Why stressed, Mom?: Exploring Family Reflection on Social and Emotional Sensor Data through Family Informatics

Hyesoo Park
School of Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
hye@gatech.edu

Sueun Jang
School of Computing
KAIST
Daejeon, Republic of Korea
sueun.jang@kaist.ac.kr

Hyunsoo Lee
Graduate School of Data Science
KAIST
Daejeon, Republic of Korea
hslee90@kaist.ac.kr

Jennifer G Kim*
School of Interactive Computing
Georgia Institute of Technology
Atlanta, Georgia, USA
jennifer.kim@cc.gatech.edu

Uichin Lee*
School of Computing
KAIST
Daejeon, Republic of Korea
uclee@kaist.ac.kr

Abstract

While family informatics has been developed for monitoring and tracking family-centered health data, there remains a gap in understanding how family informatics can support families in reflecting on their social behaviors and emotional dynamics. We address this gap with SELaD, a system that captures and visualizes social-emotional data from daily family interactions using audio, video, and physiological sensors. In a semi-naturalistic study with 17 families (n = 51), we investigated how this data facilitates reflection. Our findings reveal a process we term *relational reflection*, where families collaboratively interpret multimodal data to deepen their understanding of conversational dynamics and emotional influences by recalling their shared history and expectation of good communication. This process was particularly enriched by emotional data from multiple sources that families could cross-reference and reconcile. This work presents SELaD as a technology probe and empirically grounds the concept of relational reflection, positioning it as a foundation for designing future reflective technologies.

CCS Concepts

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

Keywords

Family Informatics, Relational Reflection, Multimodal Sensing, Social-Emotional Data, Affective Computing

ACM Reference Format:
Hyesoo Park, Sueun Jang, Hyunsoo Lee, Jennifer G Kim, and Uichin Lee. 2026. Why stressed, Mom?: Exploring Family Reflection on Social and Emotional Sensor Data through Family Informatics. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 23 pages. https://doi.org/10.1145/3772318.3790745

*Both authors served as co-senior authors and jointly supervised the project.

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3790745

1 Introduction

HCI researchers have increasingly explored family-centered health tracking, expanding the scope of traditional personal informatics into the emerging space of *family informatics* [88]. This concept began to take shape in the HCI literature with growing interest in how families collaboratively engage with health and well-being [87, 93]. Although existing family informatics systems have primarily focused on capturing individual or aggregated health data within families (e.g., sleep [87], physical activity [68, 93], diabetes symptoms [16], and mood [68, 100]), less attention has been given to how families collectively engage with and interpret their shared social and emotional experiences. One reason for this gap is that conducting research on social and emotional experiences requires multimodal sensing of these states and presenting such data to families, which has posed technical challenges. This remains an underexplored area, despite the research communities' longstanding interest in the role of technology in shaping social dynamics, collaboration, and mutual support within close-knit groups.

Some recent studies have begun to explore this gap, showing the feasibility of using sensor-aided family informatics to support reflection on affective states during shared activities [68]. Social and emotional data of a family can reveal their unique social dynamics, such as ripple effects among family members [88] or the roles of parents and children [87]. Understanding such family interactions through deep reflection can help discover communal norms, strengthen family relationships, and foster children's social-emotional development [18]. Family interactions provide a critical context for children's social-emotional learning [86, 98], offering opportunities to reflect on social awareness, emotional awareness, and communication behaviors. Furthermore, given the inherently reciprocal nature of these interactions (e.g., modeling emotional regulation, acknowledging their own emotions, and engaging in shared reflection), parents' active participation is essential for supporting children's healthy social-emotional development. However, to date, sensor-aided family informatics has rarely been applied to observe everyday family social interaction; most deployments remain confined to special purposes such as parent-driven instructional scenarios [48, 51]. Thus, there is limited empirical evidence on how sensing can support balanced, symmetric reflection in ordinary family life. At the same time, placing sensing technology in intimate family settings introduces possible harm, such as privacy leakage, power asymmetry, and overreliance on data that can erode trust and exacerbate familial tension.

To explore this tension, we draw on the notion of dialogic reflection by Fleck and Fitzpatrick [32], originally defined as "seeing things from a different perspective and considering alternatives," by focusing on its interpersonal dimension as a process of jointly interpreting shared experiences that leads to significant shifts in social-emotional understanding. In this work, we prototype SELaD, a family informatics system designed to capture, quantify, and visualize social-emotional data from audio, video, and physiological sensors during everyday interactions (e.g., mealtime conversations) to facilitate dialogic reflection for social-emotional learning within families. SELaD thus serves both as a proof-of-concept for extending multimodal sensing into daily life and as a technology probe [47] for investigating design tensions that arise when such data are collected and presented within the family. With SELaD, we set out the following research questions to investigate how families leverage multimodal sensor data of their social-emotional interactions and how the system supports families' sensemaking of the data:

RQ1) How do families reflect on the family conversation by interpreting multimodal social and emotional data?
RQ2) What insights do families gain through reflection?
RQ3) What factors contribute to deepening the reflection?

To answer these RQs, we collected data through a user study involving 17 families (n = 51) in a semi-naturalistic home setting to examine reflective processes, evaluate usability, and conduct interviews.

Our findings show that families engaged in relational reflection when interpreting multimodal social-emotional data with SELaD. They recalled a shared family history and shared expectation of good communication, and collaboratively interpreted the meanings of data, which deepened their understanding of the conversational pattern and emotional influence (Section 6.1.) We also identified factors that facilitate this relational reflection, and observed that the nature of emotional data and how the family engaged in the reflection could shape relational reflection (Section 6.2).

Collectively, the key contributions of our study are as follows:

- We designed and implemented SELaD, a family informatics prototype that integrates audio, video, and physiological data streams from the mealtime conversation of the family into video playback, data chart, and question-type prompts that can scaffold users to reflect on their behaviors.
- We conducted a user study of 17 families (n = 51) and performed an in-depth qualitative analysis that identified the co-constructed insights about family dynamics that families gained from relational reflection, and how reflection could be facilitated by the emotional data.
- We frame *relational reflection* with the study results and discuss its design implications around (i) technologies leveraging multimodal social-emotional data, (ii) prompts to scaffold the data reflection of families, (iii) strategies to mitigate the ethical concerns.

2 Background and Related Work

2.1 From Personal Informatics to Family Informatics

Since health and well-being are inherently socially interconnected, it has been argued that future research should consider various interpersonal contexts [90]. However, prior work in HCI has primarily focused on social use of personal data and personal informatics (e.g., within significant others [70, 71], through social media [29, 108]) to drive behavior change, rather than examining data that is co-constructed or collected within interpersonal contexts.

The concept of family informatics extends the framework of personal informatics, which traditionally supports individuals in collecting and reflecting on self-tracked data, such as physical, emotional, and social metrics, to derive insights for well-being [88]. Unlike personal informatics, which centers on individual experiences, family informatics views families as interconnected networks where members share routines, behaviors, and health-related decisions [10, 20, 75]. From this perspective, family members exhibit collective responses to health outcomes [107, 109], meaning that their shared behavior patterns and coping strategies can positively and negatively influence well-being.

Family informatics leverages family-generated data collected through sensor-rich tracking tools to assess behavioral domains such as sleep [87], diet [94], and physical activity [93]. Moreover, existing family tracking systems center data related to specific health conditions of the child, such as blood glucose levels [105], mood data [100], symptoms and medications [16], which have been shown to support self-regulation and self-care of children. However, most studies focus on tracking and visualizing data rather than fostering active discussions that enable family members to gain self-knowledge and understand each other's experiences [33].

The process of engaging in meaningful discussions around shared data—termed co-reflection—is a key component of family informatics. Co-reflection refers to a collaborative data reflection process in which family members assess their behaviors at both individual and collective levels through discussion [68]. A core benefit of co-reflection is that it allows family members to develop a deeper understanding of themselves by engaging with each other's reflections, aligning with the broader notion of shared reflection [40]. Collaboratively reflecting on shared experiences can also help strengthen children's agency and critical thinking skills since it moves beyond one-way, parent-to-child instruction toward more reciprocal and dialogic interaction.

Lee et al. [68] demonstrated the feasibility of using passive sensing to support family co-reflection, showing data visualization on affective and behavioral states during family interactions can foster deeper self- and mutual understanding. However, existing family informatics research has primarily focused on capturing individual or aggregated health data within families, with less emphasis on designing systems that effectively guide families to interpret, reflect on, and assess their social and emotional interactions. Our work expands this line of research by investigating how richer forms of interaction data (e.g., conversational dynamics, emotional synchrony) captured through multimodal sensing shape the family reflection and also promote social and emotional learning.

2.2 Observation of Social-Emotional Behaviors in Family Conversations

2.2.1 Conventional approaches for family conversation assessment. Analyzing family conversations has supported the assessment and improvement of social and emotional competencies, particularly children's self-awareness, emotion regulation, and communication skills. Several therapeutic approaches have been designed for this purpose, such as psychological counseling [91, 97].

One widely used approach involves families visiting counseling centers for a professional evaluation. A child or a whole family follows a structured protocol, and experts observe their behavior or review video recordings of their interactions to assess various social and emotional behaviors [91]. Common procedures include the *Conversation Probe (CP)* for evaluating conversational abilities [79] and the *Social Performance Rating Scale (SPRS)* for assessing social behavior across different situations [35].

Not only for the assessment, but to provide feedback to families, their recorded conversations were leveraged. Family members are having a daily conversation while being recorded, and trained experts later select specific video segments during therapy sessions to examine affective communication in shared playful moments [31]. This practice systematically analyzes interactions among the father, mother, and child to identify behaviors such as affective sharing, exclusion, interference, and withdrawal. These moments are reviewed in short episodes under the therapist's guidance.

Video-based approaches have been used for decades [91] to identify verbal, nonverbal, and multimodal behaviors in family interactions and to support the development of social and emotional skills. It has been researched that video feedback could effectively enhance parenting practices and child outcomes [34]. However, they rely on professionals with specialized training and on tools for detailed behavioral coding, creating barriers for families who lack access to such resources.

2.2.2 Sensing technology for parent-child conversation. HCI researchers have also explored how technology can support family communication through sensing technology. These efforts reflect the broader view that daily conversations are central to children's social development and serve as a key setting for sharing experiences, coordinating routines, and learning norms [18].

Context-specific sensing of the family conversation has been used to provide a quantitative assessment and feedback. TalkBetter provides real-time meta-linguistic feedback to parents interacting with children with language delays, supporting goals defined by speech-language pathologists [48]. SpecialTime assists *Parent-Child Interaction Therapy* by automatically detecting dialogue acts and offering immediate feedback aligned with therapists' coding practices [46]. You et al. emphasized empathetic reflection during conflicts by visualizing how parents' nonverbal cues might be perceived by the child [110]. MAMAS monitors parent-child mealtime interactions and helps parents reflect on eating behaviors and communication patterns [51].

Yet, these sensor-based approaches have largely been confined to providing one-way instruction, where children are primarily monitored while parents take on the role of instructors. This asymmetry in participation has been noted as a potential concern in prior research, particularly regarding its impact on children's autonomy and privacy [83, 84]. Moreover, family conversation content remains underexplored, with more focus on eating habits or eating behaviors of family members.

Given the importance of daily family interactions and the potential of sensing family activities, family informatics tools present a promising approach, as their equal accessibility to all family members can shift the focus from passively following guidance to actively engaging in reflective learning of their social and emotional behaviors.

2.3 Sensor-driven Reflection on Social and Emotional Data

There has been growing interest in leveraging data-driven approaches to capture and reflect on individuals' social and emotional behaviors. Advances in sensing technologies (e.g., video, audio, and physiological data) have enabled data-driven reflection by quantifying social behaviors and visualizing patterns of interaction by detecting emotional states, such as stress.

The quantification of social behaviors provides insights into how they are perceived by others when they are interacting with them [7, 95]. Moreover, this ability to capture social aspects makes it possible to provide automated, instant feedback. Consequently, sensing has been applied to enhance social-emotional competencies, such as self-awareness, social awareness, and relationship skills [55, 56, 59]. Some systems assisted with oral presentation [14, 106], executive coaching [5], and job interview practice [44] by analyzing speech quality (e.g., pitch variety and speaking rate) and facial expression through camera sensor and providing coaching or visual/haptic feedback as users engaged in simulated presentations or mock interviews with the system. Therefore, passive sensing for social behaviors has been applied across various contexts targeting diverse populations, including interventions for individuals on the autism spectrum [42, 59, 61, 66], classroom environments to promote effective instruction [3], and family settings to analyze and support parent-child social interactions [19, 63]. Such behavioral sensing has been conducted not only in individual use cases but also in co-located settings, where attempts have been made to capture the dynamics of multi-person interactions such as conversations. For example, conversational dynamics (e.g., participation, current turn of conversation) were visualized to enhance group collaboration [9, 17, 24, 62] and encourage self-regulation of participation [1, 102]. However, in these co-located settings, relatively few studies have focused on capturing more internal states (e.g., emotions) beyond the behavioral level.

The application of combined social-emotional data from interpersonal contexts to facilitate *Technology-Supported Reflection* thus remains in its preliminary stage. While existing systems providing interventions tend to focus on improving behaviors for well-defined task goals through nudging or persuasion, *Technology-Supported Reflection* supports learning by helping individuals revisit past social interactions and derive insights for change [41]. Considering emotional as well as social cues enriches reflection, as social interaction depends on both external signals and internal states (e.g., emotions). As an example, group-based sensing of mood and stress in workplace settings [2, 37, 77, 111] have been developed and studied to support well-being and workplace learning by improving stress awareness. In such cases, collective reflective practice leverages not only the rich social context collected from advanced sensing technologies (e.g., conversational dynamics, emotional synchrony) but also the human ability to interpret nuanced information from the context. Leveraging the concept of human-as-sensors [36, 99], sensing technology has the potential to complement human memory and further facilitate the understanding of interpersonal contexts.

