Exploring Socially Acceptable 3D Feedback Design for Inclusive Small-Group Conversation

QIANJIE WEI, University of Rochester, USA
YIFAN LI, University of Rochester, USA
ZHEN BAI*, University of Rochester, USA

Prior research has shown that providing feedback on group dynamics can improve the quality of small-group conversations. However, publicly delivered feedback that targets specific individuals may reduce social acceptability due to negative experiences such as social pressure, embarrassment, and resistance. This work investigates how public feedback can be designed to remain socially acceptable while maintaining its effectiveness in XR-based small-group conversations. Informed by social science and HCI literature, we identified three design strategies (incorporate humor, reduce social attention, and use indirectness) and proposed a set of metaphor- and agent-based 3D feedback probes. We then conducted a two-phase study: (1) an online survey with 50 participants to examine how such feedback designs are perceived under public and private delivery, and (2) follow-up interviews with 14 participants who completed the survey to better understand their interpretations and elicit additional insights. Results show that metaphor- and agent-based feedback can reduce social discomfort while maintaining effectiveness. Interview findings further highlighted the role of three strategies in making feedback easier to accept, while requiring careful balance. We also discussed design implications for creating socially acceptable 3D feedback to support inclusive small-group conversations in XR.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Collaborative and social computing**; *Virtual reality*.

Additional Key Words and Phrases: Small-group Conversation, Public Feedback, Inclusiveness

ACM Reference Format:
Qianjie Wei, Yifan Li, and Zhen Bai. 2027. Exploring Socially Acceptable 3D Feedback Design for Inclusive Small-Group Conversation. *Proc. ACM Hum.-Comput. Interact.* 11, 1, Article GROUP18 (January 2027), 24 pages. https://doi.org/10.1145/3859404

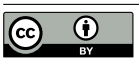
1 Introduction

Imagine you are in a conversation with two other peers. One person keeps talking and no one else can get a word in. Or maybe you notice that someone is upset, but don't know how to bring them back into the conversation. In moments like these, a feedback [65] delivered in front of everyone may help the conversation get back on track effectively. But it creates a dilemma: the feedback may also single someone out, making them feel exposed, judged, or embarrassed.

Public feedback, i.e., feedback shown to everyone in the group [57], can be especially powerful in group settings: it can increase user motivation [26, 50], improve team awareness [26], and address

*Corresponding author

Authors' Contact Information: Qianjie Wei, Department of Computer Science, University of Rochester, Rochester, New York, USA, qwei3@ur.rochester.edu; Yifan Li, Department of Computer Science, University of Rochester, Rochester, New York, USA, yli287@ur.rochester.edu; Zhen Bai, Department of Computer Science, University of Rochester, Rochester, New York, USA, zhen.bai@rochester.edu.



This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2027 Copyright held by the owner/author(s).
ACM 2573-0142/2027/1-ARTGROUP18
https://doi.org/10.1145/3859404

problematic behavior effectively due to social comparison [13] and social influence [3]. However, delivering feedback that targets individuals in front of the entire group may lead to negative psychological feelings, such as embarrassment or even resistance [2, 36], making public feedback less acceptable. Private feedback, defined as feedback shown only to the person to whom it applies [57], can avoid this psychological discomfort [32, 48], but it is often easy to be ignored and not effective [3, 56].

A substantial body of research in HCI and CSCW has investigated on providing feedback about group dynamics and supporting inclusive group conversations [4, 32, 37, 40, 67, 72], both in co-located and distributed settings. Existing feedback systems such as CoCo [68], "Observe, Ask, Intervene" [32], MeetingCoach [67], and SealMates [4] demonstrate that visualized conversational cues and behavioral patterns (e.g., speaking time, turn-taking, emotional tone) has a positive effect on the conversation. There are also emerging research of extended reality (XR), which encompasses virtual reality (VR), augmented reality (AR), and mixed reality (MR), in supporting group conversational dynamics such as turn-taking [39, 42] and balanced participation [86]. While research shows that users in XR-mediated group interaction also experience social discomfort associated with public feedback [77], there is a gap of research on exploring design space of public feedback in XR environments to reduce such discomfort.

