

Enablers and Barriers of Empathy in Software Developer and User Interactions: A Mixed Methods Case Study

HASHINI GUNATILAKE, Faculty of Information Technology, Monash University, Australia

JOHN GRUNDY, Faculty of Information Technology, Monash University, Australia

RASHINA HODA, Faculty of Information Technology, Monash University, Australia

INGO MUELLER, Faculty of Information Technology, Monash University, Australia

Software engineering (SE) requires developers to collaborate with stakeholders, and understanding their emotions and perspectives is often vital. Empathy is a concept characterising a person's ability to understand and share the feelings of another. However, empathy continues to be an under-researched human aspect in SE. We studied how empathy is practised between developers and end users using a mixed methods case study. We used an empathy test, observations and interviews to collect data, and socio-technical grounded theory and descriptive statistics to analyse data. We identified the nature of awareness required to trigger empathy and enablers of empathy. We discovered barriers to empathy and a set of potential strategies to overcome these barriers. We report insights on emerging relationships and present a set of recommendations and potential future works on empathy and SE for software practitioners and SE researchers.

CCS Concepts: • **Software and its engineering**; • **Human-centered computing** → **Empirical studies in collaborative and social computing**; **Collaborative and social computing theory, concepts and paradigms**;

Additional Key Words and Phrases: empathy, human aspects, software engineering, awareness, enablers, barriers, software developers, end users

ACM Reference Format:

Hashini Gunatilake, John Grundy, Rashina Hoda, and Ingo Mueller. 2024. Enablers and Barriers of Empathy in Software Developer and User Interactions: A Mixed Methods Case Study. *ACM Trans. Softw. Eng. Methodol.* 1, 1 (January 2024), 42 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

1 Introduction

Empathy has been recognised as a human aspect that can help to understand software developer and stakeholder human interactions [35]. It is also considered as a vital competency in many professional fields such as medicine [34, 46], healthcare [59], nursing [79], animal science [39], education [7, 53], professional writing and reviewing [16], marketing [20], and project management [27]. Empathy has been recognised as factor that improves inter-group attitudes and relationships [5].

Despite its prominence across disciplines, empathy remains a complex phenomenon with no one unified definition [12, 13, 19, 36, 57, 61]. One definition of empathy is “*the ability to experience*

Authors' addresses: Hashini Gunatilake, hashini.gunatilake@monash.edu, Faculty of Information Technology, Monash University, Australia; John Grundy, Faculty of Information Technology, Monash University, Australia, john.grundy@monash.edu; Rashina Hoda, Faculty of Information Technology, Monash University, Australia, rashina.hoda@monash.edu; Ingo Mueller, Faculty of Information Technology, Monash University, Australia, ingo.mueller@monash.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Association for Computing Machinery.

1049-331X/2024/1-ART \$15.00

<https://doi.org/10.1145/nnnnnnn.nnnnnnn>

the affective and cognitive states of another person, while maintaining a distinct self, in order to understand the other” [36]. This meta definition is derived from an inductive conceptual content analysis of the existing definitions, and provides a clearer outline of its fundamental dimensions. There are only a limited number of studies to date on empathy in SE [35]. Empathy has been discussed in the context of user experience (UX) with the use of personas [28, 51], used in combination with empathy map to improve personas development [29], as a tool for design thinking (DT) to better understand the requirements, thereby enhancing software quality [8] and addressing user privacy concerns [54]. In another study, researchers investigated the relationship between the collective empathy of software development teams and the effectiveness of their project process [2]. Here, collective empathy encompasses cognitive, affective, and behavioural empathy. They found that collective empathy impacts their project management effectiveness including team learning, product speed-to-market, and reduced project development costs. A common thread across these works is that *empathy seems vital for fostering human connections* [45] and that *having a good connection among software project stakeholders* may positively influence the success of a project. However, to the best of our knowledge, there are no studies that study empathy in the critical relationship of *software developers and end users*, exploring its enablers and barriers [35]. Due to the limited research in SE and the positive impact of empathy observed in other disciplines, we were motivated to explore this topic. To do this we conducted a mixed methods case study by exploring how empathy is practised between developers and end users in a specific software development project. Our study centred around the following research questions:

[RQ] To what extent do developers demonstrate empathy towards end users, and conversely, how empathetic are end users towards developers in the context of software development?

- What factors contribute to the empathy between software developers and end users?
- What factors impede empathy between software developers and end users?
- What steps were taken to address the factors that impede empathy between developers and end users?

In the absence of a universally accepted definition, we used the meta definition of empathy described earlier [36]. We employed the questionnaire of cognitive and affective empathy (QCAE), conducted observation of empathy cues [10, 26, 58], and interviews with developers and end users to collect both qualitative and quantitative data. We used socio-technical grounded theory (STGT) for qualitative data analysis [41] to analyse the qualitative data and measures of dispersion in descriptive statistics for our quantitative data analysis.

The key contributions of this work include the identification of:

- The awareness required to trigger empathy and enablers of empathy between developers and end users;
- Some key potential barriers to empathy of developers and end users, and strategies to overcome these barriers;
- A set of actions to manage empathy enablers and barriers in SE projects; and
- A set of recommendations derived from our findings for software practitioners and SE researchers.

The rest of this paper is organised as follows. Section 2 introduces definitions of key empathy specific terminology used throughout this paper. Section 3 presents an overview of key related studies. Section 4 describes our research methodology. Section 5 describes the key findings from this work, and Section 6 discusses our insights, implications, recommendations for practitioners & researchers and directions for future work. In Section 7, we present limitations of this study, followed by our conclusions in Section 8.

2 Glossary of Terms

We use a range of empathy-specific terminology throughout our paper and brief definitions of these terms are provided in Table 1.

Table 1. Glossary of Terms

| Term | Definition |
|--------------------------|--|
| Empathy | The ability to experience affective and cognitive states of another person, while maintaining a distinct self and understand the other [36]. |
| Cognitive Empathy | The ability of a person to consciously detect and understand the internal states of others [30, 73]. |
| Affective/Emotional Emp. | The ability of a person to perceive and share other individual's emotional states and feelings [17, 49]. |
| Behavioural Empathy | This consists of two types of empathic behaviours i.e., <i>behaviour mirroring</i> and <i>empathic communication</i> . Expressions, mannerisms, postures, & gestures of other person is referred as behaviour mirroring. Empathic communication is defined as intentional behaviour that displays cognitive and/or affective empathy towards the other person. |
| Sympathy | An emotional response stemming from the apprehension of another's emotional state or condition that is not one's own state or condition, but consists of feelings of sorrow or concern for the other [22, 23]. |
| Perspective Taking* | The ability of a person to see the situation from another person's perspective [49, 62]. |
| Online Simulation* | The ability of a person to understand and mentally represent how another person is feeling [49, 62]. |
| Emotion Contagion* | The ability of a person to reflect self oriented emotions while noting the emotional states of others [49, 62]. |
| Proximal Responsivity* | A person's emotional reaction to the moods of another person, who is physically or emotionally close to them. |
| Peripheral Responsivity* | A person's emotional reaction to the state of moods of another person, who is not close to them or they do not know the person at all [49, 62]. |
| Empathic Concern | The tendency of a person to experience feelings of warmth, compassion and concern for others undergoing similar experiences [14]. |
| Emotional Intelligence | The ability to monitor one's own and others' feelings and emotions, to discriminate among them and to use this knowledge to guide one's thinking and actions [64]. |
| Social Awareness | The awareness of others' emotions [32]. |
| Collectivism | "Societies in which people from birth onwards are integrated into strong, cohesive in-groups, which throughout their lifetime continue to protect them in exchange for unquestioning loyalty" [44]. |
| Individualism | "Societies in which the ties between individuals are loose; everyone is expected to look after themselves & their family" [44]. |

* A subscale of QCAE

3 Related Work

Empathy is widely regarded as a multidimensional construct [12, 13], with *cognitive*, *affective* and *behavioural* empathy as its key three dimensions [12]. Others have argued there are only two dimensions to empathy, cognitive and affective [13]. There are two key commonly used methods to measure empathy: measuring empathy via self-assessments, and via neurophysiological examination methods of studying different brain activities using brain images. Researchers have also measured empathy through observations, based on demonstrated empathy behaviours. These empathy behaviours can be divided into two main categories, verbal and nonverbal behaviours, that have been widely studied [10, 18, 26, 58].

Verbal empathy can be broken down into four behaviours [26], empathic understanding responses, empathic affirmations, empathic evocations, and empathic conjectures. Non-verbal empathy behaviours consider *sounds and signals* to understand whether people show interest in others [58]. In their study, attentive listening was identified as one such a behaviour including: sounds as verbal acknowledgements without interrupting the flow of other person's speech; repeated nodding, tilted head, and placing thumb and forefinger of one hand on the chin as signals for attentive listening [58]. Wide open eyes and raised eyebrows also have been classified as

signals for attentive listening [25] and for people feeling addressed and therefore actively participating [4]. These mimicking behaviours are known to occur in combination with each other and often associated with an activated posture such as bending forward towards the other person. In addition to the sounds and signals of attentive listening, Chartrand et al., explained *similar facial expression, posture or gesture* as a nonverbal empathy behaviour [10].

Research across disciplines supports strong association between *similarities* and empathic behaviour, such as in psychology [6, 24, 40, 43, 78], neuroscience [60, 76, 77], and philosophy [68]. Similarities can be defined as sharing a similar experience or demographic characteristic with another person [50]. The study by Motomura et al. highlighted the nuanced effects of familiarity on empathy valence, distinguishing between positive and negative empathy [56]. Specifically, it found that positive empathy, linked to activities like nursing or support, is less likely to occur with strangers. In contrast, negative empathy, associated with adverse situations such as danger, can readily happen even with unfamiliar individuals. Investigating empathic brain activity, the study suggested that *familiarity* acts as an enabler for positive empathy. Building on this, researchers have argued that the concept of *relatedness* is crucial for human empathy, going so far as to propose that familiarity can be more accurately defined as relatedness [1].

The studies on *culture* and empathy reported that cultural backgrounds of people influence their empathic behaviour [50]. It is commonly believed that empathic behaviours are more valued in collectivistic societies than individualistic societies [37, 52, 67]. A study found a positive relationship between affective empathy and country-level index of collectivism, but no association was found between cognitive empathy and country-level index of collectivism/individualism [11]. However, another study found that individualistic societies have higher empathy compared to collectivistic societies [9]. This was due to their ability to have an independent cognitive and emotional state by differentiating their own emotions from others. However, most of the studies on culture and empathy have found higher associations between collectivism and dispositional affective and cognitive empathy [11, 75].

Empathy barriers have been studied primarily in healthcare domain [21, 38, 48]. A study on healthcare identified *time pressure, conflicting priorities, bureaucracy* as barriers to empathy [48]. *Anxiety, inability to understand the relationship between illness and patients' emotional needs, and negative emotions due to tensions with patients* have been identified as barriers to physicians' empathy towards patients [38]. This study found that physicians become anxious due to time pressure hence there is a tendency of not listening to patients. Another study found empathy barriers of oncology nurses. This study found three types of barriers: those related to nursing, healthcare, and cancer care [69]. *Lacking compassion, disinterest in oncology nursing & self-criticism, psychological distress* were identified as barriers related to nursing. Barriers related to healthcare included *job strain, task centeredness rather than patient-centeredness, lack of formal training for empathy with cancer patients, lack of manager support, and nurse-patient gender imbalance*. Barriers related to cancer care were *difficulty of maintaining empathy with cancer patients and feeling of uselessness of care for cancer patients*. In summary, familiarity, similarities, relatedness, culture were identified as empathy enablers, and time pressure, negative emotions, disinterest, task centredness were identified as empathy barriers.

Many empathy measures, models and techniques have been developed over the years. In our previous work, we developed a preliminary empathy taxonomy for SE considering these empathy models, techniques and measures [35]. Since empathy is seen as beneficial for improving human connections across disciplines, and has not been adequately studied in SE, we became interested to explore it in the context of one of the prominent relationships in SE contexts, between software developers and end users.

4 Research Design

We wanted to understand how empathy is practised between software developers and end users. Our topic lend itself well to a mixed methods case study, where we could investigate the phenomenon in depth, from multiple angles using multiple approaches to understand it up-close and in depth. We adhered to the guidelines outlined by Runeson et al. [63] when conducting this case study.

4.1 Case Description

The Project: The selection of the *AskPCOS project* aligns with our research questions on factors influencing empathy between software developers and end users. Access to end users is notoriously challenging to arrange and often limits what can be researched and how. Our access to the AskPCOS project provided us the level of access required to do this in a real-world SE case setting. This project's context offers a practical environment to explore the dynamics of empathy, ensuring our study is grounded in real-world complexities and user engagement constraints. Thus, the AskPCOS project serves as an ideal and relevant case, allowing us to delve into the nuanced factors that contribute to or impede empathy between software developers and end users. The project involved the development of an engaging extension that integrates SMART (Specific, Measurable, Achievable, Relevant, and Time-Bound) goal setting process into the existing AskPCOS web application.¹ This solution helps women with Polycystic Ovary Syndrome (PCOS) to better manage their lifestyle. This software project intended to digitise questionnaires, checklists, and other paper based material related to PCOS lifestyle education sessions conducted by the Monash Health public PCOS clinic.

The fourth author supervised this software project, facilitating our access to it. Healthcare professionals from Monash Centre for Health Research and Implementation (MCHRI)² advertised the project in Facebook groups dedicated to PCOS patients. The PCOS patients who expressed interest were then recruited as study participants. Throughout the study, we maintained a consistent group of PCOS patients.

The development project was carried out over a period of 24 weeks and there were two iterations within it. The main stakeholders in the development project were Healthcare professionals from MCHRI (domain experts) and two academic staff from Monash Art, Design & Architecture (MADA) faculty.³ Our study had several multi-disciplinary participants: two Masters level design students from MADA (UX designers), six final year undergraduate IT students from Monash Faculty of Information Technology (FIT) (developers), and four patients with PCOS (target end users). These stakeholders and participants had multiple interactions throughout the project.

The People: Our study⁴ was based on the usability testing sessions conducted with the participation of software developers and potential end users. The purpose of these sessions were to evaluate the usability of the newly implemented SMART goal setting process of the AskPCOS web application. Usability testing sessions provided a platform for developers and users to directly interact with each other. All developers had the opportunity to directly interact with at least one user by acting as the hosting developer of the usability testing sessions. Likewise all users were able to interact with developers. All participants stated that they were able to build a better connection due to interaction and/or familiarity with each other.

¹<https://www.askpcos.org/>

²<https://www.monash.edu/medicine/mchri>

³<https://www.monash.edu/mada>

⁴Approved by Monash Human Research Ethics Committee. ERM Reference Number: 32281

Table 2. Demographics of the Interview Participants

| Participant ID | Age Group | Education | Experience in working with teams | Ex. in working with people from diverse backgrounds | Ex. in working with patients/people with sensitive cond. | Ex. in working with customers/users | Ex. in using eHealth apps | Affinity to tech. vs people |
|-------------------|-----------|---------------|----------------------------------|---|--|-------------------------------------|---------------------------|-----------------------------|
| Developers | | | | | | | | |
| P1 | 20-25 | SD*, BE* | Yes | Yes | No | No | No | SHC* |
| P2 | 20-25 | SD, BE | Yes | Yes | No | No | Yes | STC |
| P3 | 20-25 | DD*, BE+BCom* | Yes | Yes | No | No | No | STC |
| P4 | 20-25 | SD, BE | Yes | Yes | Yes | Yes | No | STC |
| P5 | 20-25 | DD, BE+BCom | Yes | Yes | No | Yes | No | STC |
| P6 | 20-25 | DD, BE+BCom | Yes | Yes | No | No | No | FTC |
| Users | | | | | | | | |
| PU1 | 20-25 | - | Yes | Yes | - | - | Yes | SHC |
| PU2 | 31-35 | - | Yes | Yes | - | - | Yes | FTC |
| PU3 | 41-45 | - | Yes | Yes | - | - | Yes | FHC |
| PU4 | 26-30 | - | Yes | Yes | - | - | No | FHC |

*SD: Single Degree, DD: Double Degree, BE: Bachelor of Engineering, BCom: Bachelor of Commerce, FHC: Fully Human Centric, SHC: Somewhat Human Centric, STC: Somewhat Technology Centric, FTC: Fully Technology Centric

Table 2 shows key participant information. We interviewed six developers who were final year SE students and four end users who were females with the PCOS condition. A majority of developers and end users were from Australia except for one developer from Hong Kong and a user from Zambia. All the developers and end users were residing in Australasia while conducting this study. All the developers belonged to the age group 20-25 and there was only one female developer among them. End users were from four different age groups including 20-25, 26-30, 31-35 and 41-45. Half of the developers were following a single degree which is Bachelor of Engineering (BE) in SE and rest of the developers were following double degrees. They were following Bachelor of Commerce majoring in different areas like business analytics, finance and econometrics along with the Bachelor of SE. All the developers and users confirmed that they had prior experience in working in a team and with people from diverse backgrounds. Only one developer had prior experience in working with patients or other groups of people who tend to have some sensitive condition. Two developers had prior experience in working with customers/users in other university classes or in industry experience. Except for one user, the rest of the users had not interacted with software developers in any form prior to the study. Only one developer had used an eHealth app before, however three out of four end users had prior experience in using eHealth apps. From all the participants we inquired the affinity to technology and people on a scale of zero to three, where zero was fully human-centric, one was somewhat human centric, two was somewhat technology-centric and three was fully technology-centric. Most of the developers self-identified as technology-centric and most of the users self-identified as human-centric.