While prior work on group sensing in workplaces has focused on loosely connected individuals, close relationships (e.g., families, friends, and couples) offer richer opportunities for emotional insight due to shared experiences and deeply entangled social dynamics [49]. However, the application of sensor-aided systems in intimate contexts remains limited. Building on this gap, our system adapts social-emotional sensing in the family context, demonstrating sensor-aided family informatics using social and emotional data of their family interactions.

3 System Design

To observe how families reflect on their interpersonal behaviors, we designed an early-stage system, SELaD (**S**ocial **E**motional **L**earning by multimodal **D**ata) using video, audio, and physiological sensors. In Section 3.1, we first outline the overall design process of SELaD based on our formative study. Section 3.2 then describes the SELaD system design in more detail.

3.1 Design Process Based on a Formative Study

The goal of the formative study was to identify the data types that can capture social-emotional behaviors during family conversations and develop the interface of family informatics. To achieve this, we conducted a literature review to select the data type (Section 3.1.1) and expert interviews on both the data types and the early system design (Section 3.1.2).

3.1.1 Data Type Selection. We conducted a literature review of existing conversation assessment methodologies. Based on this review, we selected the Conversation Probe (CP) framework [79, 80, 89] as the foundational architecture for the SELaD system. The behavior indicators of CP are broadly categorized into four composite groups: *appropriate content, paralinguistic behaviors, interactive behaviors, and nonverbal behaviors* as shown in Table 1. Each category comprises a set of coded behaviors, such as the level of "involvement" and the level of "asks questions" for the "Interactive Behaviors" section. In SELaD, we mapped the sensor data types with these coded behaviors, such as "number of questions" for "asks questions".

We chose CP over other scales for several reasons. First, CP provides low-level behavior indicators that can be measured using audio and video analysis and social signal processing technologies. For instance, CP's item "asks questions" can be easily captured via sensors, unlike Social Performance Rating Scale (SPRS) [35] items that assess vocal quality, such as "Participant demonstrates no warmth, enthusiasm, or interest in verbal expression [35]," which require more complex interpretation of nuanced social situations. Moreover, CP has the advantage of not being limited to verbal content, compared to alternative interaction assessment systems that are applicable for passive sensing [46]. Therefore, CP allows our system to capture a broader range of skills than verbal content, such as recognizing one's own emotions.

We sought to validate these data type selections as summarized in Table 1, through expert interviews, which we describe in the following subsection.

3.1.2 Expert Interview. To gather information on the feasibility of the system and its expected utility, we conducted expert interviews (n = 4). The participants were researchers and practitioners specializing in family and children's social-emotional development, including a professor and a Ph.D. candidate from a child development and family studies department, a family therapy expert, and a child counselor working with our target age group. All participants were based in Korea. This composition was intended to comprehensively cover both academic and applied perspectives. Each semi-structured remote interview lasted approximately one hour. The questions were broadly categorized into two: *(1) What family interactions should we focus on to capture their social-emotional data and provide reflection on, (2) Feedback and design suggestions for SELaD.* The prototype depicted in Fig. 1 was provided to the experts as a visual aid to envision and suggest potential design improvements. Participants received compensation of 80,000 KRW (55 USD), and the study was conducted with IRB approval. Researchers performed thematic coding [13] and derived themes from the interview response transcriptions for each question category. Common themes mentioned by interviewees are summarized as follows:

As an answer to the first question, 3 out of 4 interviewees mentioned that family mealtime conversation is one of the most fundamental contexts for family interactions and, therefore, needs to be examined. They highlighted that social-emotional aspects are accumulated through interactions involving language. Given that family mealtime conversations are often used in research to naturally assess children's social and emotional competencies [6, 30], we planned to conduct our study in a mealtime setting as well.

After reviewing our selected data types and the early prototype (right of Fig. 1), experts suggested incorporating additional forms of emotional data. They recommended identifying emotions (e.g., happiness, anger) from conversation content, noting that the ability to choose emotion-aligned words during dialogue is an important indicator of children's social development. They also noted that physiological data (e.g., stress signals, muscle tension, brainwaves) can offer families meaningful insights to help them become more aware of feelings not openly expressed in words or facial expressions during conversations. However, experts also highlighted that interpreting physiological data can be ambiguous (e.g., whether emotional arousal stems from joy or anger) and literacy can vary among people, requiring additional explanation in the experiment. Reflecting this feedback, and informed by prior work in affective computing that highlights the value of a comprehensive understanding of affective states [104], we expanded our data types accordingly.

3.2 SELaD Design

We now illustrate the overall system architecture and detailed interface design of SELaD, providing both a high-level overview of data sections and a detailed explanation of data section interface design.

3.2.1 Overview of Multimodal Data Sections. SELaD was designed as depicted in Fig. 2 based on the findings of the formative study. For each conversational topic, the interface displays a **chart-aligned**

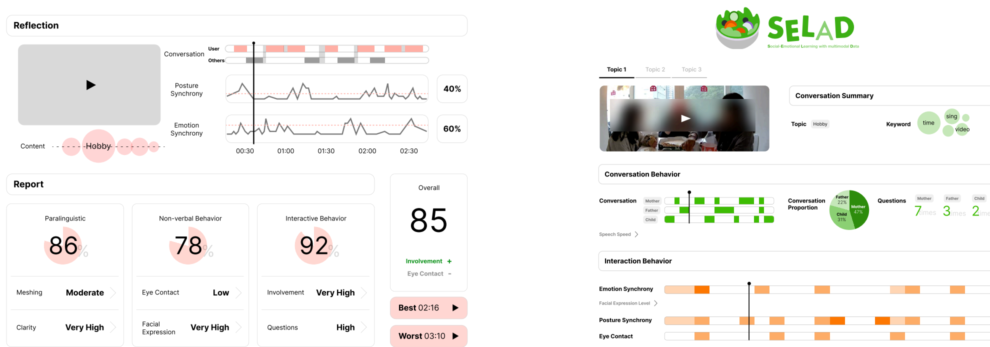


Figure 1: Early prototypes of SELaD, designed based on the selected features from the literature review. The left prototype represents an individual view that provides a summarized report intended to help users understand whether they performed well or poorly, thereby promoting a learning effect. However, due to difficulties in constructing objective assessment criteria, the design process shifted toward a more transparent presentation of raw multimodal data, as illustrated in the right-side prototype. The right prototype was subsequently used during expert interviews.

video view (S1), along with four *multimodal data sections* – **S2) Conversation Summary, S3) Conversation Behaviors, S4) Interaction Behaviors, and S5) Physiological Responses**. In Fig. 2, different sources of data are distinguished by color (i.e., S2, S3 – audio: green, S4 – video: orange, S5 – physiological data: purple).

Chart-Aligned Video. (S1) in Fig. 2 is the original video source alongside the conversation analysis data. It was placed on top of the screen to provide rich contextual information on family activities, used as video feedback [97]. This was designed to reflect experts' opinions that videos paired with data can facilitate reflection by revisiting memories and emotions from that moment. Additionally, to help family users easily navigate and correlate information from different sources, we synchronized the video with other multimodal data — such as conversation behaviors, interaction behaviors, and physiological responses — using the same timeline.

Conversation Summary. (S2) in Fig. 2 focuses on the verbal content of the conversation. To support reflection on whether their conversation aligned with the given topic, this section displays the topic of the session and highlights keywords spoken by each family member. Three representative keywords per participant are extracted and visualized. The proportion of positive words used is also shown as a percentage, providing insight into the emotional tone of the conversation. By providing these data types, Conversation Summary helps participants reflect on their social communication skills and emotional tone.

Conversation Behaviors. (S3) in Fig. 2 presents data capturing paralinguistic and interactive behaviors during the conversation using audio data. For paralinguistic aspects, a conversation timeline shows who is speaking over time, and a speaking initiation count indicates how often each member spoke. These visualizations help users assess individual and group conversational dynamics.

Table 1: Selected data types for SELaD mapped with Conversation Probe (CP)

Selected Behavior Indicators in CP		Selected Data Types for SELaD	
Category	**Behavior**	**Data Type**	**Section**
Appropriate Content	Verbal Content	Keywords	Conversation Summary
		Positive words proportion	
Paralinguistic Behaviors	Clarity	Speaking speed	Conversation Behaviors
	Fluency		
	Meshing	Conversation visualization	
		Speaking initiation	
Interactive Behaviors	Involvement	Conversation proportion	
	Asks questions	Number of questions	
Nonverbal Behaviors	Gaze	Eye contact	Interaction Behaviors
	Flat affect	Emotional expression	
	Appropriate affect	Emotion synchrony	
-		Stress	Physiological Responses
		Emotional arousal	

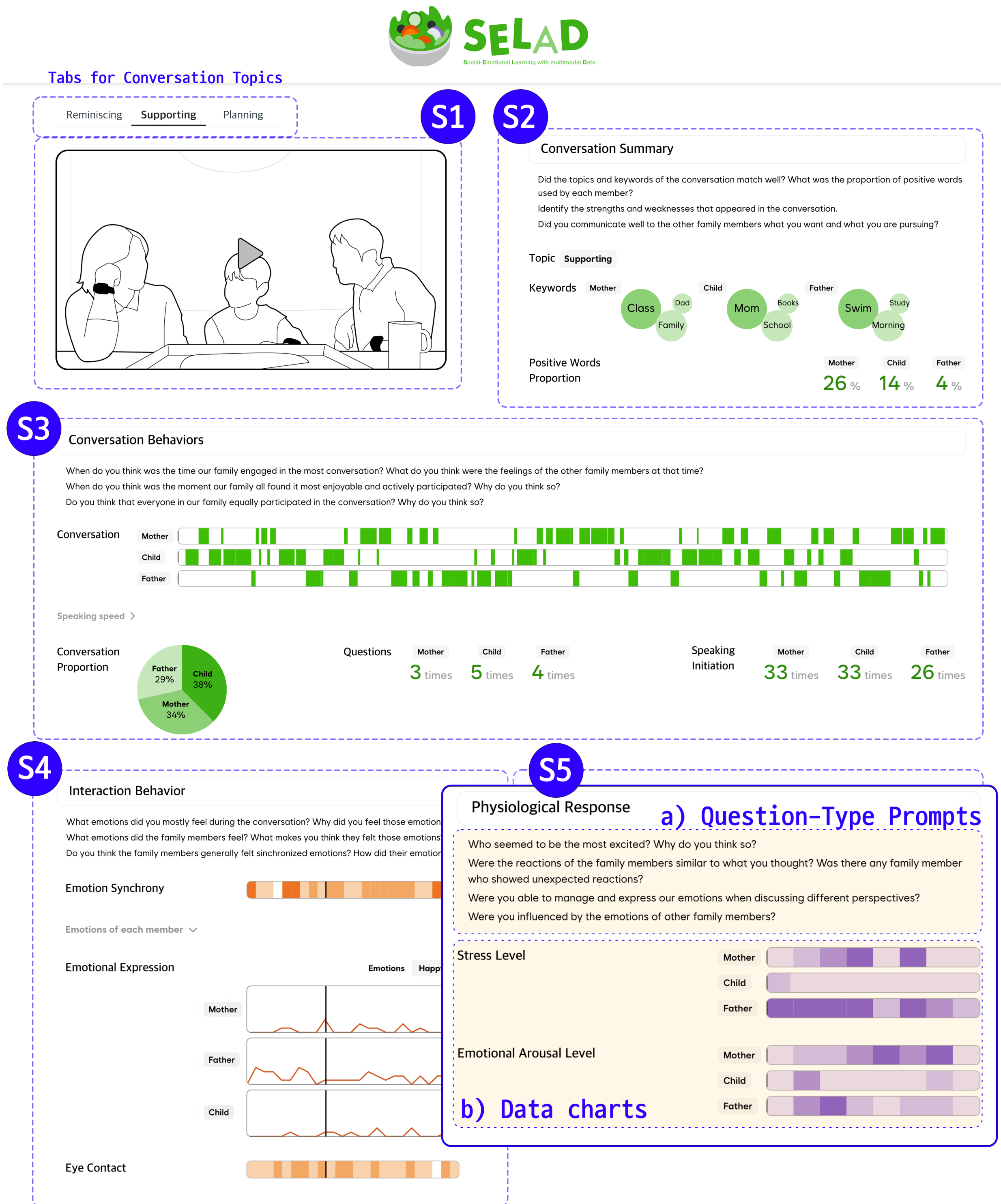


Figure 2: A screenshot of the visualization of SELaD. S1 is a chart-aligned video view of their conversation, S2-5 indicate different data sections, and each data section comprises a) Question-type Prompts and b) Data Charts.

Speaking speed is visualized through the height of each segment on the Y-axis, offering cues about fluency. To reflect CP's coded behaviors such as *involvement* and *asking questions*, a pie chart illustrates each member's conversation proportion, highlighting involvement balance and dominance within the family. The number of questions each member asks is also displayed to indicate the level of interactivity. These data types could capture social dynamics (e.g., dominance, responsiveness), situating them as social data.

