder. *Journal of consulting and clinical psychology* 72, 3 (2004), 411. https://doi.org/10.1037/0022-006X.72.3.411

[34] Max T Curran, Jeremy Raboff Gordon, Lily Lin, Priyashri Kamlesh Sridhar, and John Chuang. 2019. Understanding digitally-mediated empathy: An exploration of visual, narrative, and biosensory informational cues. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–13. https://doi.org/10.1145/3290605.3300844

[35] Vedant Das Swain, Lan Gao, Abhirup Mondal, Gregory D Abowd, and Munmun De Choudhury. 2024. Sensible and Sensitive AI for Worker Wellbeing: Factors that Inform Adoption and Resistance for Information Workers. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–30. https://doi.org/10.1145/3613904.3642716

[36] Vedant Das Swain, Lan Gao, William A Wood, Srikruthi C Matli, Gregory D Abowd, and Munmun De Choudhury. 2023. Algorithmic Power or Punishment: Information Worker Perspectives on Passive Sensing Enabled AI Phenotyping of Performance and Wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17. https://doi.org/10.1145/3544548.3581376

[37] Susan X Day and Paul L Schneider. 2002. Psychotherapy using distance technology: A comparison of face-to-face, video, and audio treatment. *Journal of Counseling Psychology* 49, 4 (2002), 499. https://doi.org/10.1037/0022-0167.49.4.499

[38] Matthew DeCamp and Jon C Tilburt. 2019. Why we cannot trust artificial intelligence in medicine. *The Lancet Digital Health* 1, 8 (2019), e390. https://doi.org/10.1016/S2589-7500(19)30197-9

[39] Arindam Dey, Thammathip Piumsomboon, Youngho Lee, and Mark Billinghurst. 2017. Effects of sharing physiological states of players in a collaborative virtual reality gameplay. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 4045–4056. https://doi.org/10.1145/3025453.3026028

[40] Barry Duncan. 2010. On becoming a better therapist. *Psychotherapy in Australia* 16, 4 (2010), 42–51. https://search.informit.org/doi/10.3316/informit.307687236000677

[41] Eylül Ertay, Hao Huang, Zhanna Sarsenbayeva, and Tilman Dingler. 2021. Challenges of emotion detection using facial expressions and emotion visualisation in remote communication. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers*. 230–236. https://doi.org/10.1145/3460418.3479341

[42] Pouyan Esmaeilzadeh. 2020. Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. *BMC medical informatics and decision making* 20, 1 (2020), 1–19. https://doi.org/10.1186/s12911-020-01191-1

[43] Driss Essabbar, Maria Zrikem, and Marc Zolgadri. 2016. Power imbalance in collaboration relationships. *International Journal of Supply and Operations Management* 2, 4 (2016), 1021–1034. https://doi.org/10.22034/2015.4.04

[44] Louis Everstine, Diana Sullivan Everstine, Gary M Heymann, Reiko Homma True, David H Frey, Harold G Johnson, and Richard H Seiden. 1980. Privacy and confidentiality in psychotherapy. *American Psychologist* 35, 9 (1980), 828. https://doi.org/10.1037/0003-066X.35.9.828

[45] Fredrik Falkenström, Fredrik Granström, and Rolf Holmqvist. 2013. Therapeutic alliance predicts symptomatic improvement session by session. *Journal of counseling psychology* 60, 3 (2013), 317. https://doi.org/10.1037/a0032258

[46] Barry A Farber. 2003. Patient self-disclosure: A review of the research. *Journal of clinical psychology* 59, 5 (2003), 589–600. https://doi.org/10.1002/jclp.10161

[47] Barry A Farber, Kathryn C Berano, and Joseph A Capobianco. 2004. Clients' Perceptions of the Process and Consequences of Self-Disclosure in Psychotherapy. *Journal of Counseling Psychology* 51, 3 (2004), 340. https://doi.org/10.1037/0022-0167.51.3.340

[48] Heather A Faucett, Matthew L Lee, and Scott Carter. 2017. I should listen more: real-time sensing and feedback of non-verbal communication in video telehealth. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–19. https://doi.org/10.1145/3134679

[49] Milou A Feijt, Joyce HDM Westerink, Yvonne AW De Kort, and Wijnand A IJsselsteijn. 2023. Sharing biosignals: An analysis of the experiential and communication properties of interpersonal psychophysiology. *Human–Computer Interaction* 38, 1 (2023), 49–78. https://doi.org/10.1080/07370024.2021.1913164

[50] Centers for Disease Control and Prevention. 2022. Health Insurance Portability and Accountability Act of 1996 (HIPAA). https://www.cdc.gov/phlp/publications/topic/hipaa.html Accessed: 2024-06.