The Process: In the first 12 weeks, design students from MADA had discussions with healthcare professionals from MCHRI and end users to get a better understanding of the manual goal setting process. Then design students created some design prototypes by aligning user needs and the identified issues. The design students conducted multiple co-design sessions with end users to refine and identify potential solutions. After this, the design students created high fidelity

wire-frames by incorporating end user feedback received during these sessions. In the next 12 weeks, project stakeholders were limited to development students and end users as the designers had completed their portion of the work. The developers started development of the application based on the finalised design prototype. Developers conducted usability testing sessions with users. Usability testing sessions were conducted with the participation of two developers and an end user. There were multiple usability sessions between developers and end users. During these usability sessions end users were instructed to follow a list of tasks to evaluate the usability of newly developed features of the application. One developer hosted the session by providing instructions to end users (*hosting developer*) and the other developer took meeting minutes (*minute taking developer*). End users provided feedback regarding their experience during these sessions. The co-design sessions, usability testing sessions and all the other sessions were conducted online via the Zoom.

The Case Selection: Empathy involves nuanced human interactions. A mixed-methods approach facilitated a comprehensive understanding of these interactions. Qualitative methods, such as interviews and observations, helped us to capture the richness of individual experiences and perceptions of developers and users, while quantitative methods such as empathy test scores helped to analyse empathy trends. By combining qualitative and quantitative data, we achieved a more holistic view into the phenomenon and subsequently a more rounded form of analysis. This approach enabled triangulation, where findings from different methods were compared and validated, which enhanced the overall reliability and validity of the study. In addition, empathy involves interpersonal dynamics that may not be fully captured through purely quantitative means. An up-close and in-depth investigation, facilitated by qualitative methods, allowed us to delve into the nuances of human interactions and gain a more in-depth understanding of the empathetic processes at play.

4.2 De-identification Process

We followed a rigorous de-identification process to protect the confidentiality of participant data and to ensure participant data is untraceable. It was essential to implement a rigorous process for de-identification as there was a limited number of participants in our case study including only one female developer. In order to respect the confidentiality of our participants, we refer to them by numbers P1-P6 for developers and PU1-PU4 for users in this paper (Table 2). We intentionally omit specific participant numbers for some quotations in the findings section to ensure participant anonymity and maintain confidentiality.

4.3 Data Collection

Figure 1 shows an overview of our study methodology and outcomes of each procedure. In the first 12 weeks, we conducted a pilot study. We observed all the co-design sessions and distributed a reflection form to the participants of each session. This reflection form served as a retrospective of co-design sessions. We asked participants what went well, what didn't went well, what could be improved, and what were the surprising events happened during the sessions. In the next 12 weeks, we used three data collection methods: empathy test, observations of the usability testing sessions between developers and end users, and semi-structured interviews. The latter provided the most in-depth insights, helping confirm, clarify, and explain many of the insights captured from the other two methods.

4.3.1 Questionnaire and Empathy Test We created two questionnaires comprising of three main types of questions: basic demographic questions, questions related to experience of working in teams and usage of eHealth apps, and an empathy test. The questionnaires were distributed

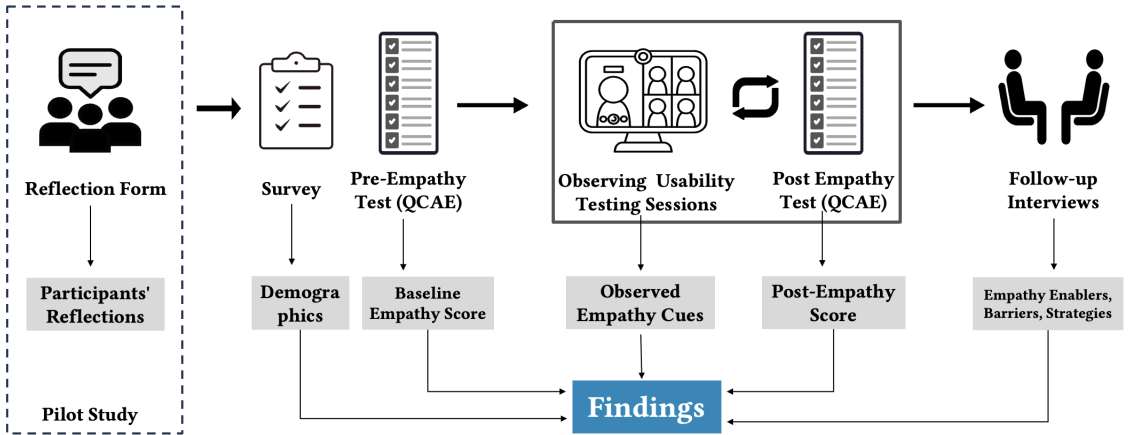


Fig. 1. Overview of the research methodology

among developers and end users via the Qualtrics online survey platform. Participants were instructed to complete the first questionnaire prior to their first usability session (round zero). We considered the empathy score calculated in round zero as the baseline empathy of the participants. Then the participants were instructed to complete a short version of the same questionnaire after each usability session. This short questionnaire was created by reducing the demographic and team related questions used in the original questionnaire. Along with the empathy test, we kept the questions related to the demographics and employment which had a possibility of changing during the course of the research.

Empathy scale selection was carried out with the help of empathy experts. We considered four prominent empathy scales: Interpersonal Reactivity Index (IRI) [15], Empathy Quotient (EQ) [3], The Jefferson Scale of Empathy (JSE) [47] and a Questionnaire of Cognitive and Affective Empathy (QCAE) [62]. We removed JSE as it has been developed specifically for health professionals and patients. We then discussed with two empathy experts (PhD in Psychology, PhD Candidate in Neuroscience Dept. of Monash Central Clinical School with special focus on empathy) to select the most appropriate scale for our study. Based on the feedback of our empathy experts – we removed EQ as it is mainly used for assessing Autism. We then removed IRI as it seems irrelevant to our study because it assesses processes broader than empathy [3, 35]. Despite some of the irrelevant items in its peripheral responsiveness subscale, QCAE appeared to be the most appropriate empathy scale to our study. Even though there are no empathy scales designed specifically for SE [35], we identified QCAE would be suitable to serve the purpose of our study especially due to its focus on cognitive empathy and perspective taking.

QCAE consists of two main components and is divided into five different subscales. These two components are cognitive empathy and affective empathy. The five subscales are perspective taking which consists of ten items, online simulation which consists of nine items, emotion contagion which consists of four items, proximal responsiveness which consists of four items, and peripheral responsiveness which consists of four items. Altogether QCAE comprised of 31 items and participants are required to indicate how each item best describes them via a Likert scale from “Strongly Agree” to “Strongly Disagree”. The scores of subscale items are summed to produce the total subscale scores. Cognitive empathy score is calculated by summing the scores

of two cognitive subscales and affective empathy score is produced by summing the scores of three affective subscales. The focus of each of these subscales are explained in Table 1.

4.3.2 Observations We observed the usability testing sessions in the second 12 weeks. The purpose of our observations was to identify the empathy cues demonstrated during the interactions between software developers and users. We found a set of verbal and nonverbal empathy cues from existing literature [10, 26, 58]. Two of the authors independently observed the usability testing sessions and manually recorded these verbal and nonverbal empathy cues demonstrated by developers and users. We observed seven usability testing sessions with each session lasting approximately 45 minutes. We used a Google form to collect responses and the two observers filled out this form separately during each session. At the end of our case study, the first author compared and collated the observation notes of both observers. The first author also re-watched all the video-recorded observation sessions multiple times to record the occurrences of empathy cues (see Tables 6, 7, 8, 9). Apart from the empathy cues, we also recorded some supplementary details such as the mood of the session, how developers responded to the communication challenges of users and how developers responded to negative and positive emotions of users or vice-versa.

We implemented several measures to prevent our observations from impacting the usability sessions, aiming to maintain an unobtrusive and non-disruptive process [70, 71]. To minimise any potential influence, observers maintained a passive role by refraining from speaking and keeping their videos off throughout the sessions. Additionally, participants were informed in advance that the observations were solely for research purposes and would not use them for any other intent.

Verbal Empathy Cues: We observed following four verbal empathy behaviours [26] and these were identified as the verbal empathy cues during this study (see Table 3). These four verbal empathy behaviours were demonstrated in the usability testing sessions as: answers with reference to the content of the other person; content-consistent repetition of the statement of the counterpart; questioning to better understand what the other person is saying with reference to the content of the other person's statement; confirmation of the other person or showing understanding for their statement.

Nonverbal Empathy Cues: We identified facial, vocal and posture related signals as nonverbal cues. Nonverbal behaviours were identified from two existing studies [10, 58]. We observed the following three manifestations of these nonverbal empathy behaviours and these three manifestations were identified as the nonverbal cues during this study (see Table 4).

4.3.3 Semi-Structured Interviews We collected data by conducting online, semi-structured interviews with ten developers and end users using open-ended questions. The interviews were approximately 30-45 minutes long. These interviews were conducted at the end of the 24-week period after all the usability testing sessions were completed. Due to the pandemic and geographical distribution of the participants, all of the interviews were conducted online via Zoom, and all were audio recorded.

Interviews focused on the experience of the participants during the usability testing sessions, in particular what made them empathetic or what inhibited their empathy towards the other group. We created two interview guides for developers and end users comprised of core questions related to enablers and barriers of empathy as well as some associated questions derived from the observations. These interview guides were employed during the interviews together with appropriate follow-up questions based on the flow of the conversation.

Table 3. Verbal Empathy Cues

| Verbal Cue | Description |
|----------------------------------|---|
| Empathic Understanding Responses | <i>A person responds with reference to the content of the other person.</i> The first behaviour is an empathic statement that demonstrates a person has understood what the other person is saying. It is a response related to the content of the person's statement. The person answers the other person, but in doing so the person refers to the content of the pre-statement of the other person. For example, the person could demonstrate this by <i>repeating or summarizing the content of the other person's statement when answering.</i> |
| Empathic Affirmations | <i>One person repeats the statement of the other person with the same content.</i> The second behaviour is verbal confirmation from the other person, in which the person tries to <i>express understanding</i> for the other person. This behaviour is demonstrated <i>by summarising the content of the statement made by the other person.</i> In this behaviour, one person repeats the statement of the other person without any evaluation or answer. In this way, the first person shows that they have understood what the other person said. |
| Empathic Conjectures | <i>A person asks in order to better understand what the other person is saying, with reference to the content of the other person's statement.</i> The third empathetic behaviour is a substantive question about the statement of the other person. It is a <i>general question about the content of the other person's statement</i> , in which one person tries to understand part of the content of the other person. When demonstrating this behaviour, the first person asks the other person a question, but this shows that they have understood the content of what the other person said. |
| Empathic Evocations | A person confirms the other person or shows understanding for their statement. The fourth behaviour is expressing the experience of the other person in different words, where the person tries to <i>summarize the experience of the other person</i> . This is often done in the <i>form of a question</i> . This empathy behaviour consists of affirming or showing understanding to the other person. When demonstrating this behaviour the person refers back to the statement of the other person and the person could, for example, praise the other person or could show understanding for a difficult situation of the other person. |

Table 4. Nonverbal Empathy Cues

| Nonverbal Cue | Description |
|---|---|
| Lute for careful listening | Involves sounds for attentive listening. This behaviour provides a verbal acknowledgement without interrupting the flow of the other person's speech or speaking directly themselves. For example this <i>involves sounds like "Mmm", "Ah", short, monosyllabic words such as "Ok", "Yes", "Oh yeah".</i> The sounds show that an empathic person is listening to the other person or can understand what the other person is saying [58]. |
| Nonverbal signals for attentive listening | Refers to the facial expressions and attitudes of the empathic person. We observed five signals for attentive listening: <i>consists of four signals related to facial expressions and one signal related to the attitude of the empathic person</i> . We recorded five signals for attentive listening as <i>multiple nods, wide open eyes & raised eyebrows which sometimes create transverse wrinkles on the forehead, a head tilted slightly to one side, the thumb & forefinger of one hand on the chin, and an empathic posture where a person bends forward towards the other person and builds up slight tension in the upper body.</i> This empathic posture is an indicator that the person feels addressed or actively participates in the conversation. |
| Similar facial expressions, posture, gestures | This indicates having similar facial expressions, posture, gestures of the empathic person as their counterpart. Ekman assigned specific facial signals to four basic emotions [25]. These emotions namely joy, anger, fear, mourning were used to identify similar facial expressions. In addition to similar facial expressions, we also observed similar body postures and similar gestures between developers and users [10]. |

The questions included in the interview guides (see Appendix B) are generally straightforward, with the exception of those related to the concept of empathy. Recognising the potential for ambiguity in the interpretation of empathy, we provided a clear definition of empathy (refer Table 1) at the beginning of the interview to guide participants. Additionally, we maintained an open line of communication during the interviews, encouraging participants to seek clarification on

any unclear or ambiguous questions. To further ensure the questions' clarity and appropriateness, the first author sought input from other, more experienced, co-authors. Their feedback helped in refining the wording of questions and addressing any potential confusion. Due to the nature of this case study, conducting a traditional pilot interview study was challenging, as the questions were closely tied to the project's details and participants' involvement. While a traditional pilot study was not feasible, these measures were taken to ensure the clarity and appropriateness of the interview questions within the unique context of this case study. Further during the interviews, we didn't directly inquire about participants' empathy enablers and barriers. Instead, our approach involved prompting them with queries about instances when they perceived successful or challenging empathetic experiences. This method was chosen to allow participants to naturally share insights into their empathy without explicitly focusing on enablers and barriers. While this approach may have limitations, we believe it contributed to the authenticity of the data we gathered.

4.3.4 Triangulation Three types of triangulation were used in this study [63]: (a) *Data source triangulation* was used in two ways: by collecting data from multiple sources such as interviews, questionnaires, observations and by collecting same data at different occasions for example administering empathy test to developers and users at different stages; (b) *Observer triangulation*, by involving multiple researchers for the analysis of same data sources such as having two researchers to observe usability testing sessions and involving two researchers for the analysis of interview transcripts; and (c) *Methodological triangulation*, by combining different types of data collection methods such as qualitative (interviews, observations) and quantitative methods (empathy test).

4.4 Mixed Methods Data Analysis

We used mixed methods to analyse data. *Qualitative data analysis* was mainly used to analyse interview data using STGT for data analysis [41]. *Quantitative data analysis* was mainly employed to calculate the scores of empathy test and also to analyse trends using these calculated scores.

4.4.1 Qualitative Data Analysis We used qualitative data analysis for analysing interview and observational data. We transcribed the recordings of interviews, then stored and analysed them using NViVo software. We used STGT [41] to analyse interview transcriptions. STGT has been formalised as a method particularly suitable for technology-intensive domain studies such as SE [33].