XR-based communication provides more possibilities for the use of digital representations (e.g., agents) [45] and contextual cues (e.g., embodied metaphors) [45, 82], which demonstrates the potential to offer novel feedback forms in group conversation. Inspired by this, we proposed our research question (RQ): **How could we design socially acceptable and effective public feedback for supporting inclusive small-group conversations in XR environments?**

To address this RQ, we first ground our design in theories and findings from social science and HCI literature, and synthesize three strategies for reducing embarrassment in public settings, including **incorporating humor** [11, 23, 36, 62], **reducing social attention** [11, 27, 66], and **using indirectness** [12, 31, 61, 72]. Based on these strategies and XR's potential to provide novel feedback forms, we proposed a series of metaphor- and agent-based 3D feedback designs across three conversational scenarios in group conversations (someone is dominating, someone is under-participating, someone is in negative emotion). To understand how such feedback designs are perceived in small-group conversations, we first conducted an online survey involving 50 participants, systematically comparing metaphor- and agent-based 3D feedback with direct text messages through the design probe method [25, 28]. Subsequently, we conducted follow-up interviews with 14 participants to gain a deeper understanding of their interpretations.

Our findings indicated that metaphor- and agent-based feedback was associated with lower perceived embarrassment and greater acceptance than the direct text messages, while maintaining usefulness and fairness. Interview results suggested that incorporating humor, reduced attention, and indirectness can reduce embarrassment in public feedback. Participants indicated that humor incorporation is effective when it is playful yet appropriate, reduced social attention can be supported by disconnecting feedback from individuals (e.g., through environmental cues or redirecting attention to an agent), and indirectness usage benefits from a balance between being subtle and remaining understandable. They also envisioned additional socially acceptable 3D feedback designs and interaction mechanisms to support inclusive small-group conversations.

The main contribution of our work lies in exploring how public feedback can be designed to be more socially acceptable while keeping its effectiveness for small-group conversations in XR: (1) We identified a set of design strategies, proposed corresponding 3D feedback probes to mitigate embarrassment caused by public exposure, and applied them to three conversational scenarios. (2) We conducted a two-part user study, involving an online survey and follow-up interviews, to understand how people perceive various feedback and strategies in small-group conversations

and to explore any additional design opportunities. (3) We further discussed design implications to inform the future work of inclusive feedback systems for small-group conversations in XR environments.

2 Related Work

2.1 Public versus Private Feedback in Small-Group Conversations

In the context of group communication, providing feedback on performance for the group members has been proven as a powerful way to shape group dynamics [68, 80]. Feedback occupies a major portion of the total effort and serves an anchoring function in communication process [71] in small-group conversations (i.e., three to ten equal-status people engaged in unstructured conversations [6]). Prior work has noted that the impact of feedback is influenced by its visibility [57, 80], which can be presented either privately (i.e., shown only to the person to whom it applied) or publicly (i.e., shown to everyone in the group) [57].

Public feedback supports social comparison [22] and social influence [3], which is effective for addressing problematic behavior [13] and has a positive effect on facilitating performance [1, 17]. Despite the benefits, public corrective feedback often elicits embarrassment [2, 36]. It can single individuals out in front of others, therefore heightening concerns about negative evaluation and loss of face. **Private feedback** is less likely to generate the social discomfort [32] by not pointing individual out to the entire group [48]. However, prior literature also noted that private feedback has less social influence [3] and is less effective in facilitating performance because of the absence of observers [56]. Overall, these prior work highlights the trade-off between public and private feedback: while public feedback can leverage social influence to facilitate performance, it may also expose individuals to embarrassment about negative evaluation; private feedback can reduce such social discomfort, but may be less effective due to its limited visibility.

2.2 Existing Feedback Systems in HCI

Prior research has extensively explored how to provide feedback on group dynamics across various settings, including video conferencing, social VR, and in-person communication [5, 68, 69, 72, 73]. Existing work has made efforts to provide feedback on a range of conversational features, including participation [4, 5, 43, 68], turn-taking [16, 76], interruption [69], agreement [33, 40], and emotional valence [18, 74]. Meanwhile, researchers also investigated how different feedback formats shape users' experience. For example, Schiavo et al. [72] conducted a study comparing overt, text-based directives with subtle directives delivered through graphical displays, and examined their effects on group dynamics. Beyond text and visual graphics, another playful feedback format is introducing agent as an external mediator [4, 44]. Liu et al. [44] has demonstrated that a biosignal-driven agent could support remote communication between romantic distributed partners. Similarly, Armstrong et al. [4] introduced a behavior-drive agent to promote balanced participation in video conferencing.

As an emerging medium for collaboration, XR enables group members to interact in both remote and co-located settings [24, 53, 83, 85]. Researchers have explored several approaches to provide feedback on conversational dynamics in XR. For instance, Lee et al. [39] presented a multimodal attention-guidance method to facilitate turn-taking in VR-based group conversations. Yang et al. [86] designed a virtual moderator to promote group members' balanced participation. Despite technological interventions, the social discomfort associated with public feedback may also persist in XR-based group conversations. Slater et al. compared small-group behavior in the virtual and real environment, and found that socially conditioned responses such as embarrassment could also be generated in the virtual meeting [77].