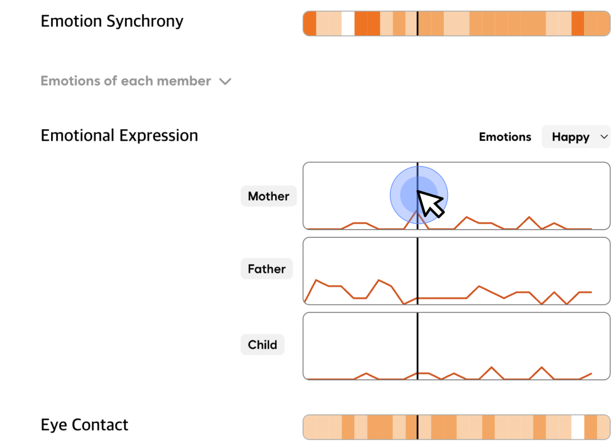
Interaction Behaviors. (S4) in Fig. 2 highlights nonverbal behaviors such as *gaze* and *appropriate affect* based on video data. To detect gaze, we analyzed whether family members were looking at one another during the conversation. We also included emotional synchrony to examine whether family members displayed appropriate affect in response to shared moments. Additionally, each member's facial expressions are visualized over time, allowing users to observe whether their expressions remained flat or showed dynamic changes. Compared to the previous sections, these data types more directly present the emotional information.

Physiological Responses. (S5) in Fig. 2 offers additional indicators beyond the CP-based indicators by detecting stress levels and emotional arousal using physiological sensors. Measures such as stress level and emotional arousal allow families to reflect on internal affective states that may not be expressed verbally or behaviorally, offering opportunities for deeper awareness and reflection on emotions that might otherwise remain hidden. As such, these signals primarily represent emotional data.

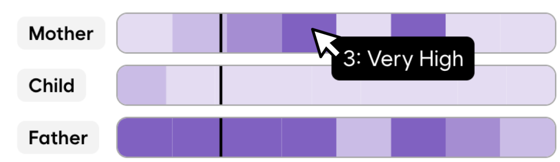
3.2.2 Detailed Data Section Interface Design. Every data section of SELaD is composed of a) Question-type Prompts and b) Data charts, as shown in the example of S5 in Fig. 2.

Question-type Prompts. To help the family better understand their communication based on the data, experts recommended adding guidance that can encourage them to reflect on their behavior. Therefore, we integrated question-type prompts (Fig. 2–a) throughout the interface in each section to guide and encourage family reflection without on-site expert intervention. For example, *"Do you think that everyone in our family equally participated in the conversation? Why do you think so?"* was used in *Conversation Behaviors* (see S2 in Fig. 2). The prompt questions were adapted from a list of exemplary social-emotional learning skills in prior research [98] to align with the data types that SELaD quantifies (see Appendix A).

Data Charts. Overall, data in SELaD are presented in time-series charts (Fig. 2–b) to facilitate understanding through alignment of the video and other data. In the time-series data charts, a vertical line is drawn over the graph to represent alignment with the video timeline (Fig. 3a). When a user clicks on a point from a chart, the video is played at that synchronized time stamp, allowing users to explore the specific moment in the video. Additionally, ordinal data such as stress, emotional arousal, emotion synchrony, and eye contact are presented using level-based charts (Fig. 3b). The level-based charts represent varying ordinal levels through color differences, and the description of the level (e.g., low, moderate) appears around the mouse when hovered over (Fig. 3b). This color saturation mapping technique was adopted based on prior research [74], a technique effective in representing sequentially ordered data.



(a) When user clicks time-series charts, timeline for charts and video aligned.



(b) When user hovers, the level indicator appeared.

Figure 3: Detailed interactive design components

4 Implementation

Based on the system design, we implemented SELaD to extract features related to social-emotional development from video, audio, and physiological sensor data collected during family conversations. These features are visualized alongside the recorded videos to support reflective exploration. Video and audio were captured from a PC-connected camera and microphone, respectively, and processed according to the data analysis pipeline as in Fig. 4. Physiological data were collected via Empatica's E4 wristband and accessed via the Empatica Connect platform for data analysis.

SELaD was developed as a web-based system with server-side and client-side components. On the server side, we used Python 3.9 for data analysis, executed within Docker containers, and employed Celery for asynchronous task management and batch processing. MongoDB handled user authentication and data storage, while Redis was used for caching. On the client side, we built the web interface described in Fig. 2 using Next.js and React.js, and implemented the interactive visualization components using d3.js.

The overall data analysis pipeline is illustrated in Fig. 4. The tools and libraries used for processing each data type are detailed in Table 2. For brevity, we present only an overview here; the full, detailed data-processing procedures for each data type are provided in the Appendix C for full reproducibility.

Audio. SELaD first applies speaker diarization to segment the conversation into speaker-specific utterances. For each segment, the

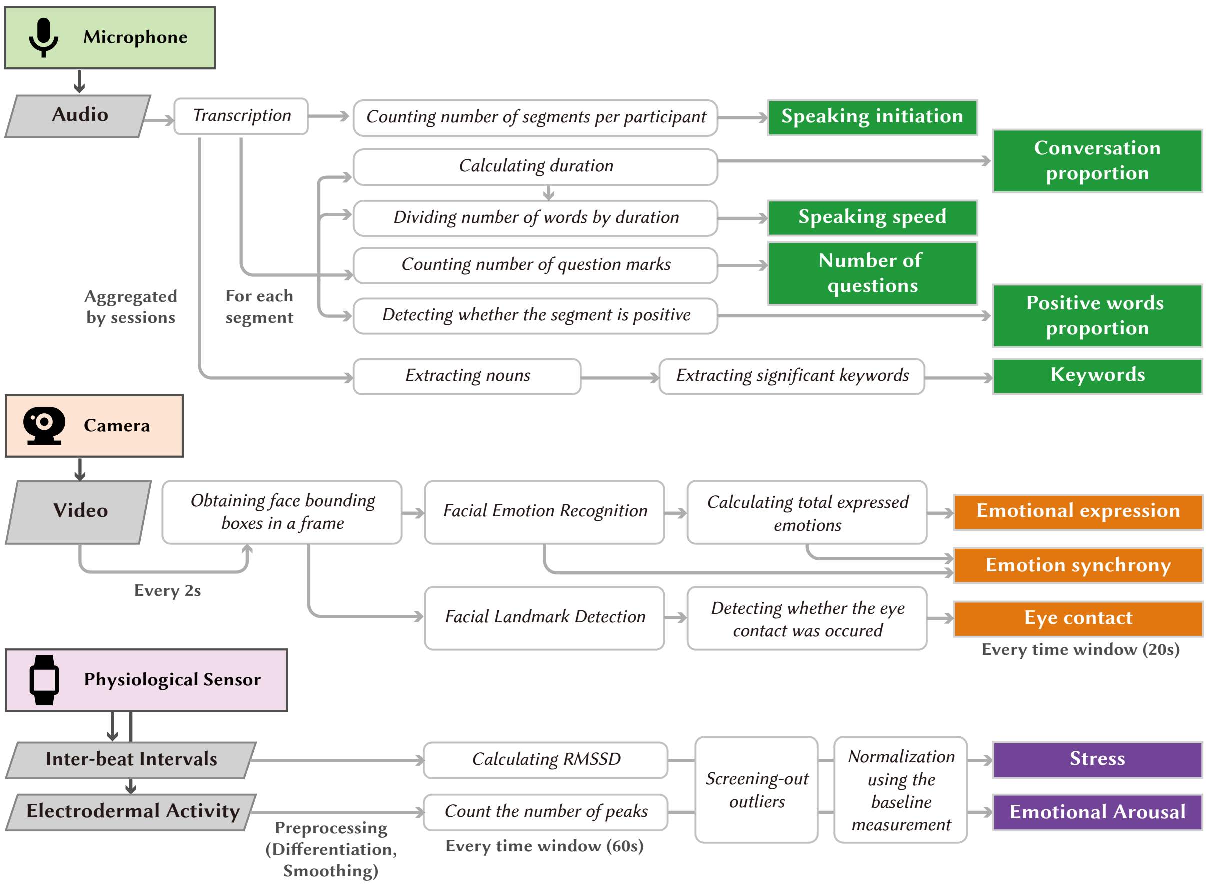


Figure 4: A flow diagram of the data processing pipeline of SELaD. The different sources of data are distinguished by color, aligning with the SELaD design (i.e., audio – green, video – orange, physiological data – purple).

system computes three features: *speaking speed* (words per second), the *number of questions* asked, and the *proportion of positive words*. *Conversation proportion* of each speaker is derived from the duration of segments, and the number of *speaking initiation* is the count of segments each speaker begins. Finally, it extracts significant *keywords* by aggregating term frequencies across the entire session.

Table 2: Implementation details used in the SELaD system.

Feature	Tool
Emotion Recognition	A CNN based PyTorch implementation on facial expression recognition [50]
Facial Landmarks	The FaceMesh model from Google's MediaPipe [73]
Speech Transcription	Naver Clova Speech AI [81]
Positive Words Proportion	Naver Clova Semantic Analysis AI
Transcript Tokenization	The Okt tokenizer from konlpy [85]
Keywords Extraction	KRWordRank [58]

Video. For every key frame, SELaD runs face bounding boxes detection, facial emotion recognition, and facial landmark extractions. Within a 20-second time window, the system computes features including *total expressed emotions*, *level of emotional synchrony* – the similarity between the emotions of participants –, and *level of eye contact*. To calculate eye contact with a single camera – suitable for in-home use – we assumed a fixed-angle setup where participants are seated side-by-side. All video analyses follow the methods established by prior studies [38].

Physiological Signals. SELaD derives *stress levels* from inter-beat intervals (IBIs) [15, 57] and infers *emotional arousal levels* from electrodermal activity (EDA) [8, 12, 21]. Following Schmidt et al. [96], the system first preprocesses each signal using differentiation and smoothing. For stress estimation, heart rate variability (HRV) is calculated from the IBI signal, using the RMSSD (root mean square of successive differences) metric within 60-second windows. For arousal estimation, the number of EDA peaks is extracted from the EDA signal within the same time windows. Finally, all feature values are normalized against a baseline measurement, outliers are removed, and the resulting metrics are level-scaled for visualization.

5 User Study

To explore how families reflect on the family conversation by interpreting family informatics of multimodal social and emotional data, we conducted a user study with 17 families, each consisting of three members ($n = 51$). The study method was inspired by the technology probe [47], which emphasizes deploying a prototype in real-world contexts to provoke insights about potential use. We adopted its exploratory orientation to investigate the situated use of SELaD, identifying opportunities and challenges in supporting family reflection through multimodal data.

5.1 Participants

We recruited 17 families consisting of three family members, i.e., a mother, a father, and a child ($n = 51$). While families with more than three members were eligible to participate, only three members per family took part in the study due to the space constraints of the testbed environment and to facilitate standardized data analysis. Based on the advice from preliminary expert interviews, we targeted families with children between the ages of 10 and 14, as this period represents *early adolescence* – a developmental stage characterized by significant transformations in parent-child relationships and development of the self [101]. Experts indicated that children in this age range would be able to comprehend and interpret the data and participate meaningfully in conversations, and because this developmental period is marked by rapid changes in social and emotional skills that make everyday family interactions informative for reflection. Families were recruited through a local online community, and all families were Korean. Additional demographics are shown in Table 3. Each family received 100,000 KRW (70 USD) as compensation for their participation in the study. The study was approved by the university's Institutional Review Board (IRB).

Table 3: Demographic information of participants (G: Girl, B: Boy)

Family ID	Father	Mother	Child
1	50	46	12 (G)
2	50	50	12 (G)
3	48	44	12 (B)
4	48	44	11 (G)
5	41	39	12 (B)
6	52	46	13 (B)
7	45	44	13 (B)
8	41	42	12 (B)
9	54	53	12 (B)
10	54	48	11 (G)
11	43	40	13 (B)
12	43	41	12 (B)
13	49	49	11 (G)
14	43	41	11 (B)
15	46	44	11 (B)
16	48	44	10 (G)
17	43	44	12 (B)

Figure 5: A testbed that resembles a home environment. The overall study was conducted in the dining room.

5.2 Procedure

5.2.1 Apparatus. The participants were invited to a testbed environment (Fig. 5) designed to resemble a house. The study was conducted in the dining room, and researchers monitored the data collection status in a separate room to allow families to freely converse. To create a natural setting for family conversations and facilitate the conversation, a meal (pizza) was provided to the participants during the data collection.

For data collection, a PC-connected camera and a MAONO BM10 USB conference computer microphone were used to capture audio and video data from the family conversations. In addition, participants wore Empatica's E4 wristbands on their non-dominant hands to capture physiological signals such as EDA and HRV. Before arriving on the testbed, participants were asked to avoid caffeine consumption on the day of the experiment due to the accuracy of HRV detection, as caffeine could cause a sudden increase in sympathetic nerve activity [22].

5.2.2 Study Process. The overall process of the user study is illustrated in Fig. 6. First, participants were briefed on the study's objectives and procedures. After signing a consent form for study participation, they were fitted with Empatica's E4 wristbands for physiological data collection.

Before the data collection, the baseline physiological data were collected in a five-minute guided meditation in a relaxed state while participants were wearing E4 wristbands. Subsequently, participants engaged in the main data collection phase with conversations on three topics, each lasting 8 to 10 minutes. Since families may find it difficult to initiate a completely open-ended conversation in the lab setting, we selected structured familiar conversation topics. These topics were selected based on formative expert interviews and existing literature about family mealtime conversation topics [30], to capture a range of social-emotional behaviors within family dynamics while mitigating the ethical risk of triggering emotionally charged or uncomfortable conversations for children, such as rebellious teenager behaviors or serious arguments among family members.

(i) *Reminiscing*. Families were encouraged to use the photos of family memories they had brought, guided by our instruction,

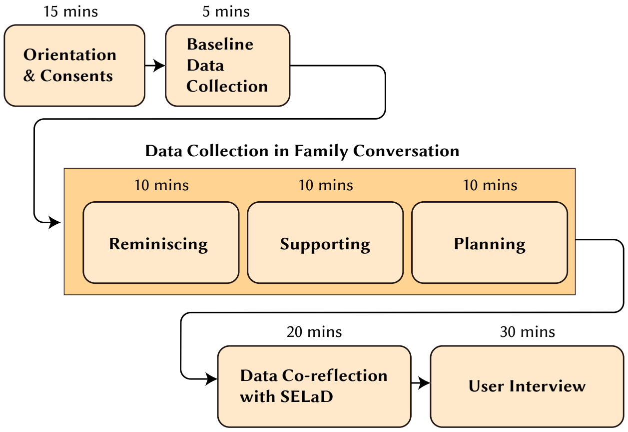


Figure 6: User study procedure.