For this study, we specifically used the *STGT for data analysis*. Our interview responses provided sufficient qualitative data for applying coding techniques but were not sufficient for full theory development. Therefore, a limited application of STGT for data analysis was found suitable for our study. We selected this approach over other qualitative analysis methods, such as thematic analysis, due to three primary reasons: (i) our study's alignment with the socio-technical (ST) research framework proposed by STGT; (ii) the rigorous nature of STGT, resulting in original, relevant, and multidimensional outcomes; and (iii) its reflective practices, such as memoing, which facilitated layered insights and reflections. As stated before, our study aligns with the dimensions of the ST research framework proposed by STGT: ST phenomenon, ST domain and actors, ST researchers, and ST tools, techniques and data. We studied how empathy is practised between developers and users which is a ST *phenomenon* because we are exploring the technical consequences of a social phenomenon i.e., empathy. The *domain* of our study is SE which is a ST domain due to "tight coupling between its social and technical aspects" [41]. Also software developers and users are the *actors* who play key roles in our ST phenomenon. In this study, these actors not only use their knowledge to create technology but also use their experience to

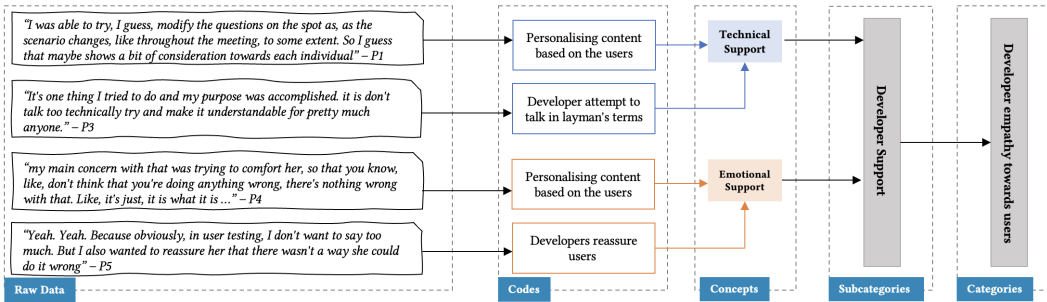


Fig. 2. Emergence of the category ‘*Developer empathy towards users*’ from raw data → codes → concepts → subcategory → category through constant comparison

increase the impact of these technological solutions. Our research team consists of ST *researchers* who have requisite knowledge and skills in qualitative research, philosophical foundations, and technical experience. As ST researchers, our team has necessary technical skills and domain knowledge. The interviews were conducted by the first author who is an experienced software practitioner and an early career researcher, supported by a supervisory team with strong research and technical skills. This enabled us to understand the processes and experiences of participants. The study used several ST *tools, techniques* and *data* including NVivo and Zoom. We used data collection techniques such as interviews, observations & questionnaires to collect both qualitative and quantitative data. STGT is particularly designed to capture this type of an ST research context which is different to the native social context of traditional GT methods [41].

Open coding was used to analyse the interview transcripts (raw data). The first author open coded all the interview transcripts and derived a preliminary code book. This preliminary code book was peer reviewed by third author who is an experienced grounded theorist. Based on the peer review discussion, the first author refined the preliminary code book and conducted constant comparison. Constant comparison was used to compare the codes within the same and across different transcripts. Similar codes were then grouped to form concepts, similar concepts were grouped to form subcategories and similar subcategories were grouped to form categories. An example of the STGT analysis is illustrated in Figure 2. For a more comprehensive example, the supplementary information package is accessible online.⁵

Memos are free-flowing ideas about the codes and emerging relationships [42]. We used basic memoing to record researcher reflections in detail. We wrote memos when we had insights about codes and their emerging relationships. Memoing also helped us to systematically document reflections on emerging concepts, subcategories, categories and possible links between them. For example, see memo in Appendix A. In addition, we qualitatively matched developer and user behaviours observed during usability testing sessions to empathy cues by referring to the identified verbal and nonverbal empathy cues (Table 3 and 4).

4.4.2 Quantitative Data Analysis We used quantitative methods to analyse empathy test scores and represent analysed qualitative empathy cues in a numerical format. For empathy test related analysis we used measures of dispersion in descriptive statistics. Measures of dispersion describe the dispersion or the spread of data [74]. We used box plots to visualise data as we can easily get an overview of the data set and identify outliers. We calculated quartiles, interquartile range

⁵<https://github.com/Hashini-G/SupplementaryInfoPackage-AskPCOSStudy>

| | | | |
|--|-------------------|---|-----|
| 5.3 | Awareness | Enablers | 5.4 |
| Developer awareness of users <i>Emotional: User needs, Concerns, Reactions, Feelings</i> <i>Technical: User understanding, Technical proficiency</i> Importance of empathy Developer self-awareness <i>Technical jargon, Differences in empathy, Less developer awareness of user's statements</i> | | Enablers of developer empathy <i>Reflection in action, Commonality, Familiarity, Understanding, Appreciation, Connection, Interaction, Awareness, User confusion, Involvement, Gender similarities, Similar level of technical ability</i> Enablers of user empathy <i>Familiarity & Connection, Connection to app's goal, Appreciation, Understanding of developer thought process, Understanding of developer needs, Mode of communication</i> | |
| Empathy in Developer - User Interactions | | | |
| 5.6 | Strategies | Barriers | 5.5 |
| Developer support <i>Technical support, Emotional support</i> Prospective developer acts of empathy <i>Better preparation, More follow-up questions, Slow down the session, Support users more, Taking time to implement user suggestions</i> | | Barriers to developer empathy <i>Less interest, Unfamiliarity, Limited time for reflection in action, Less interaction, Difficulty to resonate, More concentration on the task, Poor connection with users, Developer difficulties</i> Barriers to user empathy <i>Poor connection, User nervousness, Users feeling awkward, Confusion on terminology</i> | |

Fig. 3. Overview of Key Interview Findings derived using STGT Analysis.

(IQR), maximum and minimum values of empathy test scores. This data was then used to draw box plots. We drew box plots to visualise the dispersion of baseline empathy scores, round one empathy scores and round two empathy scores of both developers and end users. These scores were visualised considering the empathy scores of five subscales of QCAE. We also illustrated total empathy scores, total cognitive scores and total affective scores of each round using box plots. In addition to these group based visualisations, we used line graphs to illustrate empathy scores of each individual participant. These line graphs were used to visualise individual scores in each round considering five subscales of QCAE, total cognitive, total affective and total empathy scores. We further analysed qualitative empathy cues by converting them to a quantitative measure by calculating occurrences of each empathy cue. We recorded how many times each of these empathy cues were demonstrated by developers and users during each usability session. This quantitative representation is explained in Section 5.5.

5 Findings

Figure 3 outlines the key findings of our interview study analysis. Through the interview analysis, we found evidence for developer awareness (Section 5.1), enablers of developer & user empathy (Section 5.2) as well as barriers to developer & user empathy (Section 5.3). We also found a set of potential strategies developers may use to improve empathy (Section 5.4). We report the empathy cues identified during our observation study (Section 5.5) and trends discerned from our empathy test (Section 5.6). We include pertinent quotations derived from our interviews that provide valuable insights into the underlying concepts. We provide references to all the participants whose interviews generated one or more codes for a particular concept. However, due to confidentiality issues we cannot include all the underlying quotations from our interviews.

5.1 Developer Awareness

It became apparent through our data analysis, that awareness in terms of *developer awareness of users* and *developer self-awareness of own actions and emotions* was a first step towards enabling empathy.

5.1.1 Developer Awareness of Users From the answers of developers, we noticed that they had a good *emotional and technical awareness of users*. When considering emotional awareness, developers noticed both *positive and negative user emotions*, and were more concerned about negative user emotions. We asked developers about the moments that they enjoyed the most during the sessions. Developers stated that they enjoyed the session introduction more as they felt users were more relaxed during that period compared to the rest of the session [P1, P3]. Occasionally, developers sensed that users might have felt like they were being tested during the usability sessions [*negative emotions*]. Developers said that it was not nice to observe such user emotions as the underlying purpose of conducting these sessions was only to measure the usability of the app and not the capabilities of users [P5]. When asked, developers noted that they could have empathised better with users when the users were seen to be stressed or under pressure [*negative emotions*]. However, developers faced a dilemma of whether to act on their empathy by helping users resolve their confusions or to let them struggle and find their own way, which would allow elicitation of unadulterated user feedback on the software. Developers stated that users had quite amusing reactions [*positive emotions*] when they explained the purpose of the session, which is to only test the app and not the users. Developers said that users seemed much comfortable when the developers explained there were no right or wrong answers. Overall, they were able to resonate with user emotions and reasons for these emotions based on how the users acted, despite not being able to fully understand all their concerns regarding the app [P3, P6].

“...in user testing, I don't want to say too much. But I also wanted to reassure [user] that there wasn't a way [user] could do it wrong. Like we weren't testing [user]. But I think [user] felt like that, which wasn't that nice.” - P5

As observers of these usability sessions, we noticed developer remarks to users on not testing user spellings (“*it's okay, we are not testing your spellings*”) when users apologised for entering the wrong spellings. When we asked developers about the intention of their remarks, developers explained their genuine intention was to inform users that they don't have to be perfect. However, reflecting on these remarks developers stated that their words might have sounded judgemental from a user's perspective [P2]. This suggests that developers were considerate about user feelings, not just during the usability sessions, but even when reflecting on their remarks after these sessions.

Developers had a good understanding of *users' needs and concerns*. Developers stated that they understood the *need* of the users to have options to check-in for goals both daily and weekly. Developers stated this was due to user's ADHD (Attention-deficit/hyperactivity disorder) condition that made user well focused on some days and out of focus on other days. Even without having ADHD, they could imagine how users would feel in those instances [P3]. This may be an instance where the developers exhibited *perspective taking*. Another developer said that the interaction during usability sessions helped them to understand user confusions [*concerns*]. The developer said that they understood the reasons for such user confusions by observing mouse movements of users, listening to users' explanations and where they are leaning towards in those situations, enabling their empathy towards users [P4].

It was clear that the developers were *aware of users' technical competence*. Some users were seen to be technically proficient [P2] while others suggested perceived gaps in technical understanding that did not hinder the tasks and app use [P6]. P4 said having family members who were not very tech savvy allowed them to understand their users, demonstrating *perspective taking*. Users listed difficulties in understanding some terms used in the app. However, they didn't want to discourage the developers and make them feel bad about their app [PU3]. This indicates some evidence of users' awareness of developer emotions which can be a gesture of goodwill and/or an indication of the connection they share.

5.1.2 Developer Self-Awareness of Own Actions & Emotions When asked, one developer stated that they thought there was not much to empathise with end users in this particular project where the only connection between developers and users was during the usability sessions [P5]. However, upon continuing the discussion on the potential acts of empathy, the same developer stated that they were able to empathise with end users fairly well. Another developer stated that they didn't observe any significant user issues during the usability sessions, hence they didn't feel the need to empathise with users [P6]. Having said that, the same developer stated that they were able to understand users' perspective and feelings. The developer said that being in a team also helps with empathising as a whole, recognizing that different individuals empathise in distinct ways. On the contrary, all the other developers stated that they were able to empathise with users [P1-P4] without doubting the need for expressing empathy. This suggests that some developers might not have a proper understanding of empathy or might not be aware of the need of empathy towards users or the impact of empathising.

🗨️ *"...I don't know if there was much to empathise with the customers about in this particular project. Because the only contact I really had with them was user testing. It was more just observing." - P5*

Developers demonstrated their awareness of using technical jargon when inquired about users' understanding of explained technical aspects. Developers acknowledged that they might have occasionally delved into too much technical detail, but they rephrased their explanations upon noticing user confusion. Upon reflecting on the use of technical jargon, developers said that they hadn't given it much thought at the time. However, developers stated that they could have provided better explanations to enhance user understanding [P1, P2]. When asked developers about being unresponsive to some user apologies, they explained that it happened because they were distracted due to some unexpected events [P2]. This suggests that sometimes developers were well aware of their actions and behaviours, and tried their best to accommodate the users. However, developers occasionally lacked self-awareness, i.e., how some of their actions impact the users and, developers realised it only when inquired during the interviews.

5.2 Enablers of Empathy

Both developers and users identified several enablers of their empathy. Enablers of developer empathy include: *Reflection in action, Commonality, Familiarity, Improved user understanding, Awareness of user actions and emotions, Appreciation, User confusion, Developer involvement, Technology literacy related issues, Gender similarities, Better connection, User interaction and Similar technical ability*. Enablers of user empathy include: *Mode of communication, Understanding of developer thought process, Understanding of developer needs, Developer familiarity and connection, Connection to the app's goal and appreciation of developer efforts*. We describe these enablers in detail in the following subsections, emphasising how they led to developer and user empathy.

5.2.1 Enablers of Developer Empathy There were different enablers of developer empathy towards users. Developers explained their perceptions of these different enablers of empathy when asked about how well they were able to empathise with users. We asked about their views on any special reasons for their empathy and how meeting roles impacted their ability to empathise. We also posed questions derived from our observations, such as why they felt more connected to some users over others. We mapped the demonstrated behaviours and expressed ideas to different empathy types such as cognitive and emotional empathy. We also mapped the empathy cues to the ideas expressed in interviews.

Developers stated that *commonality* is one of the enablers of empathy [P2, P4]. App feature reasoning was indicated by a developer as a key enabler of their empathy. Developers stated that talking about the app itself and providing suggestions regarding potential new features was one of the instances where they felt a commonality with the users [P2]. Another developer described that meeting users with a similar personality and same background (both were university students) were the sources of commonality which reinforced empathy [P4]. The developer further stated that the technical issues users faced were also another form of commonality as the developer had experience in solving these technical issues with their family members who are not very familiar with technology.

🗨️ *"I think I did empathise with them. I did understand some of the problems they have. And I understood where they came from with the point of view that they came from. Because from my understanding is that like, I guess my observation was that most of them that I did the interviews with, they were mostly not technology like knowledgeable I guess, as I am. So I could see it most of the time, because I have family members who are also not very technologically savvy and stuff like that. So yeah, I think I did a pretty good job of understanding what they were like feeling and like using, how to feel it and how they acted when they were using the app." - P4*

Gender similarities were also an enabler of empathy. A developer indicated that their gender might have assisted in understanding the thought process of the users and empathising with them.

🗨️ *"...I think that they were women probably contributed to me being able to empathise with them more, because I could maybe relate to what their thought process bit better" - Anonymous for participant privacy*

Observing user confusion was another enabler of empathy [P2, P4, P6]. A developer stated that users had difficulties in interpreting some questions. Users were occasionally confused by the terminology used in the app and the instructions provided during the usability sessions [P6]. The developer added that it was natural for first time users to feel confused about both app and usability testing process, as everything was completely new to them. Another developer admitted that the app would have been confusing to the users as they do not have access to all the information developers had while developing the app [P2] which reinforced the statement of P6. The developer further stated that it was easier for them to see these confusions from the users' perspective. Developers stated that the direct interaction with users helped them to understand the confusing features and the users' behaviour in such instances [P4]. Developers also stated that this helped them to understand users' perspective. Both P2 and P4 stated that user confusions assisted in better understanding the users' perspective which exhibits the practice of *perspective taking* which comes under the umbrella of *cognitive empathy*.

🗨️ *"...I was more easily able to see from like her perspective, it would be confusing, because she doesn't have all the information that I have, as I was making it." - P2*

Technical issues and having a *similar technical proficiency* were also triggers of developer empathy [P6]. Developers expressed their ability to empathise with users' technical difficulties,

drawing from their personal encounters with technical issues while using other apps. Developers stated that they were more empathetic towards technical issues encountered by technically competent users, as they believed they could personally face similar confusion if they were using the application. We could argue that this is another instance where developers demonstrated *cognitive empathy* by trying to put themselves in the shoes of the users.

Developers stated that different meeting roles helped them with *reflection in action*, another enabler of empathy [P2, P5]. A developer stated that it was easier for them to reflect and understand user feedback while taking the meeting minutes [P2]. Another developer stated that it was easier to reflect while hosting the session, as there was an opportunity to ask follow-up questions to clarify user responses [P5].

🗨️ *“... I think maybe in the session with [user], I empathised with [user] more. Maybe mostly because I was the one running it. So I was able to ask the questions that were coming into my head. Whereas the one where [Other developer] was hosting it, [Other developer] wasn't following up with questions that I was thinking of. So that was kind of a bit less of a connection for the two of us. And I think it was easy to understand what they were trying to say when I was hosting because I could pull up with different questions.” - P5*

Developers expressed their ability to understand user thoughts [P2], confusions [P4], and feelings when using certain features due to different conditions like ADHD [P3]. Developers believed that this *improved user understanding* helped them to empathise with users [P2-P4]. This may also be seen as an act of *cognitive empathy*. Developers found it easy to engage with the users. Developers emphasised that users excelled in both task performance and providing feedback. Developers stated that their *awareness of user actions and emotions* helped them to be more empathetic towards users [P3]. This may also be a way of displaying *emotional empathy*.

🗨️ *“when we're doing some of the things like check ins and [user] said that [user] had ADHD. And [user] was describing like, sometimes [user] is really focused, [user] wants to check in every day. And if [user] is not really focused, [user] wants to check in only once a week. I don't have ADHD, but I could imagine how [user] would feel like in those instances, where [user] would feel the urge to be really regular with it, and then sometimes not want to be at all.” - P3*

User appreciation was another trigger of developer empathy. Developers expressed appreciation for users, as they were nice, helpful, and sacrificed their time to support the developers [P3, P5]. Developers said that this sense of appreciation boosted their ability to understand and empathise with users. This also corroborates one of the verbal empathy cues we found in the literature which is praising the other person or showing understanding for a difficult situation of the other person (Verbal empathy cue 4). Developers said that hosting helped them to be more involved in the usability sessions and this *involvement* aided them to empathise with users [P3]. *User interaction* was another enabler of empathy. Developers highlighted that the form of user interaction in these sessions allowed them to focus on users' thoughts, needs, and feedback, which in turn, boosted developer empathy [P1, P6].