Although existing studies have explored how to promote inclusive conversations through technological assistance, most of them deliver feedback in a public way while paying little attention to the negative effects of public exposure, such as embarrassment, social pressure, and resistance [2, 17, 27, 36]. Acknowledging the trade-off between public and private feedback, our work delves deeper into how public feedback in XR-based small-group conversations can be designed to be more socially acceptable while keeping its effectiveness.

2.2.1 XR to Support Novel Feedback Forms. XR provides more possibilities for the use of digital representations (e.g., embodied agents) and contextual cues [45, 84], which demonstrates its opportunity for providing novel feedback forms in group conversations. **One promising direction is the use of embodied agents as feedback carriers.** Prior work has identified multiple types of agent representations, including authentic, non-anthropomorphic, and other novel forms [53]. The appearance and behavior of these virtual characters can shape users' perceptions and behaviors in XR environments [53]. Animal-like agents, in particular, present as an emerging application in both fully and partly immersive systems. For example, Norouzi et al. [54] described a virtual dog in AR and found that its simulated behaviors affected participants' perception, locomotion, and social interaction with others. Similarly, Na et al. [52] leveraged a virtual cat in MR to reduce people's mental stress and induce positive emotions.

Another relevant direction is the use of embodied metaphors in XR. Metaphors express ideas in a familiar and acceptable substitution for a wide range of people [12, 19, 20, 46]. Building a metaphor helps users understand abstract concepts by relating them to something familiar or tangible [10, 35]. Recent XR studies have begun to leverage this potential. For example, Wang et al. [82] showed that the metaphor of "breaking the walls" in VR could support individual creative performance. Other work has used metaphorical interaction to help users cope with negative emotions in XR environments [29, 63]. Rasch et al. [63], for instance, enabled users to engage with representations of their thoughts by gradually organizing an inner space through the metaphor of cleaning the room".

2.3 Design Strategies: Reduce Embarrassment in Social Contexts

To inform more socially acceptable feedback design, several approaches to reduce embarrassment in social contexts were identified as our feedback design strategies. **Strategy 1: Incorporate humor.** Researchers have documented humor as an effective strategy for alleviating embarrassment in interpersonal communication [11, 23, 36]. Humor is a common remedial strategy in embarrassing interactions involving equal status individuals [23]. Press and Erel [62] showed that even peripheral humorous interventions (like robotic gestures) can significantly reduce awkwardness between strangers. Playfulness is considered the basis of humor [64].

Strategy 2: Reduce social attention. Empirical evidence suggests that social attention can play a causal role in amplifying awkwardness: Clegg [11] found that inducing explicit social attention would increase reported feelings of social awkwardness. This aligns with the spotlight effect, in which people overestimate how much others notice their actions and appearance, particularly in embarrassing situations [27]. As a potential remedy, the appearance of a new object may capture and direct people's social attention [87, 88], thereby reducing their perceived social exposure.

Strategy 3: Use indirectness. Prior HCI work shows that subtle interaction is typically described as socially acceptable with less embarrassment [61]. Schiavo et al. [72] also found that subtle directives can influence group behavior and are more positively accepted, compared to overt, text-based directives. This rationale is consistent with linguistic and pragmatics research: indirect expressions, especially euphemisms, contribute to saving the public self-image of people in communication [12, 31]. Metaphor is the most central device in the formation of euphemism

[12, 19], "in which one thing is referred to by a word or expression normally denoting another thing, idea, or action, so as to suggest some common quality shared by the two". In light of these prior work, we formulated our Strategy 3.

3 Designing 3D Feedback as Probes

Guided by the design strategies and emerging feedback forms in XR, we adopted the design probe approach [25, 28] and proposed a series of metaphor- and agent-based feedback probes for three conversational scenarios, aiming to mitigate the social discomfort caused by public exposure (Figure 3). Drawing inspiration from prior research on human and virtual cat interaction in XR [52], we selected a cat as the embodied agent for our agent-based feedback designs. The textual messages were also included as a baseline, reflecting the typical and direct feedback commonly used in existing systems [72]. Section 3.1 explains the rationale for selecting the conversational scenarios, while Section 3.2 describes how the feedback probes were developed, which we characterize by ***Text (Baseline)***, ***Metaphor,*** and ***Cat***.