[51] Malin Fors. 2021. Power dynamics in the clinical situation: A confluence of perspectives. *Contemporary Psychoanalysis* 57, 2 (2021), 242–269. https://doi.org/10.1080/00107530.2021.1935191

[52] Shari Geller. 2021. Cultivating online therapeutic presence: strengthening therapeutic relationships in teletherapy sessions. *Counselling Psychology Quarterly* 34, 3-4 (2021), 687–703. https://doi.org/10.1080/09515070.2020.1787348

[53] Vanessa Germain, André Marchand, Stéphane Bouchard, Marc-Simon Drouin, and Stéphane Guay. 2009. Effectiveness of cognitive behavioural therapy administered by videoconference for posttraumatic stress disorder. *Cognitive Behaviour Therapy* 38, 1 (2009), 42–53. https://doi.org/10.1080/16506070802473494

[54] Vanessa Germain, André Marchand, Stéphane Bouchard, Stéphane Guay, and Marc-Simon Drouin. 2010. Assessment of the therapeutic alliance in face-to-face or videoconference treatment for posttraumatic stress disorder. *Cyberpsychology, Behavior, and Social Networking* 13, 1 (2010), 29–35. https://doi.org/10.1089/cyber.2009.0139

[55] Marvin R Goldfried, Lisa A Burckell, and Catherine Eubanks-Carter. 2003. Therapist self-disclosure in cognitive-behavior therapy. *Journal of clinical psychology* 59, 5 (2003), 555–568. https://doi.org/10.1002/jclp.10159

[56] Thomas G Gutheil and Glen O Gabbard. 1993. The concept of boundaries in clinical practice: theoretical and risk-management dimensions. *The American journal of psychiatry* 150, 2 (1993), 188–196. https://doi.org/10.1176/ajp.150.2.188

[57] Jeff Hancock, Jeremy Birnholtz, Natalya Bazarova, Jamie Guillory, Josh Perlin, and Barrett Amos. 2009. Butler lies: awareness, deception and design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 517–526. https://doi.org/10.1145/1518701.1518782

[58] Lori D Harach and Leon J Kuczynski. 2005. Construction and maintenance of parent–child relationships: Bidirectional contributions from the perspective of parents. *Infant and Child Development: An International Journal of Research and Practice* 14, 4 (2005), 327–343. https://doi.org/10.1002/icd.393

[59] S Marie Harvey, Linda J Beckman, Carole H Browner, and Christy A Sherman. 2002. Relationship power, decision making, and sexual relations: An exploratory study with couples of Mexican origin. *Journal of Sex Research* 39, 4 (2002), 284–291. https://doi.org/10.1080/00224490209552152

[60] Mariam Hassib, Daniel Buschek, Paweł W Wozniak, and Florian Alt. 2017. HeartChat: Heart rate augmented mobile chat to support empathy and awareness. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 2239–2251. https://doi.org/10.1145/3025453.3025758

[61] Jennifer R Henretty and Heidi M Levitt. 2010. The role of therapist self-disclosure in psychotherapy: A qualitative review. *Clinical psychology review* 30, 1 (2010), 63–77. https://doi.org/10.1016/j.cpr.2009.09.004

[62] Clara E Hill and Sarah Knox. 2001. Self-disclosure. *Psychotherapy: Theory, Research, Practice, Training* 38, 4 (2001), 413. https://doi.org/10.1037/0033-3204.38.4.413

[63] Clara E Hill, Barbara J Thompson, and Nicholas Ladany. 2003. Therapist use of silence in therapy: A survey. *Journal of clinical psychology* 59, 4 (2003), 513–524. https://doi.org/10.1002/jclp.10155

[64] Tad Hirsch, Christina Soma, Kritzia Merced, Patty Kuo, Aaron Dembe, Derek D Caperton, David C Atkins, and Zac E Imel. 2018. "It's hard to argue with a computer" Investigating Psychotherapists' Attitudes towards Automated Evaluation. In *Proceedings of the 2018 Designing Interactive Systems Conference*. 559–571. https://doi.org/10.1145/3196709.3196776

[65] Adam O Horvath and Lester Luborsky. 1993. The role of the therapeutic alliance in psychotherapy. *Journal of consulting and clinical psychology* 61, 4 (1993), 561. https://doi.org/10.1037/0022-006X.61.4.561

[66] Adam O Horvath and B Dianne Symonds. 1991. Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of counseling psychology* 38, 2 (1991), 139. https://doi.org/10.1037/0022-0167.38.2.139

[67] Yueqin Huang, YU Wang, Hong Wang, Zhaorui Liu, Xin Yu, Jie Yan, Yaqin Yu, Changgui Kou, Xiufeng Xu, Jin Lu, et al. 2019. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *The Lancet Psychiatry* 6, 3 (2019), 211–224. https://doi.org/10.1016/S2215-0366(18)30511-X

[68] Joris H Janssen, Jeremy N Bailenson, Wijnand A IJsselsteijn, and Joyce HDM Westerink. 2010. Intimate heartbeats: Opportunities for affective communication technology. *IEEE Transactions on Affective Computing* 1, 2 (2010), 72–80. https://doi.org/10.1109/T-AFFC.2010.13

[69] Jacinta Jardine, Caroline Earley, Derek Richards, Ladislav Timulak, Jorge E Palacios, Daniel Duffy, Karen Tierney, and Gavin Doherty. 2020. The experience of guided online therapy: a longitudinal, qualitative analysis of client feedback in a naturalistic RCT. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–15. https://doi.org/10.1145/3313831.3376254