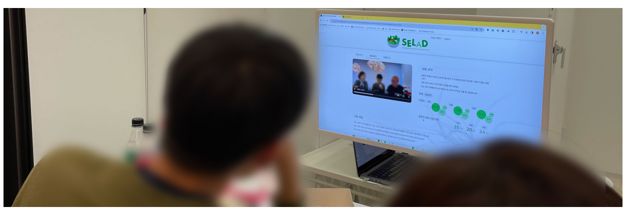


Figure 7: Participants collaboratively reflect their data using SELaD.

such as *"Let's recall family memories together."*, *"Share your favorite moment."*, and *"Use the photos to help you remember the past events and your emotions at that time."*

(ii) *Supporting*. Families were asked to share their recent experiences, encompassing both challenges and celebrations, and were further encouraged to discuss how they could support or be supported by each other for such experiences.

(iii) *Planning*. Families were tasked with planning a family trip. They were also provided with starter questions such as *"Where would you like to go for a family trip?"*, *"What activities would you like to do during the trip?"*, and *"How will you get there, and where will we stay?"*

At the beginning of the data co-reflection phase, we briefly introduced the SELaD interface with types of data and how to use SELaD. Then family members freely interacted with SELaD, engaging in discussions about their social-emotional behaviors with multimodal data, as depicted in Fig. 7. Families were encouraged to explore the data freely rather than being provided with specific directions for reflection.

Following the data co-reflection phase, we conducted a semi-structured interview with each family. The interview questions were designed to understand the collaborative reflection experiences of families using SELaD (See Appendix B). As all family members participated in the interview at the same time, we gathered both individual perspectives from each member and any consensus reached during the discussion of each interview question.

5.3 Data Analysis

Our analysis using thematic analysis [13] focused on the families' reflection pattern, the use of various data types, the insights the family gained from the reflection, and the factors that contributed to the reflection. The entire reflection dialogue of a family lasted around 20 minutes, and the post-interviews about 30 minutes were recorded and analyzed. Screen recordings of user interactions with SELaD were also utilized to provide a contextual understanding of the reflection process. We first transcribed all family conversations from the data reflection phase and identified "reflection instances" where instances are delineated based on conversational boundaries, such as pauses longer than three seconds, shifts in focus to a different time point in the data, or clear topic transitions. Families differed in how actively they reflected during the reflection session. On average, families engaged in 19.24 reflection instances (SD = 8.49). The family with the fewest instances (Family 6) showed 9 instances of reflection, whereas the family with the most (Family 10) showed 42 instances. In total, we identified 327 reflection instances.

Table 4 summarizes how frequently each data type was referenced during family reflections. For example, during a single reflection instance, a family noticed a high point in the stress level graph and watched the corresponding video clip to understand what happened, and finally compared it with the facial expression data. Then, the data types used in that instance were coded as "Stress" and "Facial expression." Families differed widely in which data types they attended to and how actively they used them. For example, the "Keywords" data was used by 11 families, and 6 families did not refer to it for their reflection. Across all sessions of all families, keywords were referenced 23 times, indicating that some families used this data repeatedly during the reflection. This corresponds to an average frequency of 2.09 uses per family.

Subsequently, we revisited each reflection instance and coded the types of insights families derived from the reflection, the types of relational cues they used, and whether a deeper understanding

Table 4: Usage metrics for each data type in reflection: total number of instances, number of families (percentage of all families), and average usage frequency per family.

Data Type	# Instances	# Families	Freq
Keywords	23	11 (64.7%)	2.09
Positive words proportion	27	12 (70.6%)	2.08
Conversation visualization	35	16 (94.1%)	2.19
Speaking speed	4	2 (11.8%)	2.00
Speaking initiation	5	5 (29.4%)	1.00
Conversation proportion	26	11 (64.7%)	2.27
Number of questions	13	8 (47.1%)	1.62
Eye contact	15	9 (52.9%)	1.67
Facial expression	81	16 (94.1%)	5.06
Facial synchrony	9	7 (41.2%)	1.29
Stress	70	16 (94.1%)	4.38
Emotional arousal	25	13 (76.5%)	1.92

or discovery about their family was observed. The interview answers about their insights into the families' social and emotional behaviors provided contextual grounding for this coding process. Two researchers separately read and coded approximately 50% of the transcripts. After assigning initial codes, they reviewed the codes, discussed discrepancies, merged similar codes, and prioritized themes based on their frequency of occurrence. They then reached a consensus on the final coding results. In the following result sections, we will present key themes regarding how families engaged with these data types to reflect on their conversations and social-emotional experiences.

6 Findings

In our study, we observed how **family members interpret their own and others' social and emotional behaviors, and share the interpretation in social or familial contexts**. We call this practice *Relational Reflection.* This concept builds on the idea of *dialogic reflection*, which has been defined as "looking for relationships between pieces of experience and knowledge during cycles of interpretation and questioning [32]." *Relational reflection* goes beyond this definition by involving not only connections across one's own experiences and data but also the use of **shared histories and shared expectations on family interactions** to interpret data within a social or familial setting collaboratively. Below, based on the results of the qualitative analysis of reflection sessions and the responses from the interviews, in Section 6.1, we first show how families gain insights about their speech patterns and emotional influences from the relational reflection. Section 6.2 then highlights how this relational reflection could be facilitated from our observation.

For clarity, we denote each family member with the following abbreviations: F for father, M for mother, and C for child, with the family ID number (e.g., M01 represents the mother of the first family). Also, when family members refer to one another by name, we denote them as [C01], for example.

6.1 Co-constructed Insights about Social-Emotional Dynamics through Relational Reflection

We observed that families used SELaD to examine their social-emotional dynamics, such as how they talk to each other, who dominated the conversation, and how their talk was shaped by others' influence. Through relational reflection using SELaD, families collaboratively developed these insights into their group dynamics from their previous conversation, with the use of shared histories and shared expectations of family interactions.

In our study, families gained two types of insights: conversational patterns and emotional influence. In the following sections, we will illustrate how the key elements of relational reflection – **shared histories, shared expectations on family interactions** – enabled families to develop these insights.

6.1.1 *Recognizing and Evaluating Families' Conversational Patterns.*

One notable finding was that families used their existing shared knowledge about how they usually talk, along with their expectations of balanced family interaction, to recognize and evaluate conversation patterns surfaced through SELaD. These included identifying the dominant speaker in conversation, assessing the family's turn-taking practices, and noticing positive language usage. Since the conversational patterns are widely recognized as indicators of communication quality, families focused on the speech data components (i.e., Conversation Summary and Conversation Behaviors) to interpret their conversational patterns.

To interpret speech data, families frequently recalled their shared history, such as other family members' **habits, preferences, and memories they've experienced together**. Using shared family history in reflection added relational context, which facilitates families to feel more personally engaged in the reflection. For instance, M12 pointed out her husband's conversational habit from the keywords data, *"I thought this was hilarious. (pointing to the keyword 'anyway' from the conversation summary section) He says the word 'anyway' so often."* This shows how M12 used her prior knowledge of her husband's frequent word choice to make sense of the keyword data. Similarly, M04 commented on her low proportion of positive words, recalling her existing speaking habit: *"To be honest, I'm definitely not a positive person. I know that I often push kids to the extreme and sometimes use negative language. But I didn't realize I was also doing so in the previous conversation. Maybe my default condition is too negative, and that's why the positive words proportion data shows that I did not use positive words."* By drawing on her memory of using negative language toward other family members, she could reconfirm that she did not use positive words.

Additionally, families often referred to their expectations of **balanced family conversation**, which led them into the relational reflection about their balance of conversation – when they identified a dominant speaker or noticed an imbalance, they connected it to these expectations with self-evaluation. The dominant speaker attempted to understand their own turn-taking practices by examining the data types of conversation proportion, speaking initiation, and number of questions. When one family member dominated the conversation, families noted an imbalance and questioned it. For example, M12, noticing her own high proportion of speaking time, remarked, *"Oh, look at this – I'm dominating the whole conversation. Is it desirable?"* She believed that balanced conversation is desirable; however, the data challenged her expectation of balance, prompting reflection while questioning both herself and other family members about whether this distribution was desirable. Families also tried to understand the family's turn-taking practices by looking at the conversation visualization data. For instance, F03 identified and evaluated that their family's turn-taking dynamic was good: *"When I look at the conversation between the three of us, dad, mom, and child, it seems like we're constantly taking turns back and forth. I talk, then mom talks, and the child talks in the middle. It's like a toss over another toss. I think that's how we talked well."* In this case, F03 valued back-and-forth turn-taking as a desirable way of communicating, and because their data followed this pattern, he concluded that they had talked well.

Beyond recognition and evaluation of their own conversation, the relational reflection on conversational patterns often led families to feel a need for positive changes, such as practicing listening more or being more responsive to other family members. For instance, M09, who was the dominant speaker in the family, reflected on her dominance in conversation and noted the need to practice listening more: *"I need to talk less. Yes, I already knew I talked too much. However, looking at the conversation proportion, I realized I really need to look back on myself. I think I need to practice listening more."* M09's reflection was also grounded in her expectation for balanced conversation, which motivated her to commit to self-improvement by speaking less and listening more.

6.1.2 Discovering Emotional Influence within Families. A key finding was that families leveraged their prior knowledge of how they usually reveal their emotions, connecting with their expectations about emotional synchrony and the appropriate expression, focusing on their emotional dynamics using SELaD. Since expressing appropriate emotions and identifying and responding to each other's feelings is an important nonverbal skill of empathetic conversation, families paid attention to these emotional influences while interpreting the emotional data, such as facial expressions and stress levels. Prior research in family informatics showed that families want to understand the *ripple effects* across family members, how their behaviors and moods affect each other [88]. Our system supported these needs by enabling families to contextualize individual emotions within the family setting and the discovery of emotional influences often fostered empathy, feelings of connection, and perspective-taking.

We observed that families' expectations of good family communication involved **emotions being well synchronized**. When this expectation was connected with their **shared history about emotional tendency**, they could reveal patterns they had not expected or previously recognized. For example, the father, F04, a person who displayed fewer expressive facial cues, however, his physiological data exhibited fluctuations in emotional arousal level. This led the family to interpret that he actively engaged in the conversation even if he did not outwardly express. M04 highlights this discovery: *"I'm the one talking the most. It means there must be others in the family who don't talk much, such as a teenager in adolescence or a blunt husband. I might get to know them over time, but it's still hard to understand their feelings and thoughts if they don't talk much. But seeing SELaD, I can see that [F04] is also experiencing emotional fluctuations or emotional synchronization when I'm feeling sad. It makes me realize that this person is not a robot, and it can actually be comforting."* In this example, the family drew on their shared history of F04 being less expressive, and combined it with the expectation that emotions should be synchronized and shared within the family. Although F04 rarely showed his emotions outwardly, the emotional data revealed that he was resonating with M04's sadness, in fact. This understanding allowed M04 to find comfort in recognizing their emotional connection, which was difficult to explicitly express in words.

We also observed families' expectations about a **positive influence on one another**. When their data contradicted this expectation, they often expressed surprise and sometimes questioned whether they had caused a negative effect. For instance, M12 was interpreting C12's peak of stress level with the conversation visualization and realized that she was the one speaking when the child's stress was high. She was surprised and remarked, *"Why [C12] is stressed... Is it when I'm talking? Because of me?"* Then, upon checking the video, both she and the child became engaged in identifying the cause together. Another example involved participants reflecting on their negative facial expressions and, in doing so, recognizing criticisms from others that they had previously dismissed. Seeing themselves in the data helped them to take these remarks more seriously and to acknowledge the impact of their expressions. For example, M06, upon seeing her emotion data and facial expressions toward the child in the video, realized the impact of her negative manner and reflected: *"Do I usually talk with this face?"* and C06 responded *"Usually, your angry face looks even harsher than the video."* In the interview, she described her insights about that moment, *"When I reflected on the way I was talking, I realized I was trying to take control of other family members. I finally understand what my son meant when he said I'm 'always angry' and 'like an angry bird,' which I hadn't recognized the significance until now."* This shows that relational reflection could uncover dismissed behavioral cues as meaningful insights. In this example, M06 drew on the shared history of her facial expression, frequently commented "always looks angry," to understand its seriousness after seeing her expressions and their effects on others in the data.

6.2 Facilitators of Relational Reflection

In our study, not all the family dialogues we observed led to the relational reflection previously illustrated, nor did all families experience them in the same way. In some cases, reflection practices stopped at merely looking at the data charts, explaining what they've seen. Based on these differences, we analyzed when the reflection moved beyond superficial observations toward a deeper understanding and discovery about their family. Our analysis identified four key factors that facilitate relational reflection – two related to the characteristics of emotional data itself (Section 6.2.1 and 6.2.2), and two related to how families used these data (Section 6.2.3 and 6.2.4) – which we elaborate on in the following sections.

6.2.1 Ambiguity of Emotional Data as Catalyst. The reason families use emotional data to discuss more than speech data types is its inherent ambiguity, making them harder to interpret and more open to differing interpretations among family members. Families engaged with emotional data, such as facial expression and stress level, more frequently and for longer periods than with speech-related data. Table 4 shows that emotional data were recalled more often than speech data (i.e., with a higher frequency of usage). Unlike straightforward conversational data, these emotional data carried ambiguous information, and the interpretation of emotional signals highly relies on shared contextual understandings, aligning with the prior research about the stress signal [49].