3.1 Conversational Scenarios Selection

Both **balanced participation** and **positive emotions** have been identified as important and recurring themes in prior work on group conversations [32, 43, 67, 69]. Dominance and under-participating behavior of one or a few members are common in collaboration [38, 58, 60] and can negatively affect group interaction by limiting perspectives diversity and reducing opportunities for inclusive collaboration [70]. Beyond participation balance, prior work has also examined the role of emotional valence in group conversations. Negative emotions, such as sadness, anger, and anxiety, were viewed as potentially having a negative impact on decision-making or behavioral control in group conversations [21, 51]. Drawing on this line of work, we included three representative conversational scenarios in this study: **(1) someone is dominating the conversation**, **(2) someone is under-participating in the conversation**, and **(3) someone is in negative emotion during the conversation**. The first two scenarios are associated with imbalanced participation, while the third scenario concerns emotional valence.

3.2 Design Rationale

3.2.1 Text (Baseline). As for the Baseline, we chose explicit and straightforward textual messages, which are common found in existing systems [15, 32, 37, 72]. We created scenario-specific text messages presented on a 3D UI panel (*"You should include everyone in the conversation!"* for Scenario 1 [72], *"You are not participating enough. Say something!"* for Scenario 2 [72], *"It's okay to take a moment."* for Scenario 3), to compare with the following designs.

3.2.2 Metaphor. **Scenario 1 and 2: Participation imbalance.** Previous research has used the seesaw to represent imbalance [49, 59] (Strategy 3: Use indirectness). In a study designing visual metaphors for providing feedback to clinicians, Patel et al. [59] found that seesaw tilting offered an intuitive and straightforward way to identify who was dominating the conversation. Additionally, introducing a novel object can redirect people's attention away from the individual [87, 88] (Strategy 2: Reduce social attention). Moreover, the seesaw is associated with humor and playful interactions, particularly during moments of togetherness among peers [30] (Strategy 1: Incorporate humor). Here seesaw tilting was used to indicate participation imbalance in Scenario 1 and 2. Specifically, when someone is dominating the conversation, the increased participation was mapped onto increased weight on the seesaw. When someone is under-participating, this lower level of participation was represented by having that person's side rise upward.



Fig. 1. Overview of the feedback probes. We proposed ***Text (Baseline)***, ***Metaphor***, and ***Cat*** for three conversational scenarios: someone is dominating the conversation, someone is under-participating, and someone is experiencing negative emotion. Subfigures (a), (d), and (g) show the ***Text*** Baseline; (b), (e), and (h) show the ***Metaphor*** probes; and (c), (f), and (i) show the ***Cat*** probes.

		Strategy 1: Incorporate humor	Strategy 2: Reduce social attention	Strategy 3: Use indirectness
Scenario 1 & 2: Participation imbalance		The seesaw is associated with playful and humorous interactions.	— Not intentionally incorporated.	The seesaw metaphorically represents participation imbalance.
		Cat behaviors can create playful moments and elicit laughter.	The cat may redirect attention away from the feedback receiver.	Feedback is conveyed indirectly through the cat's behavior.
Scenario 3: Negative emotion		— Not intentionally incorporated.	The sunflower may redirect attention away from the feedback receiver.	The sunflower symbolically conveys optimism and emotional support.
		— Not intentionally incorporated.	The cat may redirect attention away from the feedback receiver.	Emotional support is conveyed indirectly through the cat's behavior.

Fig. 2. Overview of how the feedback probes reflect the three design strategies across the conversational scenarios. Not every probe incorporates all three strategies. An em dash indicates that a strategy was not intentionally included in the corresponding probe.

Scenario 3: Negative emotion. For Scenario 3, the sunflower was chosen as a metaphorical representation for providing emotional support. In everyday language, sunflowers are seen as a symbol of optimism, happiness, and good luck [41, 75] (Strategy 3: Use indirectness). Beyond its

metaphorical meaning, the sunflower also functions as an attentional cue: as a novel visual element, it may draw social attention away from the individual experiencing negative emotion [87, 88] (Strategy 2: Reduce social attention).

3.2.3 Cat. Research on companion animals shows that pets' behaviors can provide a greater source of laughter for people [78] (Strategy 1: Incorporate humor). In Scenario 1, the cat knocks over nearby objects to attract the dominant participant's attention and subtly remind them to make space for others. In Scenario 2, the cat turns away and then looks back toward the under-participating participant, using this gesture as a gentle invitation to engage in the conversation. In Scenario 3, when someone is in negative emotion during the conversation, a virtual cat will start headbutting and seek for petting. In all the scenarios, the feedback is conveyed indirectly through the cat's behavior rather than through an explicit message (Strategy 3: Use indirectness). By directing attention toward the cat, these behaviors may also reduce the extent to which the feedback recipient becomes the focus of the group's attention, helping them feel less singled out (Strategy 2: Reduce social attention).