[70] Yanqi Jiang, Xianghua Ding, Xiaojuan Ma, Zhida Sun, and Ning Gu. 2023. IntimaSea: Exploring Shared Stress Display in Close Relationships. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19. https://doi.org/10.1145/3544548.3581000

[71] Maria Karam and MC Schraefel. 2006. Investigating user tolerance for errors in vision-enabled gesture-based interactions. In *Proceedings of the working conference on Advanced visual interfaces*. 225–232. https://doi.org/10.1145/1133265.1133309

[72] Sarah Knox and Clara E Hill. 2003. Therapist self-disclosure: Research-based suggestions for practitioners. *Journal of clinical psychology* 59, 5 (2003), 529–539. https://doi.org/10.1002/jclp.10157

[73] Timothy Lamb, Nancy A Pachana, and Nadeeka Dissanayaka. 2019. Update of recent literature on remotely delivered psychotherapy interventions for anxiety and depression. *Telemedicine and e-Health* 25, 8 (2019), 671–677. https://doi.org/10.1089/tmj.2018.0079

[74] Jakob Eg Larsen, Thomas Blomseth Christiansen, and Kasper Eskelund. 2017. Fostering bilateral patient-clinician engagement in active self-tracking of subjective experience. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 427–430. https://doi.org/10.1145/3154862.3154918

[75] Di Liang, Vickie M Mays, and Wei-Chin Hwang. 2018. Integrated mental health services in China: challenges and planning for the future. *Health policy and planning* 33, 1 (2018), 107–122. https://doi.org/10.1093/heapol/czx137

[76] Fannie Liu, Laura Dabbish, and Geoff Kaufman. 2017. Supporting social interactions with an expressive heart rate sharing application. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–26. https://doi.org/10.1145/3130943

[77] Fannie Liu, Mario Esparza, Maria Pavlovskaia, Geoff Kaufman, Laura Dabbish, and Andrés Monroy-Hernández. 2019. Animo: Sharing biosignals on a smartwatch for lightweight social connection. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 1–19. https://doi.org/10.1145/3314405

[78] Fannie Liu, Geoff Kaufman, and Laura Dabbish. 2019. The effect of expressive biosignals on empathy and closeness for a stigmatized group member. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–17. https://doi.org/10.1145/3359303

[79] Fannie Liu, Chunjong Park, Yu Jiang Tham, Tsung-Yu Tsai, Laura Dabbish, Geoff Kaufman, and Andrés Monroy-Hernández. 2021. Significant otter: Understanding the role of biosignals in communication. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15. https://doi.org/10.1145/3411764.3445200

[80] Amy Lopez, Sarah Schwenk, Christopher D Schneck, Rachel J Griffin, and Matthew C Mishkind. 2019. Technology-based mental health treatment and the impact on the therapeutic alliance. *Current psychiatry reports* 21 (2019), 1–7.

https://doi.org/10.1007/s11920-019-1055-7

[81] Samuel D Lustgarten, Yunkyoung L Garrison, Morgan T Sinnard, and Anthony WP Flynn. 2020. Digital privacy in mental healthcare: current issues and recommendations for technology use. *Current opinion in psychology* 36 (2020), 25–31. https://doi.org/10.1016/j.copsyc.2020.03.012

[82] John C Markowitz, Barbara Milrod, Timothy G Heckman, Maja Bergman, Doron Amsalem, Hemrie Zalman, Thomas Ballas, and Yuval Neria. 2021. Psychotherapy at a distance. *American Journal of Psychiatry* 178, 3 (2021), 240–246. https://doi.org/10.1176/appi.ajp.2020.20050557

[83] Andrew McStay. 2020. Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy. *Big Data & Society* 7, 1 (2020), 2053951720904386. https://doi.org/10.1177/2053951720904386

[84] Albert Mehrabian. 1971. Silent messages (Vol. 8, No. 152, p. 30). *Belmont, CA: Wadsworth* (1971).

[85] Clara Moge, Katherine Wang, and Youngjun Cho. 2022. Shared user interfaces of physiological data: Systematic review of social biofeedback systems and contexts in hci. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–16. https://doi.org/10.1145/3491102.3517495

[86] Orla Moran, Julie Doyle, Suzanne Smith, Oonagh Giggins, and John Dinsmore. 2022. Investigating the needs and concerns of older adults with multimorbidity and their healthcare professionals for conceivable digital psychotherapeutic interventions. *Digital Health* 8 (2022), 20552076221089097. https://doi.org/10.1177/20552076221089097

[87] Jim Mullin, Anne H Anderson, Lucy Smallwood, Matthew Jackson, and E Katsavras. 2001. Eye-tracking explorations in multimedia communications. In *People and Computers XV—Interaction without Frontiers: Joint Proceedings of HCI 2001 and IHM 2001*. Springer, 367–382. https://doi.org/10.1007/978-1-4471-0353-0_22

[88] Sean Neill. 2017. *Classroom nonverbal communication*. Routledge.

[89] Helen Nissenbaum. 2004. Privacy as contextual integrity. *Wash. L. Rev.* 79 (2004), 119.