🗨️ *“being able to interact with people in that way, means that you're not focusing on yourself and trying to make sure that you're performing up to a task, like you're actually able to pay attention to what the end users want and the way that they're thinking about it...” - P6*

Familiarity is another enabler of developer empathy. Observed developer familiarity can be divided into two groups as *closeness to code* and *familiarity of user issues*. Developers found it easier to empathise with user feedback when they were familiar with a particular feature. This familiarity allowed them to ask insightful follow-up questions, improving their understanding of the user [P2]. Developers stated that their prior awareness of less user-friendly features allowed them to empathise more with users who encountered difficulties with these features [P4]. This

suggests that the developers' closeness to code enabled them to view things from users' perspective and be more empathetic towards them. Developers pointed out that they had personally experienced some of the user issues with their family members, and this familiarity increased their empathy towards users [P4].

☞ "... I was really easily able to identify with like the use of the software. Like whenever they had feedback, or thought things weren't working properly, I was really easily able to empathise with, like, yeah, they use the program, which I've been working on a lot. I suppose that's just because I was really familiar with it. Like I could see that the flaws in it, which they pointed out." - P2

Developers said that spending more time with users helped them build a *strong connection*, which improved their empathy towards users [P4]. Developers stated several reasons for establishing a good connection with users. They found it easier to connect with highly engaged users, and their first-ever usability session also fostered a strong connection [P1]. Developers also expressed that the ease of engaging with users [P1, P6] was another contributing factor. Some developers even noted that it was easier to engage with users than they had initially anticipated. Users also supported this idea by providing positive feedback. They found it easy to engage with developers and had no suggestions for changes in *user engagement*, even for future sessions [PU1, PU2, PU3]. This suggests that users were pleased with the experience they had while engaging with the developers. Developer involvement and familiarity were also significant factors in establishing a strong connection with users. Some developers felt a stronger connection when hosting sessions due their ability to build a rapport [P3,P4]. In contrast, others found a strong connection while taking minutes due to reflection in action [P1, P2]. In terms of familiarity, developers stated it would have been easier to understand and converse with more familiar users [P4, P5]. They also found it easier to connect with users who shared similar technological interests [P4, P6]. These similar interests served as a shared experience that strengthened their connection with users. This indicates that there were many instances where developers established a strong connection with users, and this connection further reinforced their empathy towards users. We also observed a positive correlation between empathy and connection.

☞ "... Yeah, I think it's just like, the longer time you spend with them [users], and the more of them that you as the interviewer, like, the "developer" was able to interact with the participant, I felt like the better it got, the more empathy I felt like I had with the like, the more closely like, I guess, you know, related, I guess it's the wrong word. Like, I felt like more like a connection with them even more than what I interacted with them." - P4

5.2.2 Enablers of User Empathy All users stated that they empathised with developers and understood their intentions, indicating their self-awareness of empathy towards developers [PU1-PU4]. They also described enablers that made them empathetic towards developers. In addition, users explained certain positive actions that might have contributed to their empathy towards developers. Ability to *communicate with developers face-to-face*, rather than through email [PU1] was an enabler of user empathy. Users stated that they *understood developers' thought processes* by interacting with the actual app, which was another enabler of user empathy [PU1]. Users stated that *understanding developers' needs* made them empathetic towards developers. Users described that they understood the information developers were seeking and the desired outcomes. This understanding allowed them to empathise with developers' requirements. Users stated that they provided the best possible feedback to developers. Users considered this as a way to support the developers in expanding the app's development and exploring new opportunities within their work [PU2, PU3]. Users were comfortable to give their blunt feedback since developers assured them not to be concerned about hurting their feelings [PU2]. *Empathy due to*

understanding of developers' thought processes and needs, demonstrates perspective taking, which is encompassed under the umbrella of cognitive empathy.

Users empathised with developers due to their *connection to the app's goal*. They found it exciting to be a part of a project that transformed an idea into a useful app particularly for females like them. Users believed the app would have a positive impact on their lives and those of other PCOS patients they knew. Users believed developers would also have empathy towards PCOS patients like them as they built this app to support such patients. Users stated this realisation helped with their empathy [PU3].

🗨️ *"I could empathise with them because I have PCOS and I could see that it's an app that would be helpful to myself and the people that I know. And so for that reason, I felt like there was a personal connection to the app. And I could understand that if they were trying to develop an app that would help people then they themselves would have empathy or an understanding that this has an impact on women. And so for that reason, I felt like I could empathise with that."* - PU3

Users were empathetic due to *appreciation of developer effort*. Users understood the developers and recognised their dedication to create an app that would be beneficial for them [PU2]. Users recognised developers' persistent passion through their commitment to the app [PU3]. Users also understood developers sought user feedback to improve the app and understood developers' intentions [PU4]. Users were appreciative towards developers due to these reasons and it also became an enabler of their empathy.

🗨️ *"I was able to sort of empathise with them, because I know that they're working hard to make, like a program that would help someone like me. So that in itself was just like, okay, I get where you're coming from. And I want this too so I'm kind of like, yeah, just understood them for just that."* - PU2

Users believed having usability sessions with the same developers would strengthen their empathy and connection with the developers due to *familiarity and established connection* [PU3, PU4].

🗨️ *"I feel previously connected to them and having liaison with them, I thought, or I would think if it's the same type of research, or, you know, maybe because they're the same type of people in that environment, they know what's going on. Maybe it might be the same. I think I'm just comfortable with doing it..."* - PU4

Users explained positive behaviours that may have enabled their empathy towards developers. Users were excited to join the usability sessions, demonstrating a clear understanding of the developers' needs and providing honest feedback [PU2]. Users found the sessions enjoyable and relaxed. The casual conversation with developers at the start helped users to be more relaxed. Users also appreciated the developers' patience and respectful engagement. They felt connected with the developers during all the sessions, allowing them to comfortably express their thoughts, as the developers were receptive to their feedback. The developers' clear communication helped users to engage with them successfully [PU3]. These user responses regarding their engagement with developers signify a positive experience for users during these interactions. This suggests that users had a positive engagement with developers, which could have fostered empathy towards the developers.

🗨️ *"Yeah, I think they were extremely patient with me. I'm not very good with technology. And so for some of the different aspects of the app that they were showing me they were very patient and could communicate with me in a way that I understood. So they showed me that kind of sensitivity around that, which was good."* - PU3

Users provided *positive feedback about usability sessions and the app*. With regards to usability sessions, users appreciated the developers' flexibility in scheduling [PU2]. Users suggested having clearer instructions on the type of device to use in the sessions would be nice, but otherwise,

they were content with the usability sessions [PU4]. With regards to the app, users enjoyed speaking out loud while using the app. They believed the app will be highly useful upon completion [PU1]. Users were also pleased with their ability to provide feedback and found the practical side of using the app enjoyable [PU2, PU3]. Users were excited to see the app for the first time and found it to be user friendly [PU3]. Users were able to reflect on their first session and provide more authentic feedback in the second session [PU3]. Users even suggested some new features they would like to have in the app [PU2].

🗨️ *“I think it’s gonna work. Like once the app is like, complete and stuff, I think it’s gonna work really good. That’s like going through like the first part of the study. And this part was good to get an insight as well into what goes on behind making that sort of thing.” - PU1*

User feedback about the app and usability sessions suggests that users enjoyed the experience they had while using the app. This indicates that there is a possibility that these users will use the app once it is released which could be an indication of the usability of the app. All these responses show the positive actions of users towards the interaction with the developers that may have reinforced their empathy.

5.3 Barriers to Empathy

Both developers and users explained barriers that they perceived to their empathy with the other. Key barriers to developer empathy include: *Less developer interest, Task centredness, Unfamiliarity of features, Limited time for reflection in action, Difficulty in understanding user struggles, Difficulty in empathising with low-technology literacy related user concerns, Less interaction and Difficulty to resonate with collective experiences of women.* Major barriers to user empathy include: *Poor user connection with developers and their Negative emotional responses.*

5.3.1 Barriers to Developer Empathy Developers perceived *less interaction with users* as a barrier to their empathy. During the minute taking (MT) role, they couldn’t ask follow-up questions, which hindered their connection with users as the hosting developer did not address these questions either [P5]. Developers could not build a rapport with the users during MT and it made understanding users difficult. Developers felt like mere observers during MT and believed only hosting developers should interact with users because of the session format [P4]. According to developers, the fast pace of some sessions may have hindered their empathy [P1].

🗨️ *“...Whereas the one where [other developer] was hosting it, [other developer] wasn’t following up with questions that I was thinking of. So that was kind of a bit less of a connection for the two of us...” - P5*

Unfamiliarity of users and hosting sessions, and limited interaction due to MT were the major reasons for poor developer-user connection [P2-P4]. Sometimes developers were unfamiliar with the information shared by users, which made it difficult to continue the conversation. Also hosting felt awkward initially as it seemed more like giving instructions than a real conversation. Developers found the MT role to be more of a background role, limiting their ability to interact with the users. These factors made them less connected to the users.

Some developers stated that if they were less involved and less interested they would not empathise much with the users. This indicates that *reduced developer interest can lead to less empathy towards users* [P1]. Developers explained their *task centredness* made them less focused on users [P1, P4].

🗨️ *“...I think I was more focused on trying to present the tasks as clearly as possible...” - P1*

Developers empathised more with user feedback when they were familiar with the software features, which demonstrates empathy due to closeness to code. However, developers were *unable to empathise with user feedback much when the features were unfamiliar* to them. They

elaborated that when a feature was familiar to them, they had already considered various ways to implement it. But if it was unfamiliar, they hadn't explored these options, which made them less empathetic towards user suggestions [P2].

🗨️ *".. I was able to empathise with their feedback a lot more, than on features where I wasn't as familiar because I wasn't, I hadn't already thought through, like, all the different options of how we could have done it for example, whereas on stuff that I was really familiar with, I may have already like thought of that." - P2*

Developers stated that *limited time for reflection in action* was a barrier to their empathy. They did not have enough time to think about user feedback while hosting, hence they were unable to empathise with those suggestions [P2].

🗨️ *"..Whereas when I was hosting it, it was very like, you know, you just talk and you don't really have time to stop and like think about that a specific feedback." - P2*

Reflection on action was also a barrier to developer empathy. Sometimes developers could not properly understand what users meant when they reflect on user feedback [P2]. Developers thought that asking follow-up questions would have been helpful in these instances. Developers stated they had a *difficulty in understanding some of the user suggestions* and they struggled to empathise with such suggestions. Developers stated that it was hard for them to understand what the users were envisioning and they could not get on the same page with users for all of their suggestions [P2].

🗨️ *"Well, I think like there are other suggestions where I didn't empathise as well. So when user wanted to have like a tick and flick, or one of the features and, like, it was, I mean, it's hard to say like, I struggled to empathise, because I just didn't really understand." - P2*

Developers had *difficulties in understanding user struggles* because they had not personally experienced these issues. This lack of personal experience hindered their empathy. Developers found it challenging to empathise with user issues that were unique to less tech-savvy users [P6]. Since developers were well-versed in technology and familiar with the app, they couldn't relate to these particular *low-technology literacy related user concerns* [P5, P6]. Developers had a *difficulty in finding the balance of technical expression*. They believed that excessive explanations could confuse users and make it hard to obtain relevant user feedback. However, developers felt that they found a balance and didn't delve into too much technical details [P6].

🗨️ *".. I felt that like I didn't really experience the problems they were having per se, like, we kind of like show them what we have. And they were like, yeah, it looks good, good job. And like, obviously, they had maybe some small issues using or whatever, but there wasn't anything significant. So it was hard to you know, really empathise with people who didn't really feel like we're struggling. I didn't feel the need to empathise with them..." - P6*

Developers stated that *difficulty to resonate with collective experiences of women* was another barrier for their empathy. However, some developers argued that they would be able to empathise the same way even without having the same gender.

🗨️ *"I think in this project, one thing that made it a little bit more difficult is just gender. Like, where end users are exclusively women. As a man, it's like, already, at least one step out of being able to empathise fully. It doesn't change, like how much I'm trying to, but I think it changes the reality of like, there are collective experiences that women have that I just don't know, haven't experienced personally. And from that perspective, it definitely makes it more difficult to fully empathise in the way that maybe you will be able to have a stronger connection with..." - Anonymous for participant privacy*

5.3.2 *Barriers to User Empathy Poor user connection with developers, and users' negative emotional responses*, were the two major barriers to user empathy. Users had a *poor connection* with

the developers who played the less interactive role of MT. Even though users understood the role of these developers, they perceived a lack of connection with them [PU4].

“I have nothing against [minute-taking developer] or anything. It’s just, you know, I didn’t have much interaction with [minute-taking developer] at all. That’s fine. Oh, definitely, because that I feel like well, previously, I did the other session at first, and then I was invited back. So I, for example, you know, I was aware of [this developer] previously, but again, I still hadn’t had much to do with [this developer].” - PU4

Users’ *negative emotional responses*, such as nervousness, confusion, and discomfort, were the other major empathy barriers. Users described various reasons for their nervousness during usability sessions. They stated feeling rushed due to *confusion* about the meeting time and work commitments, which added to their nervousness. Additionally, their *unfamiliarity* with what to expect during the sessions contributed to their nervousness. Users also expressed a *strong desire to provide high quality feedback* to developers, which increased their nervousness as they didn’t want to disappoint developers. Lastly, the *anticipation of the questions* that might be asked during the sessions heightened their nervousness. Users were also a bit *confused with terminology* used in the app and it made them *uncomfortable*. Users often felt they weren’t using the app correctly and couldn’t provide good feedback [PU3]. This was also noticed by the developers. They felt bad seeing users confused and uncomfortable with the app [P5].

“However, there was a point where I was little confused by the wording of something, so I felt then a little bit awkward, like I was not getting it right or not giving feedback...” - PU3

5.4 Strategies to Improve Empathy

We describe the approaches developers used to support the users including *technical support* and *emotional support*. We also discuss prospective strategies they suggested when asked a hypothetical question on what would they do to better empathise with users in future. Prospective developer strategies include: *Better preparation*, *Reduce the pace of the sessions*, *Ask more follow-up questions*, *Support when users are struggling*, *Taking time to implement the changes suggested by users*, and *More agile way of interacting with end users*.

5.4.1 Developer Support All developers explained different approaches for supporting users. They offered both technical and emotional support to users during these sessions.

Technical Support. During the usability sessions developers modified the questions based on the scenarios. They *personalised the content*, rephrased and clarified questions, especially when users seemed confused [P1]. They stated that it was an attempt to support users to the best of their ability [P1].

“I was able to try, I guess, modify the questions on the spot as, as the scenario changes, like throughout the meeting, to some extent. So I guess that maybe shows a bit of consideration towards each individual.” - P1

Developers implicitly *tried to talk in layman’s terms* when dealing with the users [P3]. Developers were confident that the users sufficiently understood the technical details, which was evident in the way users performed tasks during the usability sessions. Users also confirmed this, stating that they were able to clearly understand and follow developer instructions. In addition, developers explicitly *tried to phrase questions as clearly as possible at the first time* [P1, P6]. This demonstrates that the developers tried to support the users by giving the best possible instructions.

“We didn’t really try to explain too much technical detail, at least in my sessions, we didn’t really try to explain the technical details behind it unless it was like, strictly relevant. Often, it

was more like, we want to do this, but we haven't, because we haven't had the time to implement it yet. Like that was kind of the most in depth..." - P6

Emotional Support. Developers described their *urge to help users when they were struggling* during the usability sessions. However, they could not directly help the users as the goal of the sessions was to elicit unadulterated user feedback [P1]. Developers faced a dilemma of whether to empathise with users by helping them resolve their confusions, or allowing them struggle and find their own way. Developers were uncertain whether they found a balance in this dilemma. Despite feeling bad for confused users, developers often refrained from offering assistance. They believed they should let users resolve issues on their own to achieve the session's goal [P1, P5]. Based on some of the user responses, it was evident that occasionally developers have failed to find the balance in this situation [PU2, PU3]. Users said that they had requested developer assistance to clarify their confusions in such situations but it was not forthcoming.

🗨️ *"So I guess that's a scenario where I would empathise more with them when they're having struggles because I want to try to help them even and that's the is the point of the session to let them struggle to some extent so that we can get information about Yeah..." - P1*

When asked about a time where developers empathised with users, developers stated that they *paid more attention to user concerns* when users brought up their own confusions [P1]. This suggests that supporting and wanting to support users is one of the things that comes to their mind when thinking of empathy.

🗨️ *"Yes, I suppose like, as I said, given when the users gave feedback, so one of the users gave feedback on, like the wording was a bit confusing. And like when she pointed it out, it was really easy for me to see that then. Like once you pay it once it was brought to it, and I paid more attention..." - P2*

Developers wanted to *reassure users that they didn't need to strive for perfection* during usability sessions. They used phrases like "All good" to reassure users when they were worried about making mistakes, such as entering incorrect spellings or struggling to navigate the app. Some developers shared that they too had similar concerns, in order to reassure users in such situations [P2-P5].