4 Survey

To gather participants' initial perceptions of these feedback probes, we conducted an online survey. The survey was distributed through Reddit communities related to XR and group communication, to reach a broader pool of population. The first 30 participants who completed the survey received $15 in compensation, while the remaining participants were entered into a raffle for a chance to receive $50.

4.1 Survey Participants

A total of 50 participants completed the survey. Participants ranged in age from 19 to 44 years (M = 29.0, SD = 6.3). 29 participants identified as male (58%) and 21 identified as female (42%). In terms of race and/or ethnicity, participants identified as Black or African American (42%), White (32%), Asian (20%), Hispanic or Latino (4%), and other racial or ethnic backgrounds (2%).

4.2 Survey procedure

The survey was hosted on the Qualtrics online platform and took approximately 30 minutes to complete. It is composed of screening, information sheet, demographic information and the main survey. First, the participants answered three screening questions about age, ability to read from a computer screen with normal vision, and ability to understand basic written English. Eligible participants were able to view and download an information sheet that provides a detailed description of the survey study (a waiver of documented consent was approved by the Institutional Review Board). Participants who agreed to participate then proceeded to the remaining sections.

Demographic information gathered information about participants' age, gender, race and/or ethnicity, prior experiences in small-group conversations, and prior usage of XR technologies.

Main survey consisted of three scenarios and an open-ended question session (Figure 3). To mitigate potential order effects, the presentation order of the three scenarios was randomized for each participant. At the beginning of the survey, participants were asked to imagine that they were taking part in a small-group conversation with two other peers. For each scenario, participants first watched a short video depicting the situation and then evaluated three feedback designs (***Text***, ***Metaphor***, and ***Cat***). Each design was shown under two visibility conditions: public and private (Figure 4). Specifically, in the private mode, only the participant could see the feedback, while in the public mode, all group members could see it.

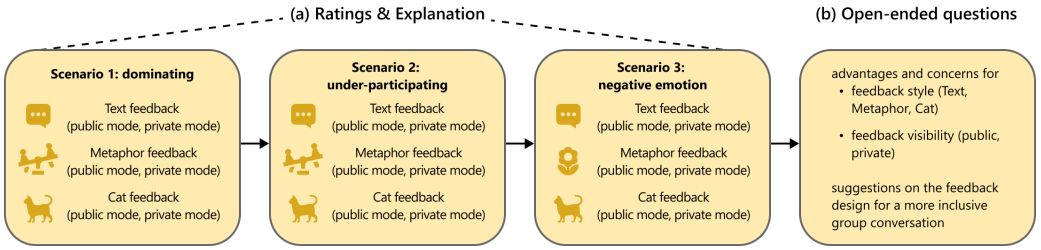


Fig. 3. Flowchart of the main survey procedure. (a) Participants were presented with three small-group conversation scenarios. For each scenario, they provided ratings and written explanation for three feedback styles—Text (Baseline), Metaphor, and Cat—with each style shown in both public and private modes. (b) Participants answered open-ended questions about the perceived advantages and concerns of different feedback styles and visibility, and offered suggestions for designing socially acceptable feedback to support more inclusive group conversations.



Fig. 4. Illustration of feedback visibility in the survey (take Cat feedback as an example). We presented side-by-side views in the survey, displaying both the participant's own perspective (your view) and another group member's perspective (Tom's view). (a) In the public mode, the cat is visible in both your view and Tom's view. (b) In the private mode, the cat is only visible in your view and does not appear in Tom's view.

After viewing each design, participants rated their perceptions using 5-point Likert-scale. We measured their perceived usefulness (*"It would help reduce dominant participation/encourage quieter speakers to contribute more/ease negative emotions during the conversation."*) [15, 37, 67], fairness (*"It would help us participate more openly and fairly."*) [37], comfort (*"I would feel comfortable receiving it during a conversation."*) [37], embarrassment (*"Receiving this feedback would make me feel embarrassed."*) [14], and distraction (*"I would be distracted by it."*) [72]. We additionally included questions asking participants what aspects of the design contributed to reducing embarrassment. Specifically, we included three items: Strategy 1 (*"It would be playful and easier to accept."*), Strategy 2 (*"It would avoid putting me at the center of attention."*), and Strategy 3 (*"It would be indirect and help me save my face."*). After completing three scenarios, participants answered open-ended questions about perceived advantages, concerns, and design suggestions for the solutions. The metrics and corresponding survey questions are listed in Appendix A.

4.3 Data Collection and Analysis

We conducted a validation process to remove responses that have duplicated IP addresses and incomplete results. The quantitative survey ratings were analyzed using descriptive statistics and inferential testing. As our data was collected on a 5-point Likert scale and each participant evaluated all the feedback probes, we conducted Friedman tests [47] to compare ***Text***, ***Metaphor***, and ***Cat*** for each dependent variable (usefulness, fairness, comfort, embarrassment, and distraction).