[90] Psychopathology Committee of the Group for the Advancement of Psychiatry. 2001. Reexamination of therapist self-disclosure. *Psychiatric Services* 52, 11 (2001), 1489–1493. https://doi.org/10.1176/appi.ps.52.11.1489

[91] Pat Ogden and Bonnie Goldstein. 2019. Sensorimotor psychotherapy from a distance: Engaging the body, creating presence, and building relationship in videoconferencing. In *Theory and practice of online therapy*. Routledge, 47–65.

[92] Hyanghee Park, Daehwan Ahn, Kartik Hosanagar, and Joonhwan Lee. 2021. Human-AI interaction in human resource management: Understanding why employees resist algorithmic evaluation at workplaces and how to mitigate burdens. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–15. https://doi.org/10.1145/3411764.3445304

[93] Erin L Pederson and David L Vogel. 2007. Male gender role conflict and willingness to seek counseling: Testing a mediation model on college-aged men. *Journal of Counseling Psychology* 54, 4 (2007), 373. https://doi.org/10.1037/0022-0167.54.4.373

[94] Deepika Phutela. 2015. The importance of non-verbal communication. *IUP Journal of Soft Skills* 9, 4 (2015), 43.

[95] He Qitong. 2023. Young Chinese Turn to Taobao for Psychological Support. (2023). https://www.sixthtone.com/news/1013845 Accessed: 2024-06.

[96] Raydiant. 2021. Face Tracking Camera. https://www.raydiant.com/blog/face-tracking-camera Accessed: 2024-06.

[97] Fujiko Robledo Yamamoto, Amy Voida, and Stephen Voida. 2021. From therapy to teletherapy: Relocating mental health services online. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–30. https://doi.org/10.1145/3479508

[98] Carl Ransom Rogers. 1995. *On becoming a person: A therapist's view of psychotherapy*. Houghton Mifflin Harcourt.

[99] Camilo Rojas, Eugenio Zuccarelli, Alexandra Chin, Gaurav Patekar, David Esquivel, and Pattie Maes. 2022. Towards Enhancing Empathy Through Emotion Augmented Remote Communication. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–9. https://doi.org/10.1145/3491101.3519797

[100] Mary Beth Rosson and John M Carroll. 2009. Scenario based design. *Human-computer interaction. boca raton, FL* (2009), 145–162. https://doi.org/10.1016/B978-044481862-1.50083-2

[101] Gillian Isaacs Russell. 2018. *Screen relations: The limits of computer-mediated psychoanalysis and psychotherapy*. Routledge. https://doi.org/10.4324/9780429479762

[102] Andrea Sabbadini. 1991. Listening to silence. *British Journal of Psychotherapy* 7, 4 (1991), 406–415. https://doi.org/10.1111/j.1752-0118.1991.tb01145.x

[103] Brett Scholz, Julia Bocking, Chris Platania-Phung, Michelle Banfield, and Brenda Happell. 2018. "Not an afterthought": Power imbalances in systemic partnerships between health service providers and consumers in a hospital setting. *Health Policy* 122, 8 (2018), 922–928. https://doi.org/10.1016/j.healthpol.2018.06.007

[104] Prabin Sharma, Shubham Joshi, Subash Gautam, Sneha Maharjan, Salik Ram Khanal, Manuel Cabral Reis, João Barroso, and Vítor Manuel de Jesus Filipe. 2022. Student engagement detection using emotion analysis, eye tracking and head movement with machine learning. In *International Conference on Technology and Innovation in Learning, Teaching and Education*. Springer, 52–68. https://doi.org/10.1007/978-3-031-22918-3_5

[105] Susan Simpson. 2009. Psychotherapy via videoconferencing: A review. *British Journal of Guidance & Counselling* 37, 3 (2009), 271–286. https://doi.org/10.1080/03069880902957007

[106] Petr Slovák, Joris Janssen, and Geraldine Fitzpatrick. 2012. Understanding heart rate sharing: towards unpacking physiosocial space. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 859–868. https://doi.org/10.1145/2207676.2208526

[107] MA Smith and C Senior. 2001. The internet and clinical psychology: a general review of the implications. *Clinical Psychology Review* 21, 1 (2001), 129–136. https://doi.org/10.1016/S0272-7358(99)00043-4

[108] Ewelina Smoktunowicz, Azy Barak, Gerhard Andersson, Rosa M Banos, Thomas Berger, Cristina Botella, Blake F Dear, Tara Donker, David D Ebert, Heather Hadjistavropoulos, et al. 2020. Consensus statement on the problem of terminology in psychological interventions using the internet or digital components. *Internet Interventions* 21 (2020), 100331. https://doi.org/10.1016/j.invent.2020.100331

[109] The British Psychological Society. 2020. Tackling gender imbalance in psychology. https://www.bps.org.uk/psychologist/tackling-gender-imbalance-psychology Accessed: 2024-06.

[110] Andriod Open Source. 2024. Sensor types - Andriod Open Source. https://source.android.com/docs/core/interaction/sensors/sensor-types Accessed: 2024-06.