This nature of emotional data provided an opportunity for relational reflection, where families could understand one another by recalling shared history or expectations. When the data diverged from their shared history or expectations – for example, when there was a shared knowledge that someone rarely expresses emotions, yet the facial expression data showed many fluctuations – they raised questions. Similarly, families usually expected that the family conversations had been conducted in a generally 'happy' mood, but the data did not reflect that; they questioned it and checked the context through video, eventually converging on an interpretation that could explain the discrepancy.

While the ambiguity of emotional data sometimes acted as a catalyst for deeper discussion, at other times, such ambiguity undermined the accuracy of emotional data, making it difficult for reflection to be sustained. When people judged the emotional data to be incorrect during the reflection, they disengaged and completely stopped using that data in subsequent reflection. Especially in the case of facial expression data, unlike physiological sensor data, people could directly validate the data, whether their expressions were accurately represented, via the video. Thus, when the facial data appeared counterintuitive or erroneous, they began to doubt the accuracy of the data itself. For example, Family 5 found the facial expression data erroneous by matching it with the video. F05 initially doubted, saying, *"I don't think this fluctuating data is correct,"* to which M05 added, *"[C05] is calm, very calm, in the video"* and C05 also denied the data chart, stating, *"Did I feel 'Disgusted'? 'Scared'? What?"* The family ultimately justified their doubt, noting, *"I think it's because the emotion detection is only based on the look on the face, so it can be low accurate."* They also criticized the emotion categorization itself. M15 stated in the interview, *"Data shows that fear, anger, and sadness came out for my emotions, but I couldn't really agree with that."*

6.2.2 Misalignment between Multiple Emotional Data. Through the reconciliation process of discussing misalignment across emotional data from different sources, families were able to actively engage in sharing their own interpretations and reflecting on the very nature of emotion. For example, families often expected that when stress data from the physiological signal was high, it would also align with facial expressions captured in the video data. However, the results did not always show such alignment. In these moments, families did not simply rely on one of the two sources; families came to realize that inner feelings might differ from outward expressions. When C04 asked a question about the stress level data, *"Why was my stress level so high at that moment, different from the facial expression?"*, M04 responded, *"Maybe you were pretending to be fine but actually weren't."*

When it is combined with the shared history of a family, these discussions further enable families to discover new aspects of a person's emotions. For example, C02 and F02 were able to collaboratively draw an insight about M02's emotion regulation from stress level data. As C02 discovered that *"[M02] seems to be highly stressed out,"* F02 added to the discussion that *"I think this was because she is suppressing her emotions. [M02] has high self-control, so her stress level came out high."* C02 further developed the discussion, moving on to another data point: *"[M02] suppresses a lot. Look at the 'Happy' graph. Our graphs are similarly high, but [M02]'s graph is different from us."* From this reflection, C02 noted that in the interview, C02 came to discover something new about his mother – that she tends to hide her emotions – in the interview. This example illustrates how families used their prior knowledge, which is the understanding that M02 often suppresses her emotions, to explain the misalignment between multiple emotional data sources.

6.2.3 Guiding Emotional Disclosure. Encouraging people in the group to share thoughts about their own emotions was an important facilitator of relational reflection, since we observed that many participants were not accustomed to revealing their feelings. When participants were reluctant to disclose their emotions, relational reflection became difficult. For example, in a few families, children refused to re-watch the video at the beginning of the reflection because they felt embarrassed, and thus did not actively participate in the interpretation phase.

In the family context, it is particularly important to guide children to actively engage in reflection, since they have difficulty recognizing their emotions [78] and often show reluctance to share their emotions in an unfamiliar environment. Thus, in our study, this guidance largely becomes the parents' role. For example, M12 consistently prompted her child to reflect about others' emotions and behavior with questions such as, *"When do you think Dad was sad? (pointing emotion expression chart)"*, *"Why do you think this data appeared?"*, and *"Aren't you curious when [C12]'s stress level was at its highest?"*. As a result of these prompts, C12 began to engage with his own data during the discussion and was able to make emotional reflections such as, *"I feel like I'm not stressed usually, but maybe there are times when I'm unconsciously stressed, while talking about the English academy."* For this parental guidance, SELaD's question-type prompts helped parents initiate discussion, reading the prompts word-for-word. For example, M14 read the prompt of the Conversation Summary section, *"(Reading the prompts) When do you think was the moment our family all found it most enjoyable and actively participated? . . . Reminiscing? Was it during reminiscing?"*

In contrast, when parents did not provide appropriate guidance, the reflection process was hindered. Few parents exhibited an authoritarian conversation style, which does not leave room for open dialogue, focusing primarily on their own data or even blaming other family members during the reflection. For example, M06 pointed to others' number of questions and said, *"I asked so many questions, but how can you have this low? Conversations require asking questions and listening. You never engage in a conversation—that's why I keep getting mad at you."* These definitive statements often limited the alternative perspectives and the opportunities for reflection.

6.2.4 Equal Access to Data Among Family Members. We found that equal access to data across the family members provided opportunities for those family members who are under the power imbalance to raise concerns about them, thereby ensuring everyone a symmetric opportunity to lead the relational reflection. Specifically, in our study, children could point out the dominance of the conversation. For example, C06, who was talking with a mother who scolded C06 because of the low number of questions, pointed out his mother's conversation proportion and responded, *"Mom does most of the talking, which is mainly nagging, and I don't have a chance to chip in."* This suggests that presenting data symmetrically to children can provide them with more opportunities to engage and voice their perspectives, even in cases of authoritarian dialogue.

However, we also found that certain factors could hinder equal access and relational reflection, most notably issues of data literacy. While participants reported that level-based charts were generally intuitive and helpful for understanding the data, some families – both parents and children – occasionally struggled to understand the meaning and analysis of physiological data types such as stress level, emotional arousal. Although these difficulties sometimes stimulated further discussion and encouraged families to construct their own interpretations, we also observed that children tended to defer to their parents' interpretations. This highlights the importance of ensuring children's symmetric access not only to the data itself but also to the interpretive process.

7 Discussion

7.1 Relational Reflection using Family Informatics

In this study, each process of *"Relational Reflection"* on family conversation data was identified as illustrated in Section 6. Based on our observation, we discuss how it differs from prior concepts of *dialogic reflection* [32], *collaborative reflection* [76], and *shared reflection* [40].

Fleck and Fitzpatrick noted that dialogic reflection can be fostered when *"an alternate perspective can also be provided by another person"* [32]. Our study extends this claim by showing how social-emotional data of conversation enabled such alternative perspectives and what insights they could gain from the reflection in the family context. In our study, we observed cycles of interpretation and questioning that align with Fleck and Fitzpatrick's description of dialogic reflection – connecting data with one's own experiences and iteratively interpreting them [32].

Building on this notion, we applied it in family settings using social-emotional data rather than individually tracked health data (e.g., sleep, diet, physical activity) that the existing family informatics have focused on [68, 87, 93]. This context revealed how *relational reflection* goes beyond dialogic reflection: families not only connected data with their own experiences but also collaboratively interpreted it by drawing on their *shared histories* and *shared expectations* of family interactions. These insights can further extend the existing family informatics literature that has focused on tracking data for pre-defined goals (e.g., achieving physical activity records) by fostering active discussions about their conversation. They could interpret and reflect on their conversational behaviors using their knowledge about conversational habits or emotional tendencies. This relational reflection could enhance self-knowledge and a deeper understanding of other family members' experiences.

Our findings also differentiate relational reflection from existing notions of shared reflection and collaborative reflection. Shared reflection refers to cases where individuals independently collect data and later exchange reflections [40]. In contrast, relational reflection centers on data from joint activities (e.g., family mealtime conversation), which enabled insights about group-level dynamics – conversational balance, emotional synchronization, and interaction styles – that cannot be captured through individually tracked data. Prior family informatics systems have primarily tracked the individual indicators rather than relational dynamics. However, understanding family dynamics could be one of families' key interests as Pina et al. found that families were interested in seeing "ripple effects" in how one member's behaviors influence others [88]. Relational reflection extends this line of work by centering reflection on joint activities, giving family members a shared context in which to exchange perspectives about how they affect one another and to discover new insights about their emotional influence. Collaborative reflection often involves collective analysis of one person's data by professionals (e.g., medical or educational teams) [76]. By contrast, relational reflection distributes both data and interpretive agency across all family members. Symmetric access, especially for children, ensured that no single perspective dominated and that even marginalized voices could contest, negotiate, or propose changes.

7.2 Multisensor-mediated Relational Reflection on Social-Emotional Behaviors

We found that emotional data played a key role in helping families to become deeply engaged in relational reflection in Section 6.1.2. This was because emotional data facilitated reflection in two ways: first, its inherent ambiguity prompted participants to engage in questioning and validating the data, thereby extending the reflection and anchoring it more closely in the relational context (Section 6.2.1); second, multiple types of emotional data enabled families to reflect on the combination of different data types, rather than the single source (Section 6.2.2). Building on this observation, we argue that relational reflection can mitigate some limitations of the emotional sensing system by enhancing the **contestability of emotion sensing**.

Existing approaches to emotion sensing have been critiqued for their narrow focus on the individual, framing emotions primarily as internal states to be tracked and optimized through behavioral adjustments and self-improvement technologies [45]. Also, the simplification of emotion classification, Ekman's seven emotion categories [27], can reduce the nuanced complexity of human emotions into discrete categories [11]. Such approaches overlook the inherently social and interactional meanings of emotional expression, reducing complex relational dynamics into decontextualized metrics. Howell's work [45] offers a critical alternative, arguing for design goals that move beyond mere detection and categorization toward human-centered emotional reflection and interpretation.

Aligning with this perspective, *relational reflection* can reduce risks of blind trust [23] by allowing discussion on the data validity. In our findings, families questioned the accuracy of the data or noted that their emotions could not be neatly categorized into predefined labels in SELaD. We also directly observed a case where skepticism raised by a family member regarding physiological data prompted others to critically reassess the system's validity in Section 6.2.1. Although such skepticism sometimes led to disengagement from the relational reflection process, providing opportunities to critically examine the validity of data also stimulated reflective dialogue.

Additionally, our findings show that multiple data types of SELaD enabled participants to avoid fully relying on a single source; when one type of data appeared questionable or difficult to interpret, they cross-referenced it with others in Section 6.2.2. These highlight the **value of multiple data types** in emotion sensing, where sensing technologies remain imperfect and emotional states lack a definitive ground truth due to their inherently ambiguous nature [11]. This practice fostered holistic reasoning through the integration of multimodal emotional sensor data, or multisensor fusion at the data level [25], addressing the limitations of each sensor data.

Based on these findings, we propose design directions for family informatics systems that visualize multimodal emotional data to better facilitate relational reflection. First, highlighting misalignments by design. If systems automatically detect inconsistencies across data types (e.g., high physiological stress with low facial arousal), users can more readily initiate relational reflection around why those discrepancies occur. Second, progressive disclosure of information. SELaD currently uses chart-based representations that expose full datasets with minimal abstraction. While comprehensive, it can impose cognitive load when families try to synthesize multiple modalities at once; conversely, overly simplified views risk undercutting interpretability. Progressive disclosure can mitigate this tension by presenting an abstracted overview with on-demand drill-downs to specific data timelines.

While we previously discussed ways to support the use of multimodal emotional data, it remains important to address the risk that over-reliance on sensor-driven feedback can diminish users' agency in interpreting their emotional data. To mitigate this, researchers recommend presenting feedback as hypotheses rather than conclusions, thereby inviting discussion and correction [45]. One approach is to allow users to confirm or re-label their data, which fosters a sense of ownership. This can be supported through features such as selective sharing via manual clicking or interactive modification of data points [28]. Systems that support the co-construction of meaning further encourage active human-data interaction, facilitating deeper reflection rather than passive acceptance [43].

7.3 Scaffolding Relational Reflection with Children

We found that both parental guidance to support children's emotional disclosure (Section 6.2.3) and providing equal, symmetric access to data for all family members (Section 6.2.4) strengthened relational reflection; in both cases, children were central to the process. For children, achieving a shared understanding of their family's social dynamics is particularly important: family reflective practices around emotion facilitate the *parental socialization of emotion* [26], the process through which parents shape children's emotional understanding and regulation, thereby supporting the development of emotional awareness and regulation skills. In our study, we found that providing question-type prompts helped families who were unsure about how to interpret the data (Section 6.2.3), supporting their engagement in reflection. This highlights broader design implications for supporting novice data interpreters, particularly children, in family settings.

For instance, we could design a system that elicits, shares, and stores family members' views about shared history and expectations, and uses them to prompt family-specific data interpretation. Designs that allow each family member to contribute personal annotations can further enable collective storytelling and shared meaning-making [53]. Similarly, *Storywell* integrates reflection into narratives, encouraging parents and children to discuss daily activities through themed questions, which improved motivation and engagement in health tracking [92]. *Reflection Companion* uses adaptive "mini-dialogues" based on scripted reflective questions to promote reflection and behavior change in families [65]. The format of such guidance can also be diversified; for instance, the potential of audio-based prompts can be explored, as speech interfaces can improve accessibility and engagement in health-related self-reporting [87].

Our findings show that symmetric access empowered children's conversational dominance and participation in steering reflection (Section 6.2.4), whereas authoritarian guidance hindered open dialogue (Section 6.2.3). However, ensuring equal access to every participant requires explicit attention to **data literacy**, especially for the children. As shown in our findings, children, in particular, often rely on parental guidance to interpret sensed data, potentially undermining their autonomy and limiting opportunities for independent reflection [4]. These tensions highlight the importance of designing systems that account for varying literacy levels within families, ensuring that all members can meaningfully participate in the process of collaborative data interpretation. To bridge this gap and support their agency, family informatics systems could incorporate child-centered scaffolded visualizations. Instead of raw graphs, systems could offer metaphoric visualizations (e.g., islands floating on the sea [49]) that map complex social-emotional behavioral data to intuitive concepts through the co-design study with children. Furthermore, systems could include literacy tutorials specifically designed for children, ensuring they understand how data is collected and processed before the reflection session begins. This could empower children to enter the conversation as informed equals.