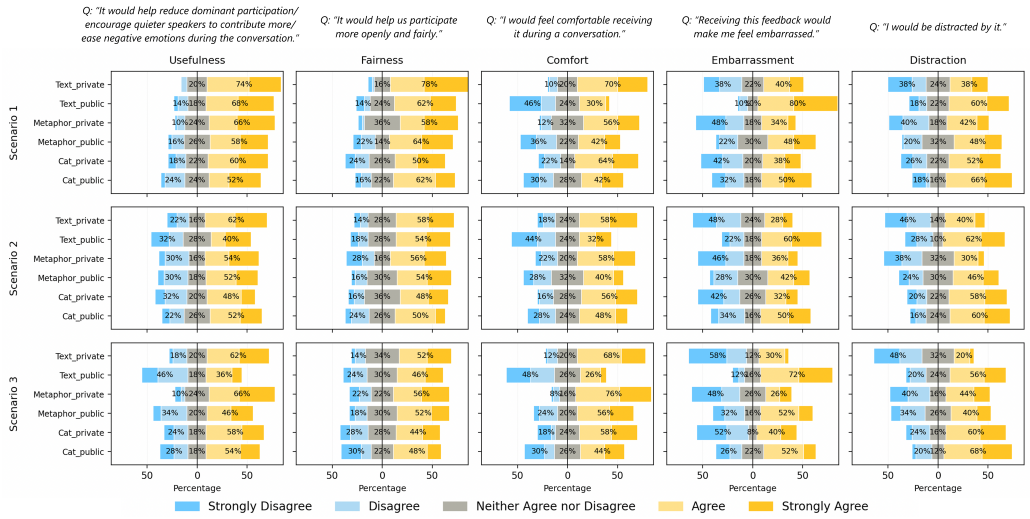


Fig. 5. Under private and public conditions, participants' ratings of ***Text***, ***Metaphor***, and ***Cat*** across three scenarios on usefulness, fairness, comfort, embarrassment, and distraction. Bars show the percentage distribution of responses on a 5-point Likert scale.

For the qualitative open-ended responses, we conducted thematic analysis [9]: two researchers independently coded participants' responses, and three researchers then grouped similar codes into broader themes.

5 Survey Results

5.1 Public vs. Private Feedback Delivery

Across all the feedback designs and scenarios, participants consistently preferred private over public delivery on socially sensitive dimensions. As shown in Figure 5, public delivery elicited higher **embarrassment** and lower **comfort** than private delivery for each design. This directly supports our motivation: publicly delivered feedback is more likely to introduce social discomfort in small-group conversations. A similar contrast emerged for **distraction**. Across all designs, public delivery was perceived as more distracting than private delivery, which suggests that making feedback visible to the group may interrupt the conversation flow.

5.2 Preference for Text, Metaphor, and Cat

Figure 5 summarizes participants' ratings of the ***Text*** (baseline), ***Metaphor***, and ***Cat*** probes across three scenarios, reporting perceived usefulness, fairness, comfort, embarrassment, and distraction under both private and public conditions. Overall, these results highlight that metaphor- and agent-based designs can reduce some of the public-delivery costs.

5.2.1 Usefulness and Fairness. In Scenario 1 and 2, *Text_private* received the highest positive ratings, with average percentage of positive responses (agree and strongly agree) of Scenario 1 at 74% and Scenario 2 at 62%. Whereas in Scenario 3, *Metaphor_private* was rated the most useful (66%). Fairness ratings were generally positive or neutral across ***Text***, ***Metaphor***, and ***Cat*** under private and public conditions, with no clear differentiation across scenarios.

5.2.2 Comfort and Embarrassment. *Text_public* consistently received the highest embarrassment and lowest comfort ratings in every scenario (comfort: 26–32%; embarrassment: 60–80%). To further examine differences among the three public probes, we conducted Friedman tests for each scenario. For comfort, the Friedman test showed a significant difference among ***Text***, ***Metaphor***, and ***Cat*** ($\chi^2 = 6.561$, $p < .05$) in Scenario 1. The Wilcoxon signed-rank tests with Bonferroni correction showed that ***Text*** was rated as significantly less comfortable than ***Metaphor*** ($p < .05$) and ***Cat*** ($p < .05$). The difference between ***Metaphor*** and ***Cat*** was not significant ($p = 1.000$). A significant difference in comfort was also found in Scenario 3 ($\chi^2 = 18.536$, $p < .001$). Post-hoc comparisons showed that ***Text*** was rated as significantly less comfortable than ***Metaphor*** ($p < .001$), whereas the differences between ***Text*** and ***Cat*** ($p = .113$), and between ***Metaphor*** and ***Cat*** ($p = .184$), were not significant. For embarrassment, the Friedman test showed a significant difference among the three public probes in Scenario 1 ($\chi^2 = 14.683$, $p = .001$). Post-hoc analysis showed that ***Text*** was rated as significantly more embarrassing than ***Metaphor*** ($p = .001$) and ***Cat*** ($p < .005$). The difference between ***Metaphor*** and ***Cat*** was not significant ($p = 1.000$).