[111] Susan Sprecher and Diane Felmlee. 1997. The balance of power in romantic heterosexual couples over time from "his" and "her" perspectives. *Sex Roles* 37 (1997), 361–379. https://doi.org/10.1023/A:1025601423031

[112] Katarzyna Stawarz, Chris Preist, Deborah Tallon, Laura Thomas, Katrina Turner, Nicola Wiles, David Kessler, Roz Shafran, and David Coyle. 2020. Integrating the digital and the traditional to deliver therapy for depression: Lessons from a pragmatic study. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14. https://doi.org/10.1145/3313831.3376510

[113] Tiffany Y Sui, Shannon McDermott, Brooke Harris, and Honor Hsin. 2023. The impact of physical environments on outpatient mental health recovery: A design-oriented qualitative study of patient perspectives. *Plos one* 18, 4 (2023), e0283962. https://doi.org/10.1371/journal.pone.0283962

[114] Tiffany Thang, Alice Liang, Yechan Choi, Adrian Parrales, Sara H Kuang, Sri Kurniawan, and Heather Perez. 2021. Providing and Accessing Support During the COVID-19 Pandemic: Experiences of Mental Health Professionals, Community and Vocational Support Providers, and Adults with ASD. In *Proceedings of the 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. 1–6. https://doi.org/10.1145/3441852.3476470

[115] Karl Tomm. 1987. Interventive interviewing: Part I. Strategizing as a fourth guideline for the therapist. *Family process* 26, 1 (1987), 3–13. https://doi.org/10.1111/j.1545-5300.1987.00003.x

[116] Nick Totton. 2018. Power in the therapeutic relationship. In *The political self*. Routledge, 29–42.

[117] Minh Tran, Taylan Sen, Kurtis Haut, Mohammad Rafayet Ali, and Ehsan Hoque. 2020. Are you really looking at me? a feature-extraction framework for estimating interpersonal eye gaze from conventional video. *IEEE Transactions on Affective Computing* 13, 2 (2020), 912–925. https://doi.org/10.1109/TAFFC.2020.2979440

[118] Wen-Shing Tseng. 1999. Culture and psychotherapy: Review and practical guidelines. *Transcultural psychiatry* 36, 2 (1999), 131–179. https://doi.org/10.1177/136346159903600201

[119] Kelli Vessoyan, Gill Steckle, Barb Easton, Megan Nichols, Victoria Mok Siu, and Janette McDougall. 2018. Using eye-tracking technology for communication in Rett syndrome: perceptions of impact. *Augmentative and Alternative Communication* 34, 3 (2018), 230–241. https://doi.org/10.1080/07434618.2018.1462848

[120] John Vines, Rachel Clarke, Peter Wright, John McCarthy, and Patrick Olivier. 2013. Configuring participation: on how we involve people in design. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 429–438. https://doi.org/10.1145/2470654.2470716

[121] Bruce E Wampold. 2015. How important are the common factors in psychotherapy? An update. *World psychiatry* 14, 3 (2015), 270–277. https://doi.org/10.1002/wps.20238

[122] R Michael Winters, Bruce N Walker, and Grace Leslie. 2021. Can you hear my heartbeat?: hearing an expressive biosignal elicits empathy. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–11. https://doi.org/10.1145/3411764.3445545

[123] Christine T Wolf. 2019. Explainability scenarios: towards scenario-based XAI design. In *Proceedings of the 24th International Conference on Intelligent User Interfaces*. 252–257. https://doi.org/10.1145/3301275.3302317

[124] Richmond Y Wong, Jason Caleb Valdez, Ashten Alexander, Ariel Chiang, Olivia Quesada, and James Pierce. 2023. Broadening Privacy and Surveillance: Eliciting Interconnected Values with a Scenarios Workbook on Smart Home Cameras. (2023). https://doi.org/10.1145/3563657.3596012

[125] Sarah Woods, Michael Walters, Kheng Lee Koay, and Kerstin Dautenhahn. 2006. Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *9th IEEE International Workshop on Advanced Motion Control, 2006.* IEEE, 750–755. https://doi.org/10.1109/AMC.2006.1631754

[126] Shelly Yu, Sarah D Kowitt, Edwin B Fisher, and Gongying Li. 2018. Mental health in China: Stigma, family obligations, and the potential of peer support. *Community mental health journal* 54 (2018), 757–764. https://doi.org/10.1007/s10597-

017-0182-z

[127] Sharon Ziv-Beiman. 2013. Therapist self-disclosure as an integrative intervention. *Journal of Psychotherapy Integration* 23, 1 (2013), 59. https://doi.org/10.1037/a0031783

[128] Ofer Zur, Martin H Williams, Keren Lehavot, and Samuel Knapp. 2009. Psychotherapist self-disclosure and transparency in the Internet age. *Professional Psychology: Research and Practice* 40, 1 (2009), 22. https://doi.org/10.1037/a0014745