Relational reflection can redistribute power by scaffolding children's participation. In our findings, we observed cases where children actively commented on the conversation imbalance (Section 6.2.4), increasing their awareness of the imbalance. In prior research on personal informatics, this awareness can prompt users to change their behavior through consciousness raising, outcome expectations, self-efficacy, and self-management [54]. Therefore, future work aiming to change behavior through relational reflection could consider how these factors are supported by system design. For example, the system could track conversational dynamics and help families achieve more satisfying communication by offering suggestions for more balanced participation [69].

At the same time, family informatics systems of social-emotional data can reinforce or deepen existing hierarchies—and our findings reflected this risk. In Section 6.2.3, a family member pointed out that the others had asked too few questions. Rather than interpreting social-emotional data as an opportunity for reflection, she treated it as a metric to be achieved. This interpretation could create pressure for the child to speak in particular ways and to display specific emotions. To mitigate these risks, future designs should facilitate democratic operation. Critical actions, such as initiating recordings or launching reflection sessions, should require mutual consent instead of being controlled solely by a parent. In addition, integrating therapy-supported reflection could reduce potential harm [34]. For example, therapists can review social-emotional data alongside families and provide structured guidance. If the therapist identifies relational concerns, they might suggest communication strategies or conversational prompts that support healthier communication dynamics.

7.4 Privacy Concerns

Privacy remains a recurring concern in the literature on family informatics [52, 87, 88]. In our study, however, the majority of families' children (15 out of 17) explicitly expressed that they felt comfortable with sharing data as long as the data is shared exclusively among family members and not with others. One reason may be that the system allows children not only to show their own data but also to view their parents' data, creating a sense of symmetry in access [64].

Even if strong concerns were not evident in our study, privacy risks remain important. As noted in Section 7.3, privacy can be unevenly protected in settings involving children, especially because the system handles sensitive emotional data. According to contextual integrity theory [82], privacy expectations are shaped by the specific context and norms surrounding the flow of information. For example, children may agree to share data in one situation but feel differently in another, depending on the social boundaries of that context, such as who is involved and how the information is shared. In our study, some families noted that sharing social-emotional data felt acceptable to them, but it could introduce privacy risks for other families with more conflict – for example, a parent reacting negatively after seeing a spike in the child's stress in their conversation. Even when a family does not experience overt conflict, they may still be reluctant to share social-emotional data when the conversation topic is sensitive, such as household rules, academic performance, or peer relationships.

If future systems with social-emotional data are designed, they would need to carefully account for the contextual privacy risks discussed above, rather than assuming that such data should always be shared. One possible data-sharing strategy could be contextual consent to help navigate these boundaries. For example, the system could flag instances where a child's emotionally sensitive data, such as stress peaks or negative facial expressions, is about to be shared with extended family members, prompting caregivers to confirm whether such disclosure aligns with the child's privacy boundaries in that context. Another approach is to support granular privacy control [67]. Children should be able to blur or hide specific segments of data (e.g., a "private mode" for selected conversation topics or the option to pause data collection) without requiring parental approval.

8 Limitations

One key limitation of our study is that it was conducted in a controlled lab-based setting rather than through real-world deployment. The user study involved three pre-selected conversation scenarios conducted in a house-like testbed environment. While this setting enabled data collection under consistent conditions, it may have introduced artifacts such as participant discomfort with the camera, constrained interaction angles, or unnatural conversation flow prompted by assigned topics. In addition to environmental constraints, our sample consisted of seventeen self-selected Korean families. As a result, the study may not fully capture the diversity of parenting styles, emotional climates, or socio-cultural factors that shape family dynamics. This limited scope restricts the generalizability of our findings and highlights the need for future research involving more diverse populations. In addition, the child's position within the family can raise ethical concerns about power dynamics, especially when children feel unable to express their views or privacy concerns. Future work could address this by conducting follow-up interviews with children separately from their families or by involving child-development specialists to help elicit children's perspectives in an age-appropriate way. Moreover, our thematic analysis focused primarily on explicitly spoken reflections. While we verified these instances through screen activity and logs, non-verbal or indirect forms of reflection may have gone unrecognized. Future studies could integrate more comprehensive multimodal approaches to capture a fuller range of reflective behaviors in the analysis.

A system-level limitation of SELaD is the accuracy of its emotion recognition models, particularly for certain categories such as sadness, due to model bias and a lack of training diversity. While the current prototype served as a proof-of-concept, future iterations should improve model robustness and accommodate a wider range of settings. Beyond technical limitations, participants reported usability issues with the data components of the system. For *Conversation Summary*, participants found it difficult to understand the temporal context of the *keyword* or *positive words* (e.g., when a specific keyword was spoken). To address this, future designs should integrate timestamp-linked transcripts or visual cues that highlight keyword usage over time.

9 Conclusion

This study demonstrated how multimodal social and emotional sensing data can facilitate relational reflection in family settings. We designed and implemented a family informatics system called SELaD with multimodal data and conducted a user study with 17 families. Our results showed that the families explored their social and emotional behaviors, collaboratively interpreted their interactions, and gained deeper insights into family dynamics.

In response to RQ1-2, families reflected on multimodal social-emotional data by recalling shared histories and shared expectations, and they co-constructed insights about conversational patterns (e.g., recognizing dominant speaker, turn-taking, positive words) and discovery about emotional influence. This could facilitate not only self-evaluation and feeling a need for positive change, but also empathy and perspective-taking. In response to RQ3, we identified key facilitators: the ambiguity of emotional data as a catalyst for deeper discussion; the misalignment across data types (e.g., stress level vs. facial expression) to question, validate, and reconcile interpretations; parental guidance and question-type prompts that helped initiate reflection; and equal, symmetric access that enabled children to participate in the reflection.

By integrating these findings, we suggested design implications for relational reflection through making emotion sensing more contestable and supportive of family sensemaking, (i) technologies leveraging multimodal social-emotional data, (ii) prompts to scaffold the data reflection of families, (iii) strategies to mitigate the ethical concerns from usage within a family and from the use of social-emotional data.

Acknowledgments

Corresponding authors are Hyunsoo Lee, Jennifer G. Kim, and Uichin Lee. This research was supported by the KAIST Future Smart Home Research Center grant funded by the Taejae Research Foundation and by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) [RS-2025-02305705]. Jennifer G. Kim was partly supported through the Industrial Technology Innovation Program (P0028404) of the Ministry of Industry, Trade and Energy of Korea and by the internal fund/grant of Electronics and Telecommunications Research Institute (ETRI) [25YD1100].

References

[1] Hiroyuki Adachi, Seiko Myojin, and Nobutaka Shimada. 2015. ScoringTalk: a tablet system scoring and visualizing conversation for balancing of participation. In *SIGGRAPH Asia 2015 Mobile Graphics and Interactive Applications*. 1–5.

[2] Daniel A Adler, Emily Tseng, Khatiya C Moon, John Q Young, John M Kane, Emanuel Moss, David C Mohr, and Tanzeem Choudhury. 2022. Burnout and the quantified workplace: tensions around personal sensing interventions for stress in resident physicians. *Proceedings of the ACM on Human-computer Interaction* 6, CSCW2 (2022), 1–48.

[3] Karan Ahuja, Dohyun Kim, Franceska Xhakaj, Virag Varga, Anne Xie, Stanley Zhang, Jay Eric Townsend, Chris Harrison, Amy Ogan, and Yuvraj Agarwal. 2019. EduSense: Practical classroom sensing at Scale. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–26.

[4] Alissa N Antle and Alexandra Kitson. 2021. 1, 2, 3, 4 tell me how to grow more: A position paper on children, design ethics and biowearables. *International Journal of Child-Computer Interaction* 30 (2021), 100328.

[5] Riku Arakawa and Hiromu Yakura. 2020. INWARD: A computer-supported tool for video-reflection improves efficiency and effectiveness in executive coaching. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.

[6] Diana Baumrind. 1978. Parental disciplinary patterns and social competence in children. *Youth & Society* 9, 3 (1978), 239–267.

[7] Daryl J Bem. 1972. Self-perception theory. *Advances in experimental social psychology* 6 (1972).

[8] Mathias Benedek and Christian Kaernbach. 2010. A continuous measure of phasic electrodermal activity. *Journal of neuroscience methods* 190, 1 (2010), 80–91.

[9] Tony Bergstrom and Karrie Karahalios. 2007. Conversation Clock: Visualizing audio patterns in co-located groups. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. IEEE, 78–78.

[10] Jomara Binda, Chien Wen Yuan, Natalie Cope, Hyehyun Park, Eun Kyoung Choe, and John M Carroll. 2018. Supporting effective sharing of health information among intergenerational family members. In *Proceedings of the 12th eai international conference on pervasive computing technologies for healthcare*. 148–157.

[11] Kirsten Boehner, Rogério DePaula, Paul Dourish, and Phoebe Sengers. 2007. How emotion is made and measured. *International Journal of Human-Computer Studies* 65, 4 (2007), 275–291.

[12] Margaret M Bradley and Peter J Lang. 2000. Measuring emotion: Behavior, feeling, and physiology. (2000).

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

[14] Mark Bubel, Ruiwen Jiang, Christine H Lee, Wen Shi, and Audrey Tse. 2016. AwareMe: addressing fear of public speech through awareness. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. 68–73.

[15] A John Camm, Marek Malik, J Thomas Bigger, Günter Breithardt, Sergio Cerutti, Richard J Cohen, Philippe Coumel, Ernest L Fallen, Harold L Kennedy, Robert E Kleiger, et al. 1996. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Circulation* 93, 5 (1996), 1043–1065.

[16] Yoon Jeong Cha, Yasemin Gunal, Alice Wou, Joyce Lee, Mark W Newman, and Sun Young Park. 2024. Shared Responsibility in Collaborative Tracking for Children with Type 1 Diabetes and their Parents. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–20.

[17] Senthil Chandrasegaran, Chris Bryan, Hidekazu Shidara, Tung-Yen Chuang, and Kwan-Liu Ma. 2019. TalkTraces: Real-time capture and visualization of verbal content in meetings. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–14.

[18] Ying-Yu Chen, Ziyue Li, Daniela Rosner, and Alexis Hiniker. 2019. Understanding parents' perspectives on mealtime technology. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 1 (2019), 1–19.

[19] Eunji Chong, Katha Chanda, Zhefan Ye, Audrey Southerland, Nataniel Ruiz, Rebecca M Jones, Agata Rozga, and James M Rehg. 2017. Detecting gaze towards eyes in natural social interactions and its use in child assessment. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 3 (2017), 1–20.

[20] Enrico Coiera, Kathleen Yin, Roneel V Sharan, Saba Akbar, Satya Vedantam, Hao Xiong, Jenny Waldie, and Annie YS Lau. 2022. Family informatics. *Journal of the American Medical Informatics Association* 29, 7 (2022), 1310–1315.

[21] Hugo D Critchley. 2002. Electrodermal responses: what happens in the brain. *The Neuroscientist* 8, 2 (2002), 132–142.

[22] Narjisse Damoun, Youssra Amekran, Nora Taiek, and Abdelkader Jalil El Hangouche. 2024. Heart rate variability measurement and influencing factors: Towards the standardization of methodology. *Global Cardiology Science & Practice* 2024, 4 (2024), e202435.

[23] Maria De-Arteaga, Riccardo Fogliato, and Alexandra Chouldechova. 2020. A case for humans-in-the-loop: Decisions in the presence of erroneous algorithmic scores. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[24] Joan Morris DiMicco, Katherine J Hollenbach, and Walter Bender. 2006. Using visualizations to review a group's interaction dynamics. In *CHI'06 extended abstracts on Human factors in computing systems*. 706–711.

[25] Bruno Dumas, Denis Lalanne, and Sharon Oviatt. 2009. Multimodal interfaces: A survey of principles, models and frameworks. In *Human machine interaction: Research results of the mmi program*. Springer, 3–26.

[26] Nancy Eisenberg, Amanda Cumberland, and Tracy L Spinrad. 1998. Parental socialization of emotion. *Psychological inquiry* 9, 4 (1998), 241–273.

[27] Paul Ekman. 1992. Facial expressions of emotion: New findings, new questions.

[28] Daniel A Epstein, Alan Borning, and James Fogarty. 2013. Fine-grained sharing of sensed physical activity: A value sensitive approach. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. 489–498.

[29] Daniel A Epstein, Bradley H Jacobson, Elizabeth Bales, David W McDonald, and Sean A Munson. 2015. From" nobody cares" to" way to go!" A Design Framework for Social Sharing in Personal Informatics. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1622–1636.

[30] Anne K Fishel. 2016. Harnessing the power of family dinners to create change in family therapy. *Australian and New Zealand Journal of Family Therapy* 37, 4 (2016), 514–527.

[31] E Fivaz-Depeursinge and A Corboz-Warnery. 1999. The primary triangle. A developmental systems view.

[32] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on reflection: framing a design landscape. In *Proceedings of the 22nd conference of the computer-human interaction special interest group of australia on computer-human interaction*. 216–223.

[33] Sarah E Fox, Amanda Menking, Jordan Eschler, and Uba Backonja. 2020. Multiples over models: interrogating the past and collectively reimagining the future of menstrual sensemaking. *ACM Transactions on Computer-Human Interaction (TOCHI)* 27, 4 (2020), 1–24.

[34] Ruben G Fukkink. 2008. Video feedback in widescreen: A meta-analysis of family programs. *Clinical Psychology Review* 28, 6 (2008), 904–916.