5.2.3 Distraction. Overall, public delivery was perceived as more distracting than private delivery. Under public delivery, ***Text*** and ***Cat*** were rated as the most disruptive (*Text_public*: 60% in Scenario 1, 62% in Scenario 2, 56% in Scenario 3; *Cat_public*: 66% in Scenario 1, 60% in Scenario 2, 68% in Scenario 3)., while ***Metaphor*** was consistently rated as the least distracting (*Metaphor_public*: 48% in Scenario 1, 46% in Scenario 2, 40% in Scenario 3). The Friedman test showed a significant difference among the three probes in Scenario 3 ($\chi^2 = 10.161$, $p < 0.05$). Post-hoc analysis showed that ***Metaphor*** was rated as significantly less distracting than ***Cat*** ($p < 0.05$). There was no significance between ***Text*** and ***Metaphor*** ($p = 0.091$), as well as ***Text*** and ***Cat*** ($p = 1.000$) in this scenario.

5.2.4 In-depth Feedback. In open-ended survey responses, participants described perceived advantages, concerns, and suggested improvements associated with the feedback probes (see Appendix B for specific themes and participant counts). ***Text*** (baseline) was valued for clarity and directness, yet perceived as more likely to cause embarrassment. ***Metaphor*** was seen as less intrusive but sometimes ambiguous or insufficiently informative. ***Cat*** was described as funny, attractive, emotionally low-pressure, and as helpful for easing awkwardness in public conditions, but limited by lower information density for serious contexts. Participants also suggested offering flexible options that allow users to choose their preferred probe incorporating multimodal cues such as haptic pulses and voice reminders.

6 Interview

To gain a deeper understanding of participants' survey responses and elicit additional suggestions, the research team invited a subset of survey participants to take part in follow-up interviews.

6.1 Interview Participants

We recruited follow-up interview participants from those who had completed the survey. Participants were invited if they met the following criteria: (1) they selected "yes" to the question asking whether they were interested in a future interview, and (2) they provided relatively detailed and thoughtful responses to the survey's open-ended questions.

A total of 14 participants took part in the follow-up interviews (gender: 6 female, 8 male; age: median = 23.5, mean = 24.64, SD = 6.37). As shown in Table 1, 9 participants identified as Asian, 2 as Black or African American, 2 as White, and 1 as Hispanic or Latino. Most participants reported engaging in small-group conversations (e.g., meetings, discussions, and social gatherings) at least once a week, with only one participant reporting participation less than once a week. Regarding

Table 1. Demographic information of the participants (N=14). Each participant's background information was summarized based on their responses in the survey.

ID	Age	Gender	Race and/or Ethnicity	Frequency of Small-group Conversations	Prior XR Experience
P1	24	Male	Asian	At least three times a week	yes
P2	20	Female	Asian	At least one time a week	yes
P3	44	Male	White	Almost every day	yes
P4	24	Male	Asian	At least three times a week	no
P5	19	Female	White	Less than one time a week	no
P6	27	Female	Asian	At least three times a week	yes
P7	27	Male	Black or African American	At least once a week	yes
P8	23	Male	Black or African American	At least once a week	yes
P9	19	Male	Asian	At least three times a week	yes
P10	22	Female	Asian	At least three times a week	yes
P11	23	Female	Asian	At least three times a week	yes
P12	27	Female	Asian	At least one time a week	no
P13	19	Male	Hispanic or Latino	At least one time a week	yes
P14	27	Male	Asian	At least one time a week	yes

prior XR experience, 11 participants reported having prior experience, whereas 3 reported no prior experience. Each participant was compensated $15 for their time.