# A APPENDIX

## A.1 Participants Demographic

Table 3. Participants Demographic

| | ID | Gender | Experience on remote or hybrid psychotherapy | In-person therapy experience? | Claims |
|---|---|---|---|---|---|
| **Client Participants** | C1 | Female | 1 course of treatment, 3 months around 8 online sessions | Yes | Anxiety, Emotional Support Needed |
| | C2 | Female | 1 course of treatment, 2.5 years around 125 online sessions | No | Depression Disorder |
| | C3 | Male | 1 course of treatment, 4 months around 12 online sessions | Yes | Not Disclosed |
| | C4 | Female | 1 course of treatment, 9 months around 20 online sessions | Yes | Anxiety |
| | C5 | Male | 2 courses of treatment, 2 month around 5 online sessions | Yes | Anxiety, Interpersonal Relationship Problem |
| | C6 | Female | 1 course of treatment, 2 years around 10 online sessions | Yes | Depression Disorder, Stress, Self Exploration |
| | C7 | Female | 1 course of treatment, 2 month around 5 online sessions | No | Self Exploration |
| | C8 | Female | 2 courses of treatment, 5 weeks around 5 online sessions | Yes | Anxiety, Emotional Support Needed, Self Exploration |
| | C9 | Male | 2 courses of treatment, 6 months around 20 online sessions | Yes | Depression Disorder |
| | **ID** | **Gender** | **Experience on psychotherapy (Experience on remote psychotherapy)** | **In-person therapy experience?** | **Professionals** |
| **Therapist Participants** | T1 | Female | 7 years (3 years) | Yes | Kids and Family Relationship |
| | T2 | Female | 14 years (6 years) | Yes | Educational Psychology |
| | T3 | Female | 7 years (2 years) | Yes | Personal Growth, Interpersonal Relationship |
| | T4 | Female | 9 years (9 years) | Yes | Emotional Distress, Personal Growth, Somatization Disorder, Interpersonal Relationship |
| | T5 | Female | 2 years (1years) | Yes | General Mental Health Concerns |
| | T6 | Female | 10 years (3 years) | Yes | Kids and Family Relationship |
| | T7 | Female | 2 years (2 years) | Yes | Depression Disorder, Anxiety, Personal Growth, Interpersonal Relationship |
| | T8 | Male | 3.5 years (3 years) | Yes | General Mental Health Concerns |
| | T9 | Female | 1 year (1 year) | No | Depression Disorder, Anxiety |
| | T10 | Female | 1.5 years (1.5 years) | Yes | General Mental Health Concerns |

## A.2 Materials for In-Scenario Activities

Table 4. Sensing System Explanation Reading Materials

| Label | Name | Description | | | | Reference |
|---|---|---|---|---|---|---|
| | | The sensor(s) used | Device(s) | How it works | Usages | |
| *Sys 1* | Webcam-Based Eye Tracking | Webcam | Computer, Mobile Phone | This system uses the webcam to capture people's eye movement. Referring to this, you can see which part of the screen people are looking at. | A person may look around with an unusual eye movement because of distraction. This technology can capture and visualize that. People can infer if the tracked people are distracted. | [48, 104] |
| *Sys 2* | Webcam-Based Facial Expression Analysis | Webcam | Computer, Mobile Phone | This system uses the webcam to capture people's facial expressions and analyze their emotions. | People can see the facial expressions of the tracked person. People can infer how stressed the tracked person is based on facial expression. People can know the emotion of the tracked person referring to the analysis. People can infer if the tracked person is distracted based on the facial expression. | [99, 104] |
| *Sys 3* | Accelerometer-Based Tremor Detection | Accelerometer | Mobile Phone, Smartwatch | This system uses the accelerometer to detect the tremor of your device. Referring to this, you can see how the device moves. | A person's hand may be shaken because of nervousness. If the tracked person is holding a mobile phone for therapy, or if he is wearing a smartwatch, this technology can easily identify if their hand is shaken and show it to people. People can infer how stressed the tracked person is. | [97] |
| *Sys 4* | Verbal Analysis | Microphone | Computer, Mobile Phone | This system analyzes people's verbal based on changes in the volume, pitch, and speed of speech. It would provide speaking attributes and sentiment analysis ultimately. | People can see how the tracked person speaks, and also an analysis of how stressful based on the voice tone and speed. People can infer how stressed the tracked person is based on the sentiment. People can know the emotion of the tracked person referring to the analysis. People can infer if the tracked person is distracted based on the sentiment. | [48, 97, 99] |
| *Sys 5* | Personal Biosignal Detection | Optical Heart Rate Sensor, Pressure Sensor, Oximetry Sensor | Smartwatch | This system utilizes multiple sensors to detect people's biosignals, like heartbeat rate, blood pressure, breathing rate, etc. Based on the detection, the system will analyze people's overall status. | When people's emotions drastically change, their biosignal may also change quickly. E.g: heartbeat rate rises when nervous. This technology can show these changes immediately. People can see the overall status of the tracked person. People can infer how stressed the tracked person is. | [77] |
| *Sys 6* | Environmental Analysis | Webcam, Microphone, Light Sensor, LiDAR | Computer, Mobile Phone | This system utilizes multiple sensors to detect changes in the physical environment and social environment. | When there is a light change in the room / extra noises happen in the room, this system will notify the tracked person about that. People can change the environment based on this feedback. | [97] |