[35] Thomas Fydrich, Dianne L Chambless, Kevin J Perry, Friederike Buergener, and Maria B Beazley. 1998. Behavioral assessment of social performance: A rating system for social phobia. *Behaviour research and therapy* 36, 10 (1998), 995–1010.

[36] Mirta Galesic, Wändi Bruine de Bruin, Jonas Dalege, Scott L Feld, Frauke Kreuter, Henrik Olsson, Drazen Prelec, Daniel L Stein, and Tamara van Der Does. 2021. Human social sensing is an untapped resource for computational social science. *Nature* 595, 7866 (2021), 214–222.

[37] Sarah Gallacher, Jenny O'Connor, Jon Bird, Yvonne Rogers, Licia Capra, Daniel Harrison, and Paul Marshall. 2015. Mood squeezer: lightening up the workplace through playful and lightweight interactions. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 891–902.

[38] Tivan Varghese George, Hye Soo Park, and Uichin Lee. 2024. FT Xtraction: Feature extraction and visualization of conversational video data for social and emotional analysis. *SoftwareX* 27 (2024), 101827.

[39] Ian J Goodfellow, Dumitru Erhan, Pierre Luc Carrier, Aaron Courville, Mehdi Mirza, Ben Hamner, Will Cukierski, Yichuan Tang, David Thaler, Dong-Hyun Lee, et al. 2013. Challenges in representation learning: A report on three machine learning contests. In *Neural Information Processing: 20th International Conference, ICONIP 2013, Daegu, Korea, November 3-7, 2013. Proceedings, Part III 20*. Springer, 117–124.

[40] Lisa Graham, Anthony Tang, and Carman Neustaedter. 2016. Help me help you: Shared reflection for personal data. In *Proceedings of the 2016 ACM International Conference on Supporting Group Work*. 99–109.

[41] Chenxu Hao, Tiffany Matej Hrkalovic, Daniel Balliet, Hayley Hung, and Bernd Dudzik. 2025. Technologies Supporting Self-Reflection on Social Interactions: A Systematic Review. In *Proceedings of the 30th International Conference on Intelligent User Interfaces*. 1354–1365.

[42] Keita Higuchi, Soichiro Matsuda, Rie Kamikubo, Takuya Enomoto, Yusuke Sugano, Junichi Yamamoto, and Yoichi Sato. 2018. Visualizing gaze direction to support video coding of social attention for children with autism spectrum disorder. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces*. 571–582.

[43] Naja Holten Møller, Gina Neff, Jakob Grue Simonsen, Jonas Christoffer Villumsen, and Pernille Bjørn. 2021. Can workplace tracking ever empower? Collective sensemaking for the responsible use of sensor data at work. *Proceedings of the ACM on human-computer interaction* 5, GROUP (2021), 1–21.

[44] Mohammed Hoque, Matthieu Courgeon, Jean-Claude Martin, Bilge Mutlu, and Rosalind W Picard. 2013. Mach: My automated conversation coach. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. 697–706.

[45] Noura Howell, John Chuang, Abigail De Kosnik, Greg Niemeyer, and Kimiko Ryokai. 2018. Emotional biosensing: Exploring critical alternatives. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–25.

[46] Bernd Huber, Richard F Davis III, Allison Cotter, Emily Junkin, Mindy Yard, Stuart Shieber, Elizabeth Brestan-Knight, and Krzysztof Z Gajos. 2019. SpecialTime: Automatically detecting dialogue acts from speech to support parent-child interaction therapy. In *Proceedings of the 13th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 139–148.

[47] Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. 2003. Technology probes: inspiring design for and with families. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 17–24.

[48] Inseok Hwang, Chungkuk Yoo, Chanyou Hwang, Dongsun Yim, Youngki Lee, Chulhong Min, John Kim, and Junehwa Song. 2014. TalkBetter: family-driven mobile intervention care for children with language delay. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 1283–1296.

[49] Yanqi Jiang, Xianghua Ding, Xiaojuan Ma, Zhida Sun, and Ning Gu. 2023. IntimaSea: exploring shared stress display in close relationships. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–19.

[50] Wu Jie. 2020. *Facial-Expression-Recognition.Pytorch*. https://github.com/WuJie1010/Facial-Expression-Recognition.Pytorch Accessed: 2024-09-13.

[51] Eunkyung Jo, Hyeonseok Bang, Myeonghan Ryu, Eun Jee Sung, Sungmook Leem, and Hwajung Hong. 2020. MAMAS: supporting parent–child mealtime interactions using automated tracking and speech recognition. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–32.

[52] Mikkel S Jørgensen, Frederik K Nissen, Jeni Paay, Jesper Kjeldskov, and Mikael B Skov. 2016. Monitoring children's physical activity and sleep: a study of surveillance and information disclosure. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction*. 50–58.

[53] Tobias Kauer, Marian Dörk, and Benjamin Bach. 2025. Towards Collective Storytelling: Investigating Audience Annotations in Data Visualizations. *IEEE Computer Graphics and Applications* (2025).

[54] Elisabeth T Kersten-van Dijk, Joyce HDM Westerink, Femke Beute, and Wijnand A IJsselsteijn. 2017. Personal informatics, self-insight, and behavior change: A critical review of current literature. *Human–Computer Interaction* 32, 5-6 (2017), 268–296.

[55] Bogoan Kim, Dayoung Jeong, Mingon Jeong, Taehyung Noh, Sung-In Kim, Taewan Kim, So-Youn Jang, Hee Jeong Yoo, Jennifer Kim, Hwajung Hong, et al. 2022. Vista: User-centered vr training system for effectively deriving characteristics of people with autism spectrum disorder. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*. 1–12.

[56] Bogoan Kim, Dayoung Jeong, Jennifer G Kim, Hwajung Hong, and Kyungsik Han. 2023. V-DAT (Virtual Reality Data Analysis Tool): Supporting Self-Awareness for Autistic People from Multimodal VR Sensor Data. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–13.

[57] Hye-Geum Kim, Eun-Jin Cheon, Dai-Seg Bai, Young Hwan Lee, and Bon-Hoon Koo. 2018. Stress and heart rate variability: a meta-analysis and review of the literature. *Psychiatry investigation* 15, 3 (2018), 235.

[58] Hyun-joong Kim, Sungzoon Cho, and Pilsung Kang. 2014. KR-WordRank: An unsupervised Korean word extraction method based on WordRank. *Journal of Korean Institute of Industrial Engineers* 40, 1 (2014), 18–33.

[59] Jennifer Kim, Melinda Snodgrass, Mary Pietrowicz, Karrie Karahalios, and Jim Halle. 2013. BEDA: Visual analytics for behavioral and physiological data. In *Workshop on Visual Analytics in Healthcare. Washington DC*. 23–27.

[60] Kyung Hwan Kim, Seok Won Bang, and Sang Ryong Kim. 2004. Emotion recognition system using short-term monitoring of physiological signals. *Medical and biological engineering and computing* 42 (2004), 419–427.

[61] Sung-In Kim, So-youn Jang, Taewan Kim, Bogoan Kim, Dayoung Jeong, Taehyung Noh, Mingon Jeong, Kaely Hall, Meelim Kim, Hee Jeong Yoo, et al. 2024. Promoting self-efficacy of individuals with autism in practicing social skills in the workplace using virtual reality and physiological sensors: Mixed methods study. *JMIR Formative Research* 8 (2024), e52157.

[62] Taemie Kim, Agnes Chang, Lindsey Holland, and Alex Sandy Pentland. 2008. Meeting mediator: enhancing group collaborationusing sociometric feedback. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 457–466.

[63] Wonjung Kim, Seungchul Lee, Seonghoon Kim, Sungbin Jo, Chungkuk Yoo, Inseok Hwang, Seungwoo Kang, and Junehwa Song. 2020. Dyadic mirror: Everyday second-person live-view for empathetic reflection upon parent-child interaction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–29.

[64] Minsam Ko, Subin Yang, Joonwon Lee, Christian Heizmann, Jinyoung Jeong, Uichin Lee, Daehee Shin, Koji Yatani, Junehwa Song, and Kyong-Mee Chung. 2015. NUGU: a group-based intervention app for improving self-regulation of limiting smartphone use. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 1235–1245.

[65] Rafal Kocielnik, Lillian Xiao, Daniel Avrahami, and Gary Hsieh. 2018. Reflection companion: a conversational system for engaging users in reflection on physical activity. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 2 (2018), 1–26.

[66] Andrea E Kowallik and Stefan R Schweinberger. 2019. Sensor-based technology for social information processing in autism: A review. *Sensors* 19, 21 (2019), 4787.

[67] Scott Lederer, Jennifer Mankoff, Anind K Dey, and Christopher Beckmann. 2003. Managing personal information disclosure in ubiquitous computing environments. *Intel Research, IRB-TR-03-015* (2003).

[68] Hyunsoo Lee, Yugyeong Jung, Youwon Shin, Hyesoo Park, Woohyeok Choi, and Uichin Lee. 2024. FamilyScope: Visualizing Affective Aspects of Family Social Interactions using Passive Sensor Data. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–27.

[69] Moon-Hwan Lee, Yea-Kyung Row, Oosung Son, Uichin Lee, Jaejeung Kim, Jungi Jeong, Seungryoul Maeng, and Tek-Jin Nam. 2018. Flower-Pop: Facilitating casual group conversations with multiple mobile devices. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 4 (2018), 1–24.

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

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

[72] Patrick Lucey, Jeffrey F Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, and Iain Matthews. 2010. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *2010 ieee computer society conference on computer vision and pattern recognition-workshops*. IEEE, 94–101.

[73] Camillo Lugaresi, Jiuqiang Tang, Hadon Nash, Chris McClanahan, Esha Uboweja, Michael Hays, Fan Zhang, Chuo-Ling Chang, Ming Yong, Juhyun Lee, et al. 2019. Mediapipe: A framework for perceiving and processing reality. In *Third workshop on computer vision for AR/VR at IEEE computer vision and pattern recognition (CVPR)*, Vol. 2019.

[74] Jock Mackinlay. 1986. Automating the design of graphical presentations of relational information. *Acm Transactions On Graphics (Tog)* 5, 2 (1986), 110–141.

[75] Julia A Malia. 2006. Basic concepts and models of family stress. *Stress, trauma, and crisis* 9, 3-4 (2006), 141–160.

[76] Gabriela Marcu, Anind K Dey, and Sara Kiesler. 2014. Designing for collaborative reflection. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare*. 9–16.

[77] Akhil Mathur, Marc Van den Broeck, Geert Vanderhulst, Afra Mashhadi, and Fahim Kawsar. 2015. Tiny habits in the giant enterprise: understanding the dynamics of a quantified workplace. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 577–588.

[78] Amanda Sheffield Morris, Jennifer S Silk, Laurence Steinberg, Sonya S Myers, and Lara Rachel Robinson. 2007. The role of the family context in the development of emotion regulation. *Social development* 16, 2 (2007), 361–388.

[79] Kerrianne E Morrison, Kilee M DeBrabander, Desiree R Jones, Robert A Ackerman, and Noah J Sasson. 2020. Social cognition, social skill, and social motivation minimally predict social interaction outcomes for autistic and non-autistic adults. *Frontiers in Psychology* 11 (2020), 591100.

[80] Kerrianne E Morrison, Amy E Pinkham, David L Penn, Skylar Kelsven, Kelsey Ludwig, and Noah J Sasson. 2017. Distinct profiles of social skill in adults with autism spectrum disorder and schizophrenia. *Autism Research* 10, 5 (2017), 878–887.