🗨️ *"Yeah, because obviously, in user testing, I don't want to say too much. But I also wanted to reassure [user] that there wasn't a way [user] could do it wrong..." - P5*

Developers *reinforced the successful acts of users* to provide more support and they demonstrated it using the terms like "Perfect" [P6]. This suggests that empowering users by reassuring and reinforcing their successful acts is one of the innate empathetic behaviours of developers when directly interacting with users. This corroborates one of the verbal empathy cues which is affirming or showing understanding for the other person (Verbal empathy cue 4).

Developers described that their use of certain behaviours and words during usability sessions was driven by the intention to *comfort users and connect* with them. They used phrases like "Perfect" or "All good" when users deviated from expected task performance and occasionally used humor too [P3-P5]. This indicates that developers naturally sought to create a rapport with users. They might have done this as a way of prompting for an effective user evaluation.

🗨️ *"I think that's my main concern with that was like, trying to comfort her, so that you know, like don't think that you're doing anything wrong, there's nothing wrong with that..." - P4*

5.4.2 Prospective Acts of Empathy Answering a hypothetical question on future acts of empathy, developers expressed they would work on *phrasing their remarks in a better way* [P2] and *spend additional time to prepare* and understand the tasks. Developers also acknowledged their nervousness during hosting the sessions and indicated their desire to prepare more to better engage users during these sessions for more productive outcomes [P1]. Additionally, they emphasised

the importance of asking *follow-up questions*, particularly concerning user feedback, to gain insights into user perspectives and thought processes [P1, P2]. This corroborates one of the verbal empathy cues which is asking a substantive question about the statement of the other person to better understand what the other person is saying (Verbal empathy cue 3).

🗨️ *“...and like maybe ask follow-up questions. Especially like if a user has feedback like ask more questions about that, rather than just saying like, okay let’s move on...” - P2*

Developers preferred a *more agile approach to interact with users* as they wanted to know the real needs of users. They believed that this approach, which involves developing a working software prototype for usability sessions, would allow having consistent communication with users and the discovering their real needs. Developers highlighted the benefits of obtaining user feedback earlier in the development cycle and the potential for continuous feedback and improvement through this agile approach [P6].

🗨️ *“All we do is we’d like to work much ourselves for a while for maybe like 7,8,9,10 weeks. And then we talk to the end users for like, two weeks, and then we’d go back and do your own thing. And then we talk to the end users again...but like having a more constant interaction with end users would make a more it would have affected both the developers and users. because we would have, as a developer, we would have had more constant need to discover what the end users want. And more constant need to develop prototypes and things that they can use, to show off in their sessions. And also, like, more room constant revisions...” - P6*

Developers also wanted to *slow down the sessions* [P2] and *take time to implement the changes suggested by users* [P6]. They preferred modifying the app incorporating user feedback which would allow users to see the implementation of their vision. Developers believed it would have been more beneficial to show the modified app to users during usability sessions as they are the ones who are really going to use it [P6].

🗨️ *“we just wanted to make what we thought would be the best. And then listen to the end users and be like, that’s great. We love your input. We’ll change this thing. But like, we’re still gonna go with what we want... If we had taken a step and being like, Alright, now we know, we’re gonna try to spend some time implementing the vision that the these end users had, because ultimately, they’re the ones who are going to be using it, not us.” - P6*

Developers expressed their intention to *provide better support to users when they are struggling* or confused. They acknowledged the need for careful preparation to manage unexpected user actions or difficulties, a consideration they had not thoroughly planned for previously [P5]. These strategies represent their commitment to improve empathy with end users in future sessions. We posed the same hypothetical question to the users. The majority of users (three out of four) said they would not change anything as *“everything was quite good and enjoyable”* [PU1, PU2, PU4]. One of the users stated that they would be more confident to provide feedback in future as they have experienced it once [PU3].

5.5 Observed Empathy Cues (Observation Findings)

We identified the demonstration of several verbal and nonverbal empathy cues which are explained in Tables 3 and 4. Below we discuss the observed empathy cues, focusing on how developers and users exhibited these cues and the number of occurrences of the cues.

5.5.1 Verbal Empathy Cues We observed all four verbal cues (Table 3) from both developers and users. However, participants demonstrated these verbal cues differently, as shown in Table 5. When considering developers, we noticed that the majority of verbal cues were observed from the hosting developers. Among them, the most dominant verbal empathy cues were third (empathic conjectures: follow-up questions) and fourth cues (empathic evocations: praising other

Table 5. Demonstration of Verbal Empathy Cues

| Cue | Developer Demonstration | User Demonstration |
|----------------------------------|--|--|
| Empathic Understanding Responses | <i>H*</i> & <i>MT*</i> : Answering user questions by repeating & referring to their previous content. | Answering developers' questions by referring to their statements. |
| Empathic Affirmations | <i>H</i> : Summarising users' actions & suggestions while responding to them. Repeating users' statements. <i>MT</i> : Summarising users' suggestions & struggles. | Summarising and repeating developers' statements. Summarising developer content while explaining user preferences. |
| Empathic Conjectures | <i>H</i> & <i>MT</i> : Asking follow-up questions from users to understand their responses & suggestions. | Asking follow-up questions from developers to understand the instructions. |
| Empathic Evocations | <i>H</i> & <i>MT</i> : Praising users and showing understanding when users faced difficulties. | Praising developer effort and showing understanding. |

* *H*: Hosting, *MT*: Minute Taking

person). By demonstrating third verbal cue, hosting developers asked multiple follow-up questions to understand responses of users and suggestions provided by users. We noticed only a few follow-up questions from MT developers. Hosting developers exhibited the fourth empathy cue by praising users and showing understanding when users faced difficulties during usability testing sessions. The first cue was the least observed verbal empathy cue among hosting developers. Hosting developers displayed the first cue by repeating the content of users' statements while answering their questions. We hardly noticed verbal empathy cues from MT developers. We observed fourth verbal cue six times among them. When considering users the most observed verbal empathy cues were first and third cues. The least observed verbal empathy cue among users was the second cue where users summarised and repeated developers' statements. Occurrences of verbal empathy cues of developers and users are illustrated in Table 6 and Table 8.

5.5.2 Nonverbal Empathy Cues All nonverbal empathy cues (Table 4) were demonstrated by both developers and users during usability sessions. There were varied behaviours linked to each cue, as shown in Table 10. The second nonverbal cue was the most observed cue among both hosting and MT developers. This cue was demonstrated by developers via facial expressions for attentive listening such as nodding, raised eyebrows, thumb & forefinger of a hand on the chin and moving face closer to screen [65]. Third and first nonverbal cues were the least observed cues among hosting developers MT developers, respectively. Developers exhibited third and first nonverbal empathy cues respectively through their facial expressions similar to users and by verbally acknowledging user's speech via different sounds and short, monosyllabic words (see Table 10). When considering users, first nonverbal cue was the most observed cue where as second cue was the least observed. Occurrences of nonverbal empathy cues of developers and users are illustrated in Table 7 and Table 9.

5.6 Empathy Test Results

We calculated the empathy test scores of both developers and users in each round of development. In addition we calculated quartiles, interquartile range (IQR), maximum and minimum values of empathy test scores. These data were then used to draw box plots to demonstrate the overall dispersion of empathy scores of developers and users. We visualised empathy scores

Table 6. Occurrences of Verbal Empathy Cues of Devs

| Devs* | V* Cue 1 # | | V Cue 2 # | | V Cue 3 # | | V Cue 4 # | |
|--------------|------------|----------|-----------|----------|-----------|----------|-----------|----------|
| | H* | MT* | H | MT | H | MT | H | MT |
| P1 | 1 | 1 | 1 | 2 | 10 | 1 | 10 | 5 |
| P2 | 5 | 0 | 9 | 0 | 11 | 0 | 9 | 0 |
| P3 | 2 | 0 | 3 | 0 | 7 | 0 | 7 | 0 |
| P4 | 5 | 0 | 3 | 0 | 3 | 0 | 7 | 0 |
| P5 | 3 | 0 | 5 | 0 | 12 | 0 | 8 | 0 |
| P6 | 2 | 0 | 2 | 0 | 1 | 0 | 7 | 1 |
| Total | 18 | 1 | 23 | 2 | 44 | 1 | 48 | 6 |

* Devs: Developers, V: Verbal, NV: Nonverbal, H: Hosting, MT: Minute Taking

Table 7. Occurrences of Nonverbal Empathy Cues of Devs

| Devs | NV* Cue 1 # | | NV Cue 2 # | | NV Cue 3 # | |
|--------------|-------------|-----------|------------|------------|------------|-----------|
| | H | MT | H | MT | H | MT |
| P1 | 23 | 9 | 63 | 98 | 8 | 34 |
| P2 | 66 | 0 | 57 | 1 | 29 | 4 |
| P3 | 22 | 0 | 40 | 21 | 11 | 9 |
| P4 | 15 | 0 | 26 | 2 | 7 | 2 |
| P5 | 13 | 1 | 14 | 12 | 3 | 4 |
| P6 | 12 | 0 | 28 | 4 | 3 | 2 |
| Total | 151 | 10 | 228 | 138 | 61 | 55 |

* Devs: Developers, V: Verbal, NV: Nonverbal, H: Hosting, MT: Minute Taking

Table 8. Occurrences of Verbal Empathy Cues of Users

| Users | V Cue 1 Count | V Cue 2 Count | V Cue 3 Count | V Cue 4 Count |
|--------------|---------------|---------------|---------------|---------------|
| PU1 | 5 | 2 | 7 | 0 |
| PU2 | 4 | 3 | 4 | 6 |
| PU3 | 3 | 3 | 7 | 5 |
| PU4 | 7 | 2 | 3 | 4 |
| Total | 19 | 10 | 21 | 15 |

* V: Verbal

Table 9. Occurrences of Nonverbal Empathy Cues of Users

| Users | NV Cue 1 Count | NV Cue 2 Count | NV Cue 3 Count |
|--------------|----------------|----------------|----------------|
| PU1 | 29 | 13 | 13 |
| PU2 | 47 | 21 | 11 |
| PU3 | 11 | 21 | 28 |
| PU4 | 29 | 10 | 21 |
| Total | 116 | 65 | 73 |

* NV: Nonverbal

using several box plots. In this section we have only included the box plots drawn for total empathy, total cognitive empathy and total affective empathy scores of developers and users in each round. We explain the implications of these graphs based on the length and position of the box within and also across participant groups.

When considering the graphs that illustrate *total empathy*, we noticed that developer box plots are short compared to user box plots (see Figure 4) in both baseline and round one. This implies that overall developers have similar empathy scores whereas user empathy scores are quite diverse (Dev IQR < User IQR). It is clear that both developer and user box plots are shorter in round one compared to their baseline box plots. This implies that both developer and user total

Table 10. Demonstration of Nonverbal Empathy Cues

| Nonverbal Cue | Developer Demonstration | User Demonstration |
|---|--|--|
| Lute for careful listening | <i>H & MT</i> *: Verbal acknowledgement via sounds like “Mmm”, “Ah” & short, monosyllabic words such as “Ok”, “Cool”, “Yeah”. | Verbal acknowledgement via sounds like “Mmm”, “Ah” & short, monosyllabic words such as “Ok”, “Yeah”. |
| Nonverbal signals for attentive listening | <i>H & MT</i> : Facial expressions for attentive listening such as nods, raised eyebrows, thumb & forefinger of one hand on the chin. Moving face closer to screen. | Facial expressions for attentive listening such as nods & raised eyebrows. |
| Similar facial expressions, posture, gestures | <i>H & MT</i> : Facial expressions similar to users such as laughing/smiling [Joy]. | Facial expressions similar to developers such as laughing/smiling [Joy]. |

* *H*: *Hosting*, *MT*: *Minute Taking*

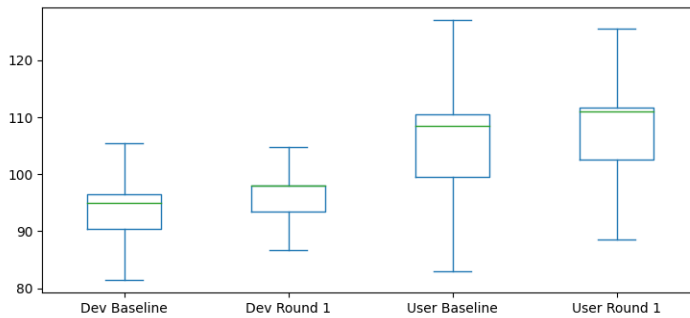


Fig. 4. Total Empathy Distribution (Removed two outliers)

scores are less dispersed in round one compared to baseline empathy scores. The user box plots are much higher than developer box plots in both baseline and round one. This suggests that overall users have higher total empathy score compared to developers. The developer round one (Dev Round 1) box plot is much higher than baseline (Dev Baseline) box plot. This same behaviour can be seen with respect to the box plots of user baseline and round one empathy. This suggests that overall developer and user empathy has increased in round one compared to their baseline empathy.

We noticed two outlier scores for a developer and a user in both baseline and round one empathy who scored below the lower whisker⁶ limit. During the interview, this developer was uncertain whether there’s any need for empathising with users in this project. It was evident from this developer’s statement “*I don’t really think, I don’t know if there was much to empathise with the customers in this particular project*”. However, we do not have a deeper understanding regarding the user with the outlier score. Based on our experience in interviewing this user, we assume this user may not have understood the scale items properly as these scale items were very generic.

When considering the *total cognitive empathy graphs* (Figure 5), it is clear that the user box plots are slightly shorter compared to the developer box plots (User IQR < Dev. IQR) in both baseline

⁶The whiskers are the two lines outside the box, that go from the minimum to the first quartile (lower whisker) and then from the third quartile to the maximum (upper whisker).

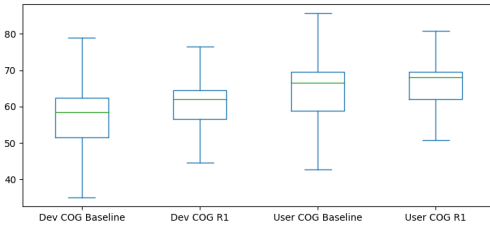


Fig. 5. Total Cognitive Empathy Distribution (Removed an outlier)

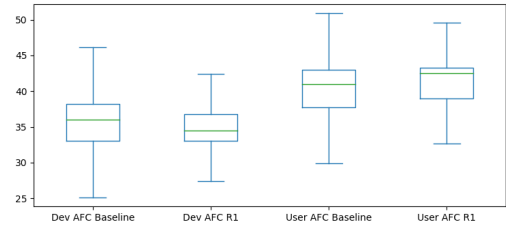


Fig. 6. Total Affective Empathy Distribution (Removed an outlier)

and round one. This implies that cognitive empathy scores of developers are more dispersed compared to users' cognitive scores. Both developer and user box plots are shorter in round one compared to their baseline box plots. This implies that both developer and user cognitive scores are less dispersed in round one compared to baseline empathy scores. User box plots are much higher than developer box plots in both baseline and round one. This suggests that overall users have higher cognitive empathy score compared to developers. The developer round one box plot is comparatively higher than developer baseline box plot. This infers that overall developer cognitive empathy scores are higher in round one compared to baseline scores. We noticed cognitive empathy box plot of users in round one is slightly higher than their baseline box plot. This implies that overall user cognitive empathy scores have increased in round one compared to their baseline scores.

We noticed an outlier score for a user in round one cognitive empathy who scored below the lower whisker limit. However, we noticed that user's cognitive score in round one has increased compared to baseline score. Despite this situation, user's score was identified as an outlier. This suggests that even though user's cognitive empathy score increased in round one, it has not increased much compared to other users.

When considering the *total affective empathy graphs* (Figure 6), both developer and user box plots are equal in length in baseline (Dev. IQR = User IQR). Despite this, the developer box plot became slightly shorter compared to user box plot in round one (Dev. IQR < User IQR). This illustrates that affective empathy scores of users are much dispersed compared to developers' in round one. Both developer and user box plots in round one are shorter compared to their baseline box plots. This implies that both developer and user affective scores are less dispersed in round one compared to baseline empathy scores. User box plots are much higher than developer box plots in both baseline and round one. This implies that overall users have higher affective empathy score compared to developers. We noticed that developer box plot in round one is comparatively lower than baseline empathy box plot. Even though the value of first quartile (Q1) is same and minimum is increased, all the other values are decreased (Q2, Q3, maximum) in round one compared baseline. This implies that overall affective empathy of developers has decreased in round one compared to baseline. However, we noticed that users' round one affective empathy box plot is higher than baseline box plot. This implies that users' overall affective empathy has increased in round one compared to baseline. Despite this trend we noticed an outlier score for a user in round one affective empathy who scored below the lower whisker limit. We identified that this user's affective empathy has reduced even when considering the numerical scores. However, we are not clear what could be the reason for this behaviour.