Table 5. Scenario Descriptions

| Label | Setting | Story | Corresponding Challenge | Reference |
|---|---|---|---|---|
| *Example Scenario* | Gloria is a therapist who works on anxiety problems. She offers a long-term consultant to Mary, a company employee who had severe insomnia because of her work. This time, they planned to have a one-hour meeting via phone. | Gloria is a therapist who works on anxiety problems. She offers a long-term consultant to Mary, a company employee who had severe insomnia because of her work. This time, they planned to have a one-hour meeting via phone. | Distraction | [82, 97] |
| *Scenario 1* | Jane is a therapist who works on general mental health problems. She was helping one of her clients, Tim, a youth who has an ADHD problem, to integrate into the school environment. They were having a one-hour meeting from both of their homes. The therapy was conducted via Zoom. | Jane met Tim online with both cameras open. The first ten minutes of the therapy went well. However, Tim then tried to pull out the phone and watch a TV show. He did not turn the volume down at all, so sometimes Jane could not hear what Tim said but could just see him looking at the phone. Besides, Tim sometimes got bored in the therapy process. He was distracted by the TV show and was disengaged with Jane. | Distraction | [82, 97] |
| *Scenario 2* | Jane is a therapist who works on general mental health problems. She was offering therapy to Jerry, who was burdened with anxiety because of marriage Issues. They were having a one-hour meeting via Zoom. | Jerry talked about his story with Jane. He was so upset that he started crying in the middle of his narrative. It was not obvious because he tried his best to restrain himself from crying. A window backlighted his face at that time. So Jane did not realize he was crying. It was about 10 minutes later when Jane noticed that. | Lack of non-verbal cues | [97, 114] |
| *Scenario 3* | Ben is a therapist who works on general mental health problems. He had a lot of clients at that time, and many of them scheduled remote therapy on that day. He usually worked from home. | Staying in the living room of his house, Ben started his remote therapy sessions. He just kept working without a break because he thought he was at home. During each therapy, he paid full attention to every detail of the clients. In the end, he felt exhausted and even cannot keep concentrate on the last therapy of the day. | Self-overlooked therapist burnout | [97] |
| *Scenario 4* | Steve is a therapist who works on trauma. This was the first time he met his client, Susan. Susan asked for his help because of PTSD after a traffic accident. | At the beginning of the meeting, Susan was super nervous so she cannot express her situation clearly. Although Steve wants to calm her down, what he can do is just say "That's fine". It took a long time for Susan to integrate into the online therapy environment. Therefore, this first-time meeting did not have as big progress as Steve once made with other clients in person. | Lack of trustworthy and safety | [82] |

## A.3 In-Scenario Activity: Design Decision Selection Responses

### *A.3.1 Scenario 1 (Distraction).*

Table 6. Scenario 1: Features the client gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 1/9 | 2/10 | 8/9 | 5/10 |
| *Feature 2 (Emotion and Sentiment)* | 1/9 | 0/10 | 7/9 | 6/10 |
| *Feature 3 (Stress)* | 1/9 | 2/10 | 4/9 | 2/10 |
| *Feature 4 (Biosignal Rate)* | 1/9 | 0/10 | 4/9 | 6/10 |
| *Feature 5 (Environmental Change)* | 1/9 | 3/10 | 6/9 | 2/10 |

Table 7. Scenario 1: Sensing system selections for the client

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 8/9 | 7/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 8/9 | 8/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 2/9 | 3/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 7/9 | 6/10 |
| *Sys 5 (Personal Biosignal Detection)* | 5/9 | 6/10 |
| *Sys 6 (Environmental Analysis)* | 7/9 | 5/10 |

Table 8. Scenario 1: Features the therapist gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 2/9 | 3/10 | 2/9 | 1/10 |
| *Feature 2 (Emotion and Sentiment)* | 2/9 | 4/10 | 5/9 | 2/10 |
| *Feature 3 (Stress)* | 3/9 | 1/10 | 3/9 | 0/10 |
| *Feature 4 (Biosignal Rate)* | 3/9 | 1/10 | 1/9 | 0/10 |
| *Feature 5 (Environmental Change)* | 1/9 | 0/10 | 1/9 | 0/10 |

Table 9. Scenario 1: Sensing system selections for the therapist

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 3/9 | 3/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 7/9 | 6/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 2/9 | 0/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 6/9 | 4/10 |
| *Sys 5 (Personal Biosignal Detection)* | 4/9 | 1/10 |
| *Sys 6 (Environmental Analysis)* | 2/9 | 0/10 |

### *A.3.2 Scenario 2 (Lack of non-verbal cues).*

Table 10. Scenario 2: Features the client gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 0/9 | 0/10 | 3/9 | 1/10 |
| *Feature 2 (Emotion and Sentiment)* | 0/9 | 1/10 | 7/9 | 8/10 |
| *Feature 3 (Stress)* | 0/9 | 1/10 | 6/9 | 4/10 |
| *Feature 4 (Biosignal Rate)* | 2/9 | 0/10 | 4/9 | 7/10 |
| *Feature 5 (Environmental Change)* | 1/9 | 2/10 | 4/9 | 6/10 |

Table 11. Scenario 2: Sensing system selections for the client

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 3/9 | 1/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 8/9 | 8/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 5/9 | 5/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 7/9 | 8/10 |
| *Sys 5 (Personal Biosignal Detection)* | 6/9 | 7/10 |
| *Sys 6 (Environmental Analysis)* | 5/9 | 8/10 |