[81] NAVER Cloud Platform. 2024. *Clova Speech*. https://www.ncloud.com/product/aiService/clovaSpeech Last accessed on May 1, 2024.
[82] Helen Nissenbaum. 2009. Privacy in context: Technology, policy, and the integrity of social life. In *Privacy in context*. Stanford University Press.
[83] Işil Oygür, Daniel A Epstein, and Yunan Chen. 2020. Raising the responsible child: collaborative work in the use of activity trackers for children. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–23.
[84] Iþil Oygür, Zhaoyuan Su, Daniel A. Epstein, and Yunan Chen. 2021. The Lived Experience of Child-Owned Wearables: Comparing Children's and Parents' Perspectives on Activity Tracking. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–12.
[85] Sungzoon Park, Eunjeong L.and Cho. 2014. KoNLPy: Korean natural language processing Python. In *Proceedings of the 26th Annual Conference on Human & Cognitive Language Technology*. Chuncheon, Korea.
[86] Evanthia N Patrikakou and Amy R Anderson. 2005. *School-family partnerships for children's success*. Teachers College Press.
[87] Laura Pina, Sang-Wha Sien, Clarissa Song, Teresa M Ward, James Fogarty, Sean A Munson, and Julie A Kientz. 2020. DreamCatcher: exploring how parents and school-age children can track and review sleep information together. *Proceedings of the ACM on Human-computer Interaction* 4, CSCW1 (2020), 1–25.
[88] Laura R Pina, Sang-Wha Sien, Teresa Ward, Jason C Yip, Sean A Munson, James Fogarty, and Julie A Kientz. 2017. From personal informatics to family informatics: Understanding family practices around health monitoring. In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing*. 2300–2315.
[89] Amy E Pinkham and David L Penn. 2006. Neurocognitive and social cognitive predictors of interpersonal skill in schizophrenia. *Psychiatry research* 143, 2-3 (2006), 167–178.
[90] Amon Rapp and Maurizio Tirassa. 2017. Know thyself: a theory of the self for personal informatics. *Human–Computer Interaction* 32, 5-6 (2017), 335–380.
[91] Sandra Rusconi-Serpa, Ana Sancho Rossignol, and Susan C McDonough. 2009. Video feedback in parent-infant treatments. *Child and Adolescent Psychiatric Clinics* 18, 3 (2009), 735–751.
[92] Herman Saksono, Carmen Castaneda-Sceppa, Jessica Hoffman, Vivien Morris, Magy Seif El-Nasr, and Andrea G Parker. 2020. Storywell: designing for family fitness app motivation by using social rewards and reflection. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.
[93] Herman Saksono, Ashwini Ranade, Geeta Kamarthi, Carmen Castaneda-Sceppa, Jessica A Hoffman, Cathy Wirth, and Andrea G Parker. 2015. Spaceship Launch: Designing a collaborative exergame for families. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1776–1787.
[94] Christopher L Schaefbauer, Danish U Khan, Amy Le, Garrett Sczechowski, and Katie A Siek. 2015. Snack buddy: supporting healthy snacking in low socioeconomic status families. In *Proceedings of the 18th acm conference on computer supported cooperative work & social computing*. 1045–1057.
[95] Marianne Schmid Mast, Daniel Gatica-Perez, Denise Frauendorfer, Laurent Nguyen, and Tanzeem Choudhury. 2015. Social sensing for psychology: Automated interpersonal behavior assessment. *Current Directions in Psychological Science* 24, 2 (2015), 154–160.
[96] Philip Schmidt, Attila Reiss, Robert Duerichen, Claus Marberger, and Kristof Van Laerhoven. 2018. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM international conference on multimodal interaction*. 400–408.
[97] Katerina Silhánová, Ruth Cave, Rubin Fukkink, Michelle Sancho, Maria Doria, Jacqueline Bristow, Denise McCartan, Henk Vermeulen, Carole S Chasle, Colwyn Trevarthen, et al. 2011. *Video interaction guidance: A relationship-based intervention to promote attunement, empathy and wellbeing*. Jessica Kingsley Publishers.
[98] Petr Slovák and Geraldine Fitzpatrick. 2015. Teaching and developing social and emotional skills with technology. *ACM Transactions on Computer-Human Interaction (TOCHI)* 22, 4 (2015), 1–34.
[99] Mani Srivastava, Tarek Abdelzaher, and Boleslaw Szymanski. 2012. Human-centric sensing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 370, 1958 (2012), 176–197.
[100] Evropi Stefanidi, Jonathan Luis Benjamin Wassmann, Paweł W Woźniak, Gunnar Spellmeyer, Yvonne Rogers, and Jasmin Niess. 2024. MoodGems: Designing for the Well-being of Children with ADHD and their Families at Home. In *Proceedings of the 23rd Annual ACM Interaction Design and Children Conference*. 480–494.
[101] Laurence Steinberg and Amanda Sheffield Morris. 2001. Adolescent development. *Annual review of psychology* 52, 1 (2001), 83–110.
[102] Janienke Sturm, Olga Houben-van Herwijnen, Anke Eyck, and Jacques Terken. 2007. Influencing social dynamics in meetings through a peripheral display. In *Proceedings of the 9th international conference on Multimodal interfaces*. 263–270.
[103] Hans Stuyck, Leonardo Dalla Costa, Axel Cleeremans, and Eva Van den Bussche. 2022. Validity of the Empatica E4 wristband to estimate resting-state heart rate variability in a lab-based context. *International Journal of Psychophysiology* 182 (2022), 105–118.
[104] Leimin Tian, Sharon Oviatt, Michal Muszynski, Brent Chamberlain, Jennifer Healey, and Akane Sano. 2022. Applied Affective Computing. (2022).
[105] Tammy Toscos, Kay Connelly, and Yvonne Rogers. 2012. Best intentions: health monitoring technology and children. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*. 1431–1440.
[106] Ha Trinh, Reza Asadi, Darren Edge, and T Bickmore. 2017. Robocop: A robotic coach for oral presentations. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 1, 2 (2017), 1–24.
[107] Jeffrey Vongmany, Tim Luckett, L Lam, and JL Phillips. 2018. Family behaviours that have an impact on the self-management activities of adults living with Type 2 diabetes: a systematic review and meta-synthesis. *Diabetic Medicine* 35, 2 (2018), 184–194.
[108] Dennis Wang, Marawin Chheang, Siyun Ji, Ryan Mohta, and Daniel A Epstein. 2022. SnapPI: Understanding everyday use of personal informatics data stickers on ephemeral social media. *Proceedings of the ACM on human-computer interaction* 6, CSCW2 (2022), 1–27.
[109] Lisa Whitehead, Elisabeth Jacob, Amanda Towell, Ma'en Abu-qamar, and Amanda Cole-Heath. 2018. The role of the family in supporting the self-management of chronic conditions: A qualitative systematic review. *Journal of clinical nursing* 27, 1-2 (2018), 22–30.
[110] Chungkuk Yoo, Seungwoo Kang, Inseok Hwang, Chulhong Min, Seonghoon Kim, Wonjung Kim, and Junehwa Song. 2019. Mom, i see you angry at me! Designing a mobile service for parent-child conflicts by in-situ emotional empathy. In *Proceedings of the 5th ACM Workshop on Mobile Systems for Computational Social Science*. 21–26.
[111] Camellia Zakaria, Rajesh Balan, and Youngki Lee. 2019. StressMon: scalable detection of perceived stress and depression using passive sensing of changes in work Routines and group interactions. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–29.

A Question-type Prompts in Each Data Section and Related SEL Competencies

Sections	Question-type Prompts	Related Competencies
Conversation Summary	Did the topics and keywords of the conversation match well? What was the proportion of positive words used by each member?	Self-awareness
	Identify the strengths and weaknesses that appeared in the conversation.	Self-awareness
	Did you communicate well to the other family members what you want and what you are pursuing?	Self-management
Conversation Behaviors	When do you think was the time your family engaged in the conversation most? What do you think were the feelings of the other family members at that time?	Relationship skills
	When do you think was the moment our family all found it most enjoyable and actively participated? Why do you think so?	Relationship skills
	Do you think that everyone in our family equally participated in the conversation? Why do you think so?	Social-awareness
Interaction Behaviors	What emotions did you mostly feel during the conversation? Why did you feel those emotions?	Self-awareness
	What emotions did the family members feel? What makes you think they felt those emotions?	Social-awareness
	Do you think the family members generally felt synchronized emotions? How did their emotions change?	Social-awareness
Physiological Response	Who seemed to be the most excited? Why do you think so?	Relationship skills
	Were you able to manage and express your emotions when discussing different perspectives?	Self-management
	Were the reactions of the family members similar to what you thought? Was there any family member who showed unexpected reactions?	Responsible decision-making
	Were you influenced by the emotions of other family members?	Responsible decision-making

B Interview Questions Used in the User Study

Perception on Social and Emotional Data
(For each type of data) What do you think it represents?
Do you find each data helpful for understanding your family conversation better?
Which information was particularly useful, and which was difficult to understand?
If other types of information were needed, what kind of data do you think would be helpful?
How do you feel about sharing physiological and behavioral data with your family members through the system?
(For children) Do you think it would be okay for your mother, father, or other family members to see your data? If not, in what situations do you think it would be inappropriate?
If this system were available for use at home, what concerns would you have?
If this type of system existed, do you think you would use it to understand your family's emotional states, communication styles, and others better?
Family Reflection Using Data Visualization
During your exploration of the system, when did you find this feature the most useful?
Do you think the data charts were useful in identifying and comparing the conversational styles of your family?
Did you use the question-type prompts to interpret the visualizations often?
Did watching the video along with the data help you understand your family's emotions and interaction context? How did it help?
If there were discrepancies in data interpretation in co-reflection, what were they about, and how did each person interpret the situation?
When there were discrepancies in data interpretation, how did you resolve them?
As you reflect on the data together, did you discover any new information through others' opinions?
Who took the lead in presenting opinions and guiding the direction of reflection?
Insights About Families' Social and Emotional Behaviors
How do you think your family normally communicates?
Was today's conversation different from your family's usual conversations? If it was different, why? What aspects were different?
Do you think this tool helped you understand how the conversation was and what areas needed improvement?
Did you gain any new insights about your or other family members' emotional states, conversational habits, or behaviors?

C Implementation Details

C.1 Video Data Analysis

The recorded video files were about 30 minutes long, formatted as 1080p, 30 fps MOV files. After data collection, these were encoded into MP4, h.264 format for compatibility with video processing. To reduce computation time, video processing was conducted every 60 frames (2 seconds) instead of frame-by-frame. For each target frame, the FaceDetection model from Google's MediaPipe library was used to obtain the face bounding box. Facial emotion recognition and facial landmarks were detected for each target frame, and based on these values, features such as total expressed emotions, emotion synchrony, and eye contact were extracted every 600 frames (20 seconds) window. Person recognition was based on order-based matching, as participants remained seated in the same positions throughout the sessions.

- *Facial Emotion Recognition* is a method to extract values for facial emotions. A pre-trained model [50] was used for facial emotion recognition, which claimed to have achieved 73.11% performance on the FER2013 dataset [39] and 94.64% on the CK+ dataset [72]. The model classifies each facial expression into one of seven emotions (anger, disgust, fear, happiness, sadness, surprise, neutral), returning an index (0-6). This index is converted to a one-hot vector (e.g., 2 into $[0, 0, 1, 0, 0, 0, 0]$), stored per target frame for further feature calculation.
- *Facial Landmarks Detection* is a method to identify key parts of the face. The facial landmarks were detected by the well-known FaceMesh model from Google's MediaPipe library [73]. This model detects 468 facial landmarks, each with their x, y, and z coordinates. Extracted facial landmarks for each target frame were saved in a variable and used for calculating derived features.

Subsequently, for each time window (20s), derived features were calculated as follows:

- *Emotional Expression*, the total expressed emotions vector of person i, denoted as TEE_i, was computed by summing the one-hot emotion vectors across all target frames within a given time window, representing the frequency of each expressed emotion.
- *Emotion Synchrony*, indicating the similarity between the emotions of person i and person j, was calculated using the Euclidean distance between their TEE vectors.
- *Eye Contact*, indicating whether one person's gaze intersects another's facial bounding box, was calculated using facial landmarks and bounding box values. To make this system feasible with a single camera – suitable for in-home use – we assumed a fixed-angle setup where participants are seated side-by-side.

For *emotion synchrony* and *eye contact*, pairwise feature values were averaged for the three pairs and then ordinally discretized into four levels: *low* (0.0-0.25), *moderate* (0.25-0.5), *high* (0.5-0.75), and *very high* (0.75-1.0).

C.2 Speech Data Analysis

We utilized Naver Clova Speech AI [81] for its high performance in Korean speech transcription. Initially, we processed transcripts for the entire audio, by setting the minimum number of speakers to 3 for speaker diarization and sending the request. Then, we used the acquired transcription.json for speech data analysis.

The transcription file provided data including the text of separated segments, the diarized speaker, and start and end times of each segment. Segment duration was calculated using start and end times, and further analysis included calculating positive word proportion, number of questions, number of speaking initiations, and speaking speed (WPM).

- *Positive words proportion* was analyzed using Clova Semantic Analysis AI [81].
- *Number of questions* was counted using question marks.
- *Speaking initiation* was counted equal to the number of segments.
- *Speaking speed* was calculated by dividing the number of words in each segment by the segment's duration in minutes.

To obtain *Keywords*, we aggregated speech segments by each person and each conversation topic and extracted nouns using the Okt tokenizer from the konlpy package [85] for morphological analysis, and extracted significant keywords for each conversation topic per person using the unsupervised learning-based keyword extraction algorithm KRWordRank [58].

C.3 Physiological Data Analysis

In the SELaD system, physiological sensor signals were used to capture emotional arousal and stress levels, complementing observable facial and linguistic emotional responses. Specifically, *electrodermal activity* (EDA) and *inter-beat interval* (IBI) data collected from Empatica's E4 wristband were used to infer affective states. While these basic metrics (e.g., EDA peak count, HRV RMSSD) may be sensitive to motion artifacts, our static, seated setup and use on the non-dominant hand helped ensure signal quality. Given their efficiency and validation in prior studies using E4 device [96, 103], we found them suitable for system deployment.

- *Emotional arousal level*, which indicates how much a user feels excited, has been studied to be associated with the phasic component of the time-series EDA response [8, 12, 21]. A well-known method to determine the degree of temporal arousal is by counting the number of peaks in the phasic component within a given time window, where more peaks are associated with higher emotional arousal levels. Following the method proposed by Kim et al. [60], we processed the EDA data by separating the phasic component, reducing noise through convolution with a Bartlett filter, detecting peaks based on zero-crossing detection, and finally counting the number of peaks within a 60s time window.
- *Stress level*, which indicates how stressed a user is, was calculated using the IBI signals the device provided. Variation between IBIs is known as heart rate variability (HRV), a well-recognized stress reactivity indicator widely used in HCI research [15, 57]. We employed a method to measure HRV by calculating the root mean square (i.e., RMSSD;

$\sqrt{T^{-1}\sum^{T}|\mathrm{IBI}_{t-1}-\mathrm{IBI}_{t}|^{2}}$) of consecutive IBIs, as it is suitable for calculating short-term HRV. A lower RMSSD indicates higher stress. We calculated the average successive RMSSD within a 60s time window using IBI.

After extracting each level, we calculated the median absolute deviation (MAD) and removed outliers exceeding the range of $\mathrm{median}(x) \pm 3\mathrm{MAD}$. Before normalizing the calculated values, we considered individual differences in physiological indicators among users. To do so, we used data from the meditation session to anchor one side of the normalization range, assuming it reflects a calm baseline. For HRV, we used the maximum from meditation and the minimum from the entire data, including conversation sessions; for EDA peaks, the minimum from meditation and the maximum from the entire data. We then applied min-max normalization using these bounds. The normalized affect levels were ordinally discretized into four levels: *low* (0.0–0.25), *moderate* (0.25–0.5), *high* (0.5–0.75), and *very high* (0.75–1.0).