6 Discussion

6.1 Collating Mixed Methods Data

Qualitative measures, including interviews and observations, were instrumental in uncovering empathy enablers, barriers, and strategies to overcome barriers. These qualitative methods also revealed the manifestation of empathy cues during direct interactions between developers and users, offering insights into how empathy is practised based on their experiences and interactions. Complementing these qualitative approaches, we employed quantitative measures, specifically empathy test scores determined by the QCAE empathy scale. These scores offered a numerical assessment of participants' empathetic tendencies, providing a comprehensive understanding of empathy in our study.

Our demographics questionnaire helped us to understand our participants better and it helped us to be more open during our observations. We observed empathy cues and behaviours of participants during usability sessions. These observations informed the interviews and we were able to inquire about different behaviours of participants. Empathy test scores provided a quantitative measure to understand participants' empathy. This measure also informed our interviews by helping us to compare & contrast participants' scores against demonstrated empathy cues and explanations provided during interviews. Hence our observations and empathy test scores drove our interviews in a better direction by helping us to form better questions and capture rich data.

We found several correlations between empathy cues, interview findings and empathy scores. We were able to map third and fourth verbal empathy cues (Table 3) with the enablers and strategies identified in interview findings. The fourth verbal cue is associated with praising or showing understanding for the other person. It is rooted in the codes "*Empathy due to appreciation*" and "*Empathy towards user confusion*". These codes belong to the enablers of empathy and usage of mixed methods assisted in unfolding the connection between empathy cues & enablers of empathy. We also identified that the fourth verbal cue is embedded in the developer strategies such as "*Reassuring users*" and "*Reinforcing the successful acts of users*". Developers used these strategies to emotionally support users. The third verbal cue which is associated with substantive questions to better understand the other person is rooted in the codes "*Developer plans to prepare more*" and "*Developer plans to ask more follow-up questions*". These codes belong to the prospective developer acts of empathy which comes under the developer strategies.

We identified correlations between empathy cues and empathy scores. While individual empathy scores for developers and users were not analysed due to insufficient data, we opted to assess the scores across two groups: developers and users. This approach allowed us to measure changes in empathy within each group over time, providing valuable insights into the general empathetic dynamics between developers and users throughout the study period. When considering developers, we noticed that the highest number of empathy cues were reported from the hosting developers. We suspect this was due to the less involvement of MT developers. Most of the developers stated that they thought they were not supposed to talk with users while taking minutes (P3, P4, P5). We rarely noticed any communication between MT developers and users while observing the usability sessions. Hence verbal empathy cues were by default low in MT role. This could be a reason for low number of overall empathy cues reported from developers while performing MT role compared to hosting. However, the highest number of nonverbal cues were also observed from hosting developers. Several developers stated that they had a minimal interaction with users while performing MT role and it was a barrier to their empathy (P3, P4, P5). This could be a reason for low number of nonverbal empathy cues. We noticed some patterns between overall empathy cue counts and empathy test scores of participants in round one

& round two. We identified three main patterns: high empathy cue count when empathy score is high; low empathy cue count when empathy score is high; and having the same empathy scores. We had three instances where participants had the same or almost the same empathy scores in all rounds. These participants were developers and in all these instances we noticed that the empathy cue count is high when hosting the session. We also identified four participants who had high empathy cue count when total empathy score is high. Two of these participants were developers and they both demonstrated high empathy cue count when hosting. We identified two participants who had low empathy cue count when empathy is high. One of them was a developer and this developer demonstrated this high empathy score while performing MT role. We also had one user who participated only in the first round which limits our ability to compare this user's empathy score with empathy cue count. Most of the time we noticed developers had a high empathy cue count when they were hosting (five out of six). Sometimes their empathy cue counts did not reflect their empathy score. However, their cue counts were more representative of the role they played in the usability testing sessions. This could be due to the nature of usability sessions.

6.2 Cognitive & Affective Empathy

Most of the empathy enablers that we identified in this study can be categorised as cognitive and affective. Cognitive empathy enablers triggered cognitive empathy. Likewise affective empathy enablers triggered affective empathy. Enablers such as *reflection in action, familiarity, improved user understanding, user confusions, user interaction, similar levels of technical ability, understanding of developer thought process and understanding of developer needs* were identified as cognitive empathy enablers. These helped participants to detect and understand others' perspectives. For instance, familiarity was a cognitive empathy enabler for developers. In this context developer familiarity refers to closeness to code and familiarity of user issues. This familiarity helped developers to better understand user suggestions and user issues, and is a demonstration of enabling cognitive empathy.

Enablers such as *awareness of user actions and emotions, appreciation, technology literacy related issues, gender, familiarity of developers & connection, connection to the app's goal and appreciation of developer efforts* were identified as affective empathy enablers. These enablers triggered participants to react to others' emotions and share an emotional experience. For instance, familiarity of developers & connection was an affective empathy enabler for users. This familiarity & connection made users feel emotionally close to developers, and users were able to establish a rapport with developers, demonstrating enabling of affective empathy.

6.3 Empathy Enablers and Barriers

As illustrated in Figure 3, our STGT data analysis uncovered various developer awareness types (see Table 11), empathy enablers (see Table 12), barriers to developer and user empathy (see barriers in Table 15 and Table 16), and strategies for overcoming these barriers (see Table 14). These identified enablers and barriers align with established empathy literature, particularly drawing parallels with enablers such as familiarity and commonality [6, 24, 40, 43, 56, 60, 68, 72, 76–78] and barriers including time pressure, disinterest, and task-centeredness [48, 69]. We identified that developer awareness of their own emotions & actions, as well as emotional & technical awareness of users, is a first step of triggering empathy. This awareness may assist in enabling empathy. Goleman also endorsed the notion that empathy builds on self-awareness [31]. If empathy is not triggered, then this awareness most likely leads developers to identify barriers to empathy. When developers identify these empathy barriers, they may apply some strategies to overcome these barriers.

Developer awareness reinforced enablers of empathy and developers used some strategies to overcome their empathy barriers. For instance, developers had the technical awareness of the users' understanding. This improved understanding enabled developers' empathy towards users' technology literacy related issues (see Table 11, 13). However, some developers identified the difficulty in empathising with low-technology literacy related user concerns as a barrier to their empathy. To overcome this empathy barrier, developers are planning to follow a set of strategies in future including ask more follow-up questions, pay more attention to user concerns, and connect more with users (see Table 14, 15).

Table 13 provides examples of required developer awareness and empathy enablers while Table 15 presents developers' empathy barriers and strategies to overcome these barriers. We established these connections based on the analysis of data obtained through the interviews. Participants were asked about their experiences with empathy enablers, empathy barriers, the actions they took to overcome these challenges, and their proposed strategies for addressing such barriers in the future. Our STGT data analysis of these responses allowed us to pinpoint specific enablers for each type of awareness, and specific barriers and their associated strategies. We present these details exclusively for developers as our data is not sufficient to compile a similar guide for users. We propose some potential strategies developers can follow that may minimise barriers to users' empathy in Table 16. For instance, users highlighted confusion on terminology as a barrier to empathy. When developers observe users are confused about certain terminology, we propose developers to employ strategies such as rephrasing & clarifying, using simple language and avoiding overloading users with technical details.

Table 11. Types of Developer Awareness

| ID | Type | Nature of Awareness |
|----|---------------------|--|
| A1 | Self Awareness | Developer awareness of technical jargon |
| | | Developer awareness of differences in empathy |
| A2 | Emotional Awareness | Users being stressed due to recording |
| | | Users feeling pressured |
| | | Developers noticing user reactions |
| | | Users might have felt developers sounded judgmental |
| | | Developer awareness of user actions and emotions |
| | | Developer understanding on user concerns |
| A3 | Technical Awareness | Improved developer understanding on user needs |
| | | Developer awareness of the importance of empathy towards users |
| | | Developers' awareness of user's technical proficiency |
| | | Developer awareness of user understanding |

Table 12. Enablers of Developer Empathy

| ID | Enablers of Developer Empathy |
|-----|---|
| E1 | Empathy towards user confusion |
| E2 | Empathy due to commonality |
| E3 | Empathy due to familiarity |
| E4 | Empathy due to the awareness of user actions and emotions |
| E5 | Empathy due to appreciation |
| E6 | Empathy due to more developer involvement |
| E7 | Empathy due to gender similarities |
| E8 | Empathy due to better connection |
| E9 | Empathy due to user interaction |
| E10 | Empathy due to improved user understanding |
| E11 | Empathy for technology literacy related issues |
| E12 | Empathy due to similar levels of technical ability |
| E13 | Empathy due to reflection in action |

Table 13. Developer Awareness and Enablers of Empathy

| Type of Awareness | Enablers of Developer Empathy | | | | | | | | | | | | |
|---------------------|-------------------------------|----|----|----|----|----|----|----|----|-----|-----|-----|-----|
| | E1 | E2 | E3 | E4 | E5 | E6 | E7 | E8 | E9 | E10 | E11 | E12 | E13 |
| Self Awareness | ☑ | | | | | | | | | | | | |
| Emotional Awareness | | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | ☑ | | | | |
| Technical Awareness | ☑ | ☑ | ☑ | | | | | | | ☑ | ☑ | ☑ | ☑ |

Table 14. Developer Strategies

| ID | Strategies to Overcome Barriers | ID | Strategies to Overcome Barriers |
|----|--|-----|---|
| S1 | Developer plans to prepare more | S10 | Developer thoughts to better phrase the re |
| S2 | Developer attempts to connect more with users | S11 | Reassure users |
| S3 | Provide better support when users are struggling | S12 | Rephrase and clarify when users seemed co |
| S4 | Pay more attention to user concerns | S13 | Developer attempt to talk in layman's term |
| S5 | Developer plans to slow down the session | S14 | Not overloading users with technical detail |
| S6 | Developer plans to ask more follow-up questions | S15 | Reinforcing the successful acts of users |
| S7 | Developers having the urge to help users when they struggle to use the app | S16 | Taking time to implement user suggested o |
| S8 | More agile way of interacting with end users | S17 | Think thoroughly before talking |
| S9 | Developer attempts to make users more comfortable | S18 | Personalising content based on the users |

Table 15. Developer Empathy Barriers and Strategies

| Developer Empathy Barriers | Strategies to Overcome Barriers | | | | | | | | |
|--|---------------------------------|----|----|----|----|----|----|----|----|
| | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 |
| Less developer interest | ✓ | ✓ | | | | | | | |
| Task centredness | | ✓ | ✓ | ✓ | | | | | |
| Unfamiliarity of features | ✓ | | | | | | | | |
| Limited time for reflection in action | | | | | | ✓ | ✓ | | |
| Difficulty in understanding user struggles | | ✓ | | ✓ | | ✓ | ✓ | | |
| Difficulty in empathising with low-tech. literacy related issues | | ✓ | | ✓ | | ✓ | ✓ | | |
| Less interaction | | ✓ | | | | | | ✓ | ✓ |
| Difficulty to resonate with collective experiences of women | ✓ | ✓ | | | | | | | |

Table 16. User Empathy Barriers and Strategies

| User Empathy Barriers | Strategies to Overcome Barriers | | | | | | | | | | | | |
|---|---------------------------------|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| | S2 | S4 | S7 | S9 | S10 | S11 | S12 | S13 | S14 | S15 | S16 | S17 | |
| Poor user connection with developers | ✓ | | | | | | | | | | | | |
| User nervousness due to time confusions | | | | ✓ | | | | | | | | | |
| Users being nervous by thinking about session questions | | | | ✓ | ✓ | ✓ | | | | | | | |
| User nervousness due to unfamiliarity of sessions | | | | | ✓ | | ✓ | ✓ | ✓ | ✓ | | | |
| User nervousness about providing useful feedback | | | | | | ✓ | | | | ✓ | ✓ | | |
| User confusion on terminology | | | | | | ✓ | ✓ | ✓ | ✓ | | | ✓ | |
| Users feeling awkward due to confusion | | ✓ | ✓ | ✓ | | ✓ | | | | | | | |

6.4 Key Insights & Reflections

We determined the following insights, reflections and emerging relationships based on our data analysis.

🧩 Grateful users: We noticed the grateful nature of the users. Users expressed their gratitude in many forms, both during the usability sessions directly to developers and in the interviews. It seems that the context of the application (eHealth) is one of the factors which influenced this grateful behaviour. Users were very grateful to developers for building an application which would assist them as PCOS patients. During the usability sessions users commended the developers for bringing their vision into life. Users seemed very pleased by the overall application and available features, and they even provided constructive feedback, suggesting improvements and new features like mood tracking. Notably, users did not provide negative feedback and expressed

their intention to use the AskPCOS app when it is released. Users appreciated the opportunity of participating in the usability sessions and commended the developers' flexibility. Even during the interviews, users expressed a positive sentiment about their overall experience.

⚙️ **Reflection in action:** Reflection in action helped developers to empathise and connect with the users. However, the limited time for reflecting during the sessions (*reflection in action*) was a barrier to developer empathy. In addition, developers had difficulties in understanding user feedback due to reflection on action. This situation aligns with the idea of Schön's reflective model [66] in professional education, emphasising the benefits of reflection in action (during event) over reflection on action (post-event). Our findings suggest that reflection in action enables developer empathy and reflection on action hinders developer empathy.

⚙️ **Closeness to code:** We identified developer familiarity has two parts: closeness to code and familiarity with user issues. Software developers are known to be defensive about their code and technical approaches [55]. However, we found that closeness to code enabled developer empathy. Developers showed more empathy towards user suggestions concerning the features they were closely working on. This familiarity helped developers to better understand user suggestions and empathise with their experiences.

⚙️ **Connection Vs familiarity Vs empathy:** We noticed an emerging relationship among familiarity, connection and enablers of empathy. We found that familiarity helped to establish a better connection and having a better connection enabled empathy. We also found that familiarity triggered empathy. Further we found that unfamiliarity led to a poor connection & it was also a barrier to empathy. Although it is said that '*familiarity breeds contempt*', we noticed that *familiarity breeds empathy*. We noticed these emerging relationships from both developers and users, suggesting a correlation among connection, familiarity, and empathy. Our insights align with established empathy research, supporting the notion that familiarity serves as an enabler for empathy [56]. Further, the framework developed by Thompson et al. emphasises that the individuals who are psychologically closer, and thereby more familiar, may evoke heightened empathetic responses [72]. This alignment with existing research strengthens our understanding of the interplay between connection, familiarity, and empathy in the dynamic context of our study.

⚙️ **Interaction Vs connection Vs empathy:** We observed an emerging relationship between interaction, connection and enablers of empathy. We found that interaction led to a better connection and a better connection enabled empathy. We also found that interaction enabled empathy. We identified the less interaction as a barrier to developer empathy and users also identified less interaction as a cause of poor connection with developers. These insights suggest that there could be a correlation between interaction, connection, and empathy.

⚙️ **Relatedness trigger empathy:** Our observations align with existing empathy research, which suggests that shared experiences foster empathy (Section 3). Various studies corroborate that relatedness acts as an enabler for empathic behaviour [6, 24, 40, 43, 60, 68, 76–78]. We saw evidence of this on the aspects of technology and gender. In this context, developers' technical competence and their ability to empathise with users' technical difficulties, as well as shared gender experiences, facilitated better understanding and empathy toward users. This substantiates the notion that relatedness, whether in technical expertise or gender, helps in enabling empathy within the context of our study.

⚙️ **Emotions and empathy:** We observed an emerging relationship between emotions and empathy. Emotions play a significant role in enabling or inhibiting empathy, as both positive and negative emotions were linked to empathy in our study. Positive emotions seemed to enable empathy, while negative emotions could either trigger or hinder empathy depending on the context

and individual experiences. For instance, positive emotions like *feeling fulfilled* and *feeling understood* enabled developers' and users' empathy respectively. In addition, even negative emotions like *guilt* triggered developer empathy. However, negative emotions like developer *frustration* and user *nervousness* were barriers to their empathy. Building upon existing research, we draw upon an integrative framework on empathy and emotion regulation by Thompson et al. to contextualise our findings [72]. This framework views empathy as a cybernetic, emotion-generative process and emphasises the role of regulatory strategies. Regulatory strategies like situation selection and cognitive change were found to influence distinct components of empathy and the resulting affective state in observers. Our study contributes to this theoretical framework by providing empirical evidence of how emotions, both positive and negative, interact with empathy in the dynamic context of software development.