6.2 Interview Procedure

The follow-up interviews were conducted remotely over Zoom, with each session lasting approximately 40-60 minutes. Each session involved one participant and two researchers: one researcher leading the interview, while the other researcher took notes and provided assistance as the participant needed. Before the interview began, we obtained participants' consent for audio recording and reminded them that they could revise or add nuance to their survey responses during the interview. We also informed participants that their opinions during the interview did not need to be consistent with the ratings they had given in the survey, and they were free to share any additional, updated, or more nuanced thoughts. The interview consisted of three parts, as shown in appendix C. The interview consisted of three parts, including **1) revisiting public versus private feedback in small-group conversations**, **2) discussing *Text, Metaphor,* and *Cat* feedback probes**, and **3) eliciting additional design suggestions**. First, participants discussed their views on public versus private feedback based on their earlier survey responses. Second, participants explained their perceptions of these feedback designs across several dimensions, including usefulness, fairness, distraction, comfort, and embarrassment. We also asked how they interpreted the design strategies underlying these probes. Third, participants provided additional design suggestions for the socially acceptable 3D feedback design. The full interview protocol is provided in Appendix C.

6.3 Data Collection and Analysis

All audio recordings were transcribed into text paragraphs using Zoom's automatic transcription function and then manually corrected. In analyzing the data, we adopted the thematic analysis approach. Two researchers first independently reviewed the transcripts and assigned initial codes to meaningful segments of the data. They then compared coding results and iteratively grouping them

into higher-level categories and themes. Three researchers further discussed the coding outcomes, refined the thematic structure, and reached consensus on the final themes.

7 Interview Results

In general, the interview results extended the survey findings by revealing more nuanced and contextual considerations and clarifying how the feedback probes were perceived to reduce embarrassment and could be further refined.

7.1 Contextual Preferences for Feedback Visibility

Participants emphasized that preferences for public versus private feedback depend on contextual factors, including feedback valence, group chemistry, and behavioral context. **Feedback valence.** Participants noted that positive feedback is more socially acceptable and encouraging when shared publicly (P1, P3, P5), such as visible reactions like *"thumb-up"* or *"claps"* (P1). Similarly, P3 stated that *"positive reinforcement would work well in public feedback"* because *"everybody knows you're being encouraged"*, while more critical feedback should remain private to avoid discomfort. **Group chemistry.** Participants mentioned that public feedback feels more acceptable in groups with higher level of familiarity and trust (P3, P5). For example, P3 noted that receiving feedback in public is easier *"with friends that you trust"*, whereas *"it can scare the individual and hinder improvement in unfamiliar settings"*. **Behavioral context.** Participants highlighted that feedback visibility should adapt to the context of behavior. Public feedback was preferred when behaviors disrupted group progress or required collective awareness (e.g., off-topic discussion), whereas private feedback was seen as more appropriate for sensitive situations such as low participation due to stress or lack of confidence (P1). Participants also noted that feedback visibility could be influenced by individual preferences, social roles and power dynamics (P11), and formal or informal settings (P10).

7.2 Perceptions of Text, Metaphor, and Cat

7.2.1 Usefulness and Fairness. Under public delivery condition, participants generally perceived ***Text***, ***Metaphor***, and ***Cat*** as effective in helping users adjust their behavior and supporting fair and open conversation. ***Text (Baseline)*** was viewed as a direct form of feedback, allowing users to *"understand what to do next immediately"* (P1, P5, P6). **Metaphor-based designs** were also considered useful for addressing participation imbalances (P2, P10) and providing emotional support (P4, P5, P8, P10). The seesaw was regarded as effective reminder because it offered a *"perspective change"* (P2), while the sunflower blooming *"creates a pleasant atmosphere"* (P4), *"brings the sense that something good is happening"* (P8), and *"is a soft, implicit, but powerful universal cultural symbol"* (P3). **Cat-based designs** were described as capable of *"breaking the dominant person's flow of conversation"* (P2), *"encouraging people to participate in the conversation"* (P3, P5), and easing negative emotions because of pets' emotional support characteristics. (P4, P6, P9). Despite these positive perceptions, some participants considered the metaphor- and cat-based designs indirect or somewhat vague, noting that they might take time to interpret or could confuse users (P3, P5, P14).

7.2.2 Comfort and Embarrassment. ***Text (Baseline)*** was often perceived as socially intrusive and emotionally harsh, with participants describing it as *"cold"*, *"like a warning"* (P2), *"a sense of blame and criticism"* (P10), and embarrassing when publicly visible (P3, P4, P5, P6, P8, P9, P11, P14). However, participants consistently emphasized that **metaphor- and cat-based designs** provide a more comfortable and socially acceptable experience compared to ***Text*** in public delivery. This was largely attributed to their indirect and less confrontational nature (P1, P3, P4, P5, P6, P8, P9, P10, P14). As P1 said, *"Cat and Metaphor let me feel more comfortable, because they are expressed in an*

indirect and interesting way." P4 similarly noted, *"In the negative-emotion scenario, flowers and cat are both a soothing way, because they don't expose your negative emotions directly to others."*

7.2.3 Distraction. ***Text