Table 12. Scenario 2: Features the therapist gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 3/9 | 4/10 | 2/9 | 1/10 |
| *Feature 2 (Emotion and Sentiment)* | 2/9 | 3/10 | 5/9 | 2/10 |
| *Feature 3 (Stress)* | 3/9 | 1/10 | 2/9 | 0/10 |
| *Feature 4 (Biosignal Rate)* | 3/9 | 2/10 | 1/9 | 0/10 |
| *Feature 5 (Environmental Change)* | 1/9 | 1/10 | 1/9 | 0/10 |

Table 13. Scenario 2: Sensing system selections for the therapist

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 2/9 | 4/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 7/9 | 5/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 3/9 | 0/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 6/9 | 5/10 |
| *Sys 5 (Personal Biosignal Detection)* | 4/9 | 2/10 |
| *Sys 6 (Environmental Analysis)* | 2/9 | 1/10 |

### *A.3.3 Scenario 3 (Self-overlooked therapist burnout).*

Table 14. Scenario 3: Features the client gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 0/9 | 1/10 | 7/9 | 2/10 |
| *Feature 2 (Emotion and Sentiment)* | 2/9 | 0/10 | 5/9 | 5/10 |
| *Feature 3 (Stress)* | 1/9 | 2/10 | 4/9 | 3/10 |
| *Feature 4 (Biosignal Rate)* | 0/9 | 1/10 | 4/9 | 4/10 |
| *Feature 5 (Environmental Change)* | 0/9 | 0/10 | 4/9 | 3/10 |

Table 15. Scenario 3: Sensing system selections for the client

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 5/9 | 3/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 6/9 | 5/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 3/9 | 3/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 7/9 | 5/10 |
| *Sys 5 (Personal Biosignal Detection)* | 4/9 | 5/10 |
| *Sys 6 (Environmental Analysis)* | 4/9 | 3/10 |

Table 16. Scenario 3: Features the therapist gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 0/9 | 6/10 | 7/9 | 2/10 |
| *Feature 2 (Emotion and Sentiment)* | 1/9 | 4/10 | 5/9 | 2/10 |
| *Feature 3 (Stress)* | 2/9 | 8/10 | 6/9 | 1/10 |
| *Feature 4 (Biosignal Rate)* | 2/9 | 6/10 | 6/9 | 2/10 |
| *Feature 5 (Environmental Change)* | 1/9 | 2/10 | 5/9 | 0/10 |

Table 17. Scenario 3: Sensing system selections for the therapist

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 8/9 | 5/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 7/9 | 7/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 7/9 | 4/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 8/9 | 5/10 |
| *Sys 5 (Personal Biosignal Detection)* | 8/9 | 8/10 |
| *Sys 6 (Environmental Analysis)* | 6/9 | 2/10 |

### *A.3.4 Scenario 4 (Lack of trustworthy and safety).*

Table 18. Scenario 4: Features the client gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 1/9 | 1/10 | 5/9 | 6/10 |
| *Feature 2 (Emotion and Sentiment)* | 0/9 | 0/10 | 8/9 | 7/10 |
| *Feature 3 (Stress)* | 0/9 | 0/10 | 8/9 | 10/10 |
| *Feature 4 (Biosignal Rate)* | 1/9 | 2/10 | 6/9 | 8/10 |
| *Feature 5 (Environmental Change)* | 0/9 | 1/10 | 6/9 | 5/10 |

Table 19. Scenario 4: Sensing system selections for the client

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 6/9 | 5/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 8/9 | 9/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 8/9 | 8/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 7/9 | 9/10 |
| *Sys 5 (Personal Biosignal Detection)* | 7/9 | 10/10 |
| *Sys 6 (Environmental Analysis)* | 6/9 | 6/10 |

Table 20. Scenario 4: Features the therapist gets from sensing systems

| | Get but NOT share | | Get and share | |
|---|---|---|---|---|
| | Client Response | Therapist Response | Client Response | Therapist Response |
| *Feature 1 (Concentration)* | 1/9 | 3/10 | 4/9 | 1/10 |
| *Feature 2 (Emotion and Sentiment)* | 3/9 | 5/10 | 4/9 | 3/10 |
| *Feature 3 (Stress)* | 2/9 | 7/10 | 4/9 | 1/10 |
| *Feature 4 (Biosignal Rate)* | 2/9 | 4/10 | 3/9 | 3/10 |
| *Feature 5 (Environmental Change)* | 1/9 | 0/10 | 2/9 | 0/10 |

Table 21. Scenario 4: Sensing system selections for the therapist

| | Client Response | Therapist Response |
|---|---|---|
| *Sys 1 (Webcam-based eye tracking)* | 4/9 | 4/10 |
| *Sys 2 (Webcam-Based Facial Expression Analysis)* | 7/9 | 8/10 |
| *Sys 3 (Accelerometer-Based Tremor Detection)* | 4/9 | 6/10 |
| *Sys 4 (Verbal Sentiment Analysis)* | 6/9 | 6/10 |
| *Sys 5 (Personal Biosignal Detection)* | 5/9 | 7/10 |
| *Sys 6 (Environmental Analysis)* | 3/9 | 0/10 |

Received January 2024; revised July 2024; accepted October 2024