6.5 Possible Implications for Software Practitioners

☐ Recommendation 1 – Proactive identification of empathy enablers and barriers: We have suggested some actions (Section 6.3) that might be considered by practitioners to identify the awareness required to enable empathy, determine enablers & barriers to empathy, and some strategies to overcome these barriers. Although there is no clear understanding on the impact of empathy on software engineering per se, empathy is highly beneficial for improving human connections. Having a healthy connection among software practitioners, users and even customers may positively influence the success of software projects.

☐ Recommendation 2 – Facilitate direct user feedback to developers: We noticed developers were very satisfied to receive positive user feedback. Positive feedback boosted developer confidence in their product and made them feel fulfilled. Direct interaction also helped developers better understand user pain points. Receiving feedback directly from end users may have had a more significant impact compared to the feedback developers typically receive from their managers. Thus, conducting usability testing sessions with developers and end users before launching software would be beneficial in fostering developer empathy towards user needs.

☐ Recommendation 3 – Pilot sessions to find best suited developer roles: We found that different developers preferred different session roles and their ability to interact with the users depended on their meeting role to some extent. Hence, we believe tailoring the session roles to the preferences and strengths of individual developers could have enhanced their ability to interact with users and understand feedback. In a real world setting, it would be beneficial to conduct a pilot session to identify the best suited role for each developer.

☐ Recommendation 4 – Find ways to support struggling users without compromising their feedback: We identified that developers were in a dilemma in wanting to empathise & help users, and their need to elicit unadulterated user responses to usability concerns. This created a trade-off between helping users and letting them struggle. There could be a better approach where developers can provide necessary emotional support to users without providing direct technical support, thus balancing empathy and feedback authenticity.

6.6 Implications for Researchers

🎓 Designing human experiments for empathy research: We identified several factors that made both developers and users uncomfortable. We believe these could be better managed when designing empathy research experiments involving humans. First, we determined that developers preferred different meeting roles and their ability to interact with the users depended on these roles. Having an understanding on the preferences and capacity of each participant may assist in designing a better experiment. We noticed a developer dilemma in helping users and

having proper strategies to manage such difficult situations would have been more beneficial. Hence we recommend considering these aspects when designing similar experiments.

📖 **Validate with a large sample:** A key contribution of our study is a set of actions (Section 6.3) which can be followed to build awareness as well as to identify empathy enablers, barriers and strategies to deal with these barriers. However, our findings are based on a limited number of participants and all of our developers were students. Hence researchers may consider replicating our study with more participants and including experienced software practitioners. This new study will help to validate and strengthen our findings.

📖 **Impact of empathy on SE and an SE-oriented empathy scale:** Our study did not focus on the impact of empathy on SE. It would be interesting to know the impact of empathy on software itself as well as software projects, software teams, and even other stakeholders such as customers, users, collaborators. Similarly, while designing and executing our study, we identified that it is quite difficult to apply existing empathy scales, mainly sourced from psychology. An SE-oriented empathy scale with SE-specific empathy measures or adapting existing scales to SE would be beneficial [35].

📖 **Developer awareness of user needs and emotions:** We found that developers were very attentive to users' emotional and technical needs. Direct interaction between developers and users could be a reason for this developer awareness. Developers being well aware of users' emotional and technical needs likely improved the usability of the apps. Developers noticed both positive and negative user emotions, and were more concerned about negative emotions. It is unclear how much of this behaviour is *Emotion Contagion* or *Perspective Taking* or *Empathic Concern* or is it *Social Awareness* or is it *Sympathy* or else is it just trying to be nice to users as they have a sensitive health condition. Developers were considerate about how users were feeling not just during the usability sessions, but even during interviews when reflecting on the usability sessions. Studying the impact of increasing direct interactions between developers and users may be beneficial to better understand this awareness of user needs and emotions.

📖 **Developer connection with users and grateful users:** During interviews, developers occasionally referred to empathy as connection. We noticed evidence of a relationship between empathy and connection (Section 6.4). Unfamiliarity and less interaction were the major causes of poor developer connection, and these were also identified as barriers to developer empathy. This suggests a correlation between the connection and empathy. Further, users expressed gratitude towards developers and none of the users had negative feedback regarding the application (Section 6.4). We hypothesise this may be a common behaviour among eHealth app users. Future studies could explore the relationship between empathy and connection and user gratitude in the context of eHealth apps.

📖 **Reflection in action and closeness to code:** Findings suggest that reflection in action enables developer empathy, and reflection on action hinders developer empathy (Section 6.4). We also identified closeness to code as an empathy enabler. However, due to limited number of participants, we cannot generalise these findings.

📖 **Emerging relationships:** We observed an emerging relationship between familiarity, connection and empathy as well as between interaction, connection and empathy. In addition we noted emerging relationships between relatedness & empathy, and emotions & empathy. These emerging relationships are worth studying further.

📖 **Varying effects of empathy enablers and barriers over time:** While our primary means of identifying empathy enablers and barriers relied on semi-structured interviews, it is important to note that these interviews took place at the conclusion of the 24-week study period. Therefore, we faced limitations in capturing the dynamic changes and evolving effects of enablers and barriers over time. Additionally, other contextual factors influencing these dynamics

could not be explored due to the retrospective nature of our interviews. Future research should consider longitudinal data collection methods to better understand the temporal aspects of empathy dynamics in software development contexts.

7 Limitations

In our study, data collection was limited to a specific context, with participants exclusively composed of final year undergraduate students majoring in IT. While our sample size included a substantial number of developers, it is crucial to acknowledge the limited diversity in our participant pool, as they were predominantly from Australia. Consequently, the findings derived from our study may not be universally applicable to the broader community of software practitioners and users. Given that the majority of our participants were students in their final year of undergraduate studies in IT from Australia, the external validity of our results is constrained, and the generalisability of our case study is thereby limited. It is important to recognise that our research scope focused on a particular demographic within the SE community, which may not fully represent the diverse landscape of practitioners worldwide. To enhance the robustness of our conclusions, future studies should consider replicating this research with a more heterogeneous sample. Including experienced software practitioners from various professional backgrounds, as well as expanding the study to different countries and contextual settings, would contribute to a more comprehensive understanding of the phenomena under investigation.

In addition, deeper statistical analysis could be done on future larger datasets. This approach would mitigate the limitations associated with external validity and provide a more nuanced and widely applicable insight into the broader SE community. Further the implications presented for software practitioners need to be validated through additional studies, as they are derived from this specific case study.

Although there are many empathy tests, none of the scales were designed specifically for SE context. We selected QCAE with the recommendation of empathy experts as it seems the best matching scale to our study. However, participants' understanding on the empathy test can vary. Different human aspects such as age, gender, ethnicity, and other psychological differences may influence the ratings of items on the QCAE. The self-report nature of the scale may have an impact on the overall score, thus participant empathy scores may be less reflective of how they actually demonstrated empathy. Empathy is regarded as a favourable trait and participants may have been tempted to respond in a more socially desirable way while filling the empathy test as well as during the interviews. We considered only baseline and round one empathy scores of developers and users. Even though we have round two scores of all the developers, we do not have round two score of one user. Hence we are using only baseline and round one scores while reporting our findings to make the presentation of data similar among both groups of participants.

Usability testing sessions were conducted using Zoom platform due to the COVID-19 pandemic and geographical distribution of the users. Therefore our observation study was also carried out online. We may have missed some empathy cues due to the limitations in Zoom setting while observing participants. We tried to rectify this situation by having two observers, video recording these sessions and watching these recordings multiple times. However, due to human errors we may still have missed some of the empathy cues. In the interviews, developers pointed out that users might have felt stressed or under pressure due to the recording of the usability sessions. This could have potentially limited users' capacity to express themselves openly, representing a limitation of the study. The follow-up interviews allowed participants to retrospectively assess the conduct of the usability sessions and suggest improvements. Incorporating a similar approach earlier in the study could have been beneficial.

We used STGT for qualitative data analysis. We generated concepts, subcategories and categories based on the codes. However, codes generated by a single researcher could be subjective and can lead to a potentially limited view of data. Hence after the first author conducted the initial coding and analysis, it was shared with all the other authors to resolve any conflicts. The third author who is well-experienced in STGT, peer reviewed all the codes, concepts, subcategories and categories. We discussed all the conflicts and carried out several discussion rounds to finalise the code book. We also had fortnight meetings where the first author discussed the code book with other authors, allowing for peer review and feedback.

8 Conclusion

We conducted an empirical case study to understand how empathy is practised in the interactions between developers and end users. We employed an empathy test, a demographics questionnaire, an observation study and a set of in-depth interviews to collect data. Mixed methods were used including STGT for qualitative data and descriptive statistics for quantitative data analysis. We identified some enablers of empathy and the nature of awareness needed to trigger empathy. We determined some barriers to empathy and strategies that could be employed to overcome these barriers. Based on our findings, we report a set of actions that can be used to identify the types of awareness required to enable empathy, as well as a set of strategies to overcome empathy barriers. We identified verbal and nonverbal empathy cues demonstrated by participants during the observation sessions. We determined some trends using the scores of empathy test. We report insights on emerging relationships and differentiate empathy enablers based on cognitive and affective empathy. Extending the findings of this study will be beneficial for both software practitioners and research community. We presented a set of recommendations and potential future works for software practitioners and researchers.

Acknowledgments

Gunatilake, Grundy and Mueller are supported by ARC Laureate Fellowship FL190100035. We thank all our participants for their tremendous contribution and all the other stakeholders of the project including the MCHRI and MADA.

References

- [1] Gabriella Airenti. 2015. The cognitive bases of anthropomorphism: from relatedness to empathy. *International Journal of Social Robotics* 7 (2015), 117–127.
- [2] Ali E. Akgün, Halit Keskin, A. Yavuz Cebecioglu, and Derya Dogan. 2015. Antecedents and consequences of collective empathy in software development project teams. *Information & Management* 52, 2 (2015), 247–259. <https://doi.org/10.1016/j.im.2014.11.004>
- [3] Simon Baron-Cohen and Sally Wheelwright. 2004. The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of autism and developmental disorders* 34 (2004), 163–175.
- [4] Caroline A Bartel and Richard Saavedra. 2000. The collective construction of work group moods. *Administrative Science Quarterly* 45, 2 (2000), 197–231.
- [5] C Daniel Batson and Nadia Y Ahmad. 2009. Using empathy to improve intergroup attitudes and relations. *Social issues and policy review* 3, 1 (2009), 141–177.
- [6] C. Daniel Batson, Susie C. Sympson, Jennifer L. Hindman, Peter Decruz, R. Matthew Todd, Joy L. Weeks, Geoffrey Jennings, and Christopher T. Burns. 1996. "I've Been there, Too": Effect on Empathy of Prior Experience with a Need. *Personality and Social Psychology Bulletin* 22, 5 (1996), 474–482. <https://doi.org/10.1177/0146167296225005> arXiv:<https://doi.org/10.1177/0146167296225005>
- [7] Teresa Blanco, Ignacio López-Forniés, and Francisco Javier Zarazaga-Soria. 2017. Deconstructing the Tower of Babel: a design method to improve empathy and teamwork competences of informatics students. *International Journal of Technology and Design Education* 27, 2 (2017), 307–328.

- [8] Edna Dias Canedo, Ana Carolina Dos Santos Pergentino, Angelica Toffano Seidel Calazans, Frederico Viana Almeida, Pedro Henrique Teixeira Costa, and Fernanda Lima. 2020. Design Thinking Use in Agile Software Projects: Software Developers' Perception. In *In Proceedings of the 22nd International Conference on Enterprise Information Systems (ICEIS2020)*, Vol. 978-989-758-423-7. SCITEPRESS–Science and Technology Publications, 217–224. <https://doi.org/10.5220/0009387502170224>
- [9] Tracy G Cassels, Sherilynn Chan, and Winnie Chung. 2010. The role of culture in affective empathy: Cultural and bicultural differences. *Journal of Cognition and Culture* 10, 3-4 (2010), 309–326.
- [10] Tanya L Chartrand and John A Bargh. 1999. The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology* 76, 6 (1999), 893.
- [11] William J Chopik, Ed O'Brien, and Sara H Konrath. 2017. Differences in empathic concern and perspective taking across 63 countries. *Journal of Cross-Cultural Psychology* 48, 1 (2017), 23–38.
- [12] Malissa A Clark, Melissa M Robertson, and Stephen Young. 2019. 'I feel your pain': A critical review of organizational research on empathy. *Journal of Organizational Behavior* 40, 2 (2019), 166–192.
- [13] Benjamin MP Cuff, Sarah J Brown, Laura Taylor, and Douglas J Howat. 2016. Empathy: A review of the concept. *Emotion review* 8, 2 (2016), 144–153.
- [14] Mark Davis. 1980. A multidimensional approach to individual differences in empathy. *JASAS Catalog of Selected Documents in Psychology* 10 (1980).
- [15] Mark H Davis. 1983. Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of personality and social psychology* 44, 1 (1983), 113.
- [16] Menno De Jong and Leo Lentz. 2007. Professional writers and empathy: Exploring the barriers to anticipating reader problems. In *2007 IEEE International Professional Communication Conference*. IEEE, Seattle, WA, USA, 1–8. <https://doi.org/10.1109/IPCC.2007.4464058>
- [17] Frans BM De Waal. 2008. Putting the altruism back into altruism: The evolution of empathy. *Annu. Rev. Psychol.* 59 (2008), 279–300.
- [18] Jean Decety and Philip L. Jackson. 2004. The Functional Architecture of Human Empathy. *Behavioral and Cognitive Neuroscience Reviews* 3, 2 (2004), 71–100. <https://doi.org/10.1177/1534582304267187> arXiv:<https://doi.org/10.1177/1534582304267187> PMID: 15537986.
- [19] Jean Decety and Yoshiya Moriguchi. 2007. The empathic brain and its dysfunction in psychiatric populations: Implications for intervention across different clinical conditions. *BioPsychoSocial medicine* 1, 1 (2007), 1–21.
- [20] Duleep Delpechitre. 2013. Review and assessment of past empathy scales to measure salesperson's empathy. *Journal of Management and Marketing Research* 13, 1 (2013), 1–16.
- [21] Frans AWM Derksen, Tim C olde Hartman, Jozien M Bensing, and Antoine LM Lagro-Janssen. 2016. Managing barriers to empathy in the clinical encounter: a qualitative interview study with GPs. *British Journal of General Practice* 66, 653 (2016), e887–e895.
- [22] Nancy Eisenberg and Natalie D Eggum. 2009. Empathic responding: Sympathy and personal distress. *The social neuroscience of empathy* 6, 2009 (2009), 71–830.
- [23] Nancy Eisenberg, Tracy L Spinrad, and Zoe E Taylor. 2014. Sympathy. In *The Handbook of Virtue Ethics* (1 ed.). Routledge, London, 409–417. <https://doi.org/10.4324/9781315729053>
- [24] JAKOB EKLUND, TERESIA ANDERSSON-STRÅBERG, and ERIC M. HANSEN. 2009. "I've also experienced loss and fear": Effects of prior similar experience on empathy. *Scandinavian Journal of Psychology* 50, 1 (2009), 65–69. <https://doi.org/10.1111/j.1467-9450.2008.00673.x> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9450.2008.00673.x>
- [25] Paul Ekman and Wallace V Friesen. 1971. Constants across cultures in the face and emotion. *Journal of personality and social psychology* 17, 2 (1971), 124.
- [26] Robert Elliott, Arthur C Bohart, Jeanne C Watson, and Leslie S Greenberg. 2011. Empathy. *Psychotherapy* 48, 1 (2011), 43.
- [27] Natalie Ewin, Ritesh Chugh, Olav Muurlink, Jacqueline Jarvis, and Jo Luck. 2021. Empathy of project management students and why it matters. *Procedia Computer Science* 181 (2021), 503–510.
- [28] Bruna Ferreira, Williamson Silva, Edson Oliveira, and Tayana Conte. 2015. Designing Personas with Empathy Map. In *Proceedings of the International Conference on Software Engineering and Knowledge Engineering, SEKE*, Vol. 152. Pittsburgh, USA. <https://doi.org/10.18293/SEKE2015-152>
- [29] Bruna Moraes Ferreira, Simone DJ Barbosa, and Tayana Conte. 2016. PATHY: Using empathy with personas to design applications that meet the users' needs. In *International Conference on Human-Computer Interaction*. Springer, Toronto, ON, Canada, 153–165.
- [30] Alvin I. Goldman. 2011. 313 Two Routes to Empathy: Insights from Cognitive Neuroscience. In *Empathy: Philosophical and Psychological Perspectives*. Oxford University Press, Oxford, 31–44. <https://doi.org/10.1093/>

- acprof:oso/9780199539956.003.0004 arXiv:https://academic.oup.com/book/0/chapter/151861762/chapter-ag-pdf/44974436/book_7216_section_151861762.ag.pdf
- [31] Daniel Goleman. 1996. *Emotional Intelligence why it can matter more IQ*. Vol. 53. Bloomsbury Publishing, Great Britain. 1689–1699 pages. Issue 9.
 - [32] Daniel Goleman, Richard Boyatzis, and Annie McKee. 2002. The emotional reality of teams. *Journal of Organizational Excellence* 21, 2 (2002), 55–65.
 - [33] Ulrike M. Graetsch, Hourieh Khalajzadeh, Mojtaba Shahin, Rashina Hoda, and John Grundy. 2023. Dealing With Data Challenges When Delivering Data-Intensive Software Solutions. *IEEE Transactions on Software Engineering* (2023), 1–23. <https://doi.org/10.1109/TSE.2023.3291003>
 - [34] Medical School Objectives Writing Group et al. 1999. Learning objectives for medical student education—guidelines for medical schools: report I of the Medical School Objectives Project. *Academic Medicine* 74, 1 (1999), 13–18.
 - [35] Hashini Gunatilake, John Grundy, Ingo Mueller, and Rashina Hoda. 2023. Empathy models and software engineering—A preliminary analysis and taxonomy. *Journal of Systems and Software* 203 (2023), 111747.
 - [36] Michaela Guthridge and Melita J Giummarra. 2021. The taxonomy of empathy: A meta-definition and the nine dimensions of the empathic system. *Journal of Humanistic Psychology* 0 (2021), 00221678211018015. <https://doi.org/10.1177/00221678211018015>
 - [37] Samer Halabi, John F Dovidio, and Arie Nadler. 2008. When and how do high status group members offer help: Effects of social dominance orientation and status threat. *Political Psychology* 29, 6 (2008), 841–858.
 - [38] Jodi Halpern. 2003. What is clinical empathy? *Journal of general internal medicine* 18 (2003), 670–674.
 - [39] Susan J Hazel, Tania D Signal, and Nicola Taylor. 2011. Can teaching veterinary and animal-science students about animal welfare affect their attitude toward animals and human-related empathy? *Journal of veterinary medical education* 38, 1 (2011), 74–83.
 - [40] Miriam S. Heinke and Winnifred R. Louis. 2009. Cultural Background and Individualistic–Collectivistic Values in Relation to Similarity, Perspective Taking, and Empathy. *Journal of Applied Social Psychology* 39, 11 (2009), 2570–2590. <https://doi.org/10.1111/j.1559-1816.2009.00538.x> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1559-1816.2009.00538.x>
 - [41] Rashina Hoda. 2021. Socio-technical grounded theory for software engineering. *IEEE Transactions on Software Engineering* 48, 10 (2021), 3808–3832.
 - [42] Rashina Hoda, James Noble, and Stuart Marshall. 2011. The impact of inadequate customer collaboration on self-organizing Agile teams. *Information and software technology* 53, 5 (2011), 521–534.
 - [43] Martin L Hoffman. 2001. *Empathy and moral development: Implications for caring and justice*. Cambridge University Press, Cambridge, UK.
 - [44] Geert Hofstede. 1991. *Empirical models of cultural differences*. Swets & Zeitlinger Publishers, Lisse, Netherlands, 4–20.
 - [45] Mohammadreza Hojat. 2016. *Empathy in health professions education and patient care*. Springer, Switzerland. <https://doi.org/10.1007/978-3-319-27625-0>
 - [46] Mohammadreza Hojat, Jennifer DeSantis, Stephen C Shannon, Luke H Mortensen, Mark R Speicher, Lynn Bragan, Marianna LaNoue, and Leonard H Calabrese. 2018. The Jefferson Scale of Empathy: a nationwide study of measurement properties, underlying components, latent variable structure, and national norms in medical students. *Advances in Health Sciences Education* 23, 5 (2018), 899–920.
 - [47] Mohammadreza Hojat, Salvatore Mangione, Thomas J Nasca, Mitchell JM Cohen, Joseph S Gonnella, James B Erdmann, Jon Veloski, and Mike Magee. 2001. The Jefferson Scale of Physician Empathy: development and preliminary psychometric data. *Educational and psychological measurement* 61, 2 (2001), 349–365.
 - [48] J Howick and S Rees. 2017. Overthrowing barriers to empathy in healthcare: empathy in the age of the Internet. *Journal of the Royal Society of Medicine* 110, 9 (2017), 352–357. <https://doi.org/10.1177/0141076817714443> arXiv:<https://doi.org/10.1177/0141076817714443>
 - [49] Guste Ilgunaite, Luciano Giromini, and Marzia Di Girolamo. 2017. Measuring empathy: A literature review of available tools. *BPA-Applied Psychology Bulletin (Bollettino di Psicologia Applicata)* 65, 280 (2017), 2–28.
 - [50] Parvaneh Yaghoubi Jami, David Ian Walker, and Behzad Mansouri. 2023. Interaction of empathy and culture: a review. *Current Psychology* 42 (2023), 1–16. <https://doi.org/10.1007/s12144-023-04422-6>
 - [51] Devi Karolita, Jennifer McIntosh, Tanjila Kanij, John Grundy, and Humphrey O. Obie. 2023. Use of personas in Requirements Engineering: A systematic mapping study. *Information and Software Technology* 162 (2023), 107–264. <https://doi.org/10.1016/j.infsof.2023.107264>
 - [52] Shinobu Kitayama, Hazel Rose Markus, and Masaru Kurokawa. 2000. Culture, emotion, and well-being: Good feelings in Japan and the United States. *Cognition & Emotion* 14, 1 (2000), 93–124.
 - [53] Meira Levy. 2018. Educating for Empathy in Software Engineering Course. In *Joint Proceedings of REFSQ-2018 Workshops, Doctoral Symposium, Live Studies Track, and Poster Track co-located with the 23rd International Conference on Requirements Engineering: Foundation for Software Quality (REFSQ 2018), Utrecht, The Netherlands, March 19, 2018*

- (*CEUR Workshop Proceedings, Vol. 2075*), Klaus Schmid, Paola Spoletini, Eya Ben Charrada, Yoram Chisik, Fabiano Dalpiaz, Alessio Ferrari, Peter Forbrig, Xavier Franch, Marite Kirikova, Nazim H. Madhavji, Cristina Palomares, Jolita Ralyté, Mehrdad Sabetzadeh, Pete Sawyer, Dirk van der Linden, and Anna Zamansky (Eds.). CEUR-WS.org, Utrecht, The Netherlands, 1–9. https://ceur-ws.org/Vol-2075/FIRE18_paper2.pdf
- [54] Meira Levy and Irit Hadar. 2018. The importance of empathy for analyzing privacy requirements. In *2018 IEEE 5th International Workshop on Evolving Security & Privacy Requirements Engineering (ESPRe)*. IEEE, Banff, Canada, 9–13. <https://doi.org/10.1109/ESPRe.2018.00008>
- [55] John McAvoy and Tom Butler. 2009. A failure to learn by software developers: Inhibiting the adoption of an agile software development methodology. *Journal of Information Technology Case and Application Research* 11, 1 (2009), 23–46.
- [56] Yuki Motomura, Akira Takeshita, Yuka Egashira, Takayuki Nishimura, Yeon-kyu Kim, and Shigeki Watanuki. 2015. Interaction between valence of empathy and familiarity: is it difficult to empathize with the positive events of a stranger? *Journal of physiological anthropology* 34 (2015), 1–9.
- [57] David. L. Neumann, Raymond C.K. Chan, Gregory. J. Boyle, Yi Wang, and H. Rae Westbury. 2015. Chapter 10 - Measures of Empathy: Self-Report, Behavioral, and Neuroscientific Approaches. In *Measures of Personality and Social Psychological Constructs*, Gregory J. Boyle, Donald H. Saklofske, and Gerald Matthews (Eds.). Academic Press, San Diego, 257–289. <https://doi.org/10.1016/B978-0-12-386915-9.00010-3>
- [58] Jennifer Nicolai, Ralf Demmel, and Jutta Hagen. 2007. Rating scales for the assessment of empathic communication in medical interviews (REM): Scale development, reliability, and validity. *Journal of Clinical Psychology in Medical Settings* 14 (2007), 367–375.
- [59] Paula Nunes, Stella Williams, Bidyadhar Sa, and Keith Stevenson. 2011. A study of empathy decline in students from five health disciplines during their first year of training. *Int J Med Educ* 2 (2011), 12–17.
- [60] M.A. Preis and B. Kroener-Herwig. 2012. Empathy for pain: The effects of prior experience and sex. *European Journal of Pain* 16, 9 (2012), 1311–1319. <https://doi.org/10.1002/j.1532-2149.2012.00119.x> arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1532-2149.2012.00119.x>
- [61] Stephanie D Preston and Frans BM De Waal. 2002. Empathy: Its ultimate and proximate bases. *Behavioral and brain sciences* 25, 1 (2002), 1–20.
- [62] Renate LEP Reniers, Rhiannon Corcoran, Richard Drake, Nick M Shryane, and Birgit A Völlm. 2011. The QCAE: A questionnaire of cognitive and affective empathy. *Journal of personality assessment* 93, 1 (2011), 84–95.
- [63] Per Runeson and Martin Höst. 2009. Guidelines for conducting and reporting case study research in software engineering. *Empirical software engineering* 14 (2009), 131–164.
- [64] Peter Salovey and John D Mayer. 1990. Emotional intelligence. *Imagination, cognition and personality* 9, 3 (1990), 185–211.
- [65] Daniel Schmelzer. 2015. *Empathie in Teams: Empathieverhalten und Empathiekreisläufe*. Ph. D. Dissertation. München, Technische Universität München, Diss., 2015.
- [66] Donald A Schön. 2017. *The reflective practitioner: How professionals think in action* (1st ed.). Routledge, London. <https://doi.org/10.4324/9781315237473>
- [67] Seth J Schwartz and Jennifer B Unger. 2010. Biculturalism and context: What is biculturalism, and when is it adaptive?: Commentary on Mistry and Wu. *Human development* 53, 1 (2010), 26.
- [68] Nancy E. Snow. 2000. Empathy. *American Philosophical Quarterly* 37, 1 (2000), 65–78. <http://www.jstor.org/stable/20009985>
- [69] Fariba Taleghani, Elaheh Ashouri, Mehrdad Memarzadeh, and Mortaza Saburi. 2018. Barriers to empathy-based care: oncology nurses’ perceptions. *International Journal of Health Care Quality Assurance* 31, 3 (2018), 249–259.
- [70] Steven J Taylor and Robert C Bogdan. 1984. *Introduction to qualitative research methods: The search for meanings*. Wiley, New York.
- [71] Steven J Taylor, Robert C Bogdan, and Marjorie L. DeVault. 2015. *Introduction to Qualitative Research Methods: A Guidebook and Resource*. Wiley, New York.
- [72] Nicholas M. Thompson, Andero Uusberg, James J. Gross, and Bhismadev Chakrabarti. 2019. Chapter 12 - Empathy and emotion regulation: An integrative account. In *Emotion and Cognition*, Narayanan Srinivasan (Ed.). Progress in Brain Research, Vol. 247. Elsevier, 273–304. <https://doi.org/10.1016/bs.pbr.2019.03.024>
- [73] Zachary Wallmark, Choi Deblieck, and Marco Iacoboni. 2018. Neurophysiological effects of trait empathy in music listening. *Frontiers in behavioral neuroscience* 12 (2018), 66. <https://doi.org/10.3389/fnbeh.2018.00066>
- [74] Claes Wohlin, Per Runeson, Martin Höst, Magnus C. Ohlsson, Björn Regnell, and Anders Wesslén. 2012. *Experimentation in software engineering* (1st ed.). Vol. 9783642290442. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/978-3-642-29044-2>
- [75] Shali Wu and Boaz Keysar. 2007. The effect of culture on perspective taking. *Psychological science* 18, 7 (2007), 600–606.

- [76] Xiaojing Xu, Xiangyu Zuo, Xiaoying Wang, and Shihui Han. 2009. Do You Feel My Pain? Racial Group Membership Modulates Empathic Neural Responses. *Journal of Neuroscience* 29, 26 (2009), 8525–8529. <https://doi.org/10.1523/JNEUROSCI.2418-09.2009> arXiv:<https://www.jneurosci.org/content/29/26/8525.full.pdf>
- [77] Parvaneh Yaghoubi Jami, Hyemin Han, Stephen J Thoma, Behzad Mansouri, and Rick Houser. 2021. Do histories of painful life experiences affect the expression of empathy among young adults? An electroencephalography study. *Frontiers in Psychology* 12 (2021), 689304.
- [78] Parvaneh Yaghoubi Jami, David Ian Walker, and Stephen J Thoma. 2023. Young adults’ empathic responses to others in psychological pain as compared to physical pain: does prior experience of pain matter? *Current Psychology* 42, 8 (2023), 6194–6215.
- [79] Juping Yu and Maggie Kirk. 2009. Evaluation of empathy measurement tools in nursing: systematic review. *Journal of advanced nursing* 65, 9 (2009), 1790–1806.

A Memos

Memo on “Developer Technical Support”

When we asked developers for examples of empathising with users, some stated that they modified the questions based on the scenarios and personalised the content for users during the usability sessions especially when users seemed confused [P1]. They stated that it was an act of consideration towards each individual user. Also, when asked about how well developers think users understood the shared technical details developers stated that they implicitly tried to talk in layman’s terms when dealing with the users [P3]. Developers stated that they are confident that the users sufficiently understood the technical details as clearly demonstrated by users’ accurate actions during usability session tasks. This was confirmed by the responses of the users. All the users stated that they were able to clearly understand and follow developer instructions. This suggests that modifying or personalising content for easier user understanding is one of the first things that comes to their mind when thinking of empathy. When we inquired about user’s understanding of technical details, developers stated that they didn’t want to overload users with technical details that they couldn’t understand. They also said that they explicitly tried to phrase questions as clearly as possible for first time as they get quite nervous while hosting the usability sessions [P1, P6]. This demonstrates that the developers tried to support the users by giving the best possible instructions.

B Interview Guides

B.1 Interview Guide of Software Developers

General Information:

- (1) Do you study a double degree?
- (2) (If Yes) What are your majors?
- (3) (If following a non-SE Degree) I want to learn a bit more about the non-SE degree that you follow. Do you have close contact with customers outside the university, for example during your projects and internships?
- (4) I am going to ask you to rate your affinity to technology vs people. If we consider a scale of 4, where 0 is very human-centric (more affinity to people), 1 is somewhat human-centric, 2 somewhat technology-centric and 3 fully technology-centric, how would you rate yourself?

Related to the experience of interacting with end users:

In this study, empathy refers to understanding a person from his or her frame of reference rather than one’s own, or vicariously experiencing that person’s feelings, perceptions, and thoughts.

- (5) Based on this definition, how well do you think you were able to empathise with the end users? Why do you think so? Can you share an example of a time when you were able to empathise with the end users?
- (6) Were there any instances where it was difficult for you to empathise with the end users? Can you share an example?

- (7) Thinking back now, was there a time when you could have empathised better with the end users?
- (8) If you could do it again, is there anything you would like to change with respect to engaging with the end users / empathising with them? Why?
- (9) Do you recognise any special reasons as to why you were able to empathise with end users or why you were unable to empathise with the end users? Can you share an example?
- (10) When you explain certain technical aspects or limitations in the application, do you think end users were able to fully understand what you explained? Why do you think so?
- (11) How did you feel when you were conducting the usability sessions?
- (12) Finally, is there anything else you would like to share? Any further feedback for us on your experience on this project?

B.2 Interview Guide of End Users

General Information:

- (1) Have you ever worked with software developers before? Or do you have any close contact with software developers?
- (2) I am going to ask you to rate your affinity to technology vs people. If we consider a scale of 4, where 0 is very human-centric (more affinity to people), 1 is somewhat human-centric, 2 somewhat technology-centric and 3 fully technology-centric, how would you rate yourself?

Related to the experience of interacting with end users:

In this study, empathy refers to understanding a person from his or her frame of reference rather than one's own, or vicariously experiencing that person's feelings, perceptions, and thoughts.

- (3) Based on this definition, how well do you think you were able to empathise with the developers? Why do you think so? Can you share an example of a time when you were able to empathise with the developers?
- (4) Were there any instances where it was difficult for you to empathise with the developers? Can you share an example?
- (5) Thinking back now, was there a time when you could have empathised better with the developers?
- (6) If you could do it again, is there anything you would like to change with respect to engaging with the developers / empathising with them? Why?
- (7) Do you recognise any special reasons as to why you were able to empathise with developers or why you were unable to empathise with them? Can you share an example?
- (8) When the developers explained certain technical aspects or limitations in the application, do you think you were able to fully understand what they explained? Why do you think so?
- (9) How did you feel when you were participating in the usability sessions?
- (10) Finally, is there anything else you would like to share regarding the study? Any further feedback for us on your experience on this project?

accepted 12 Jan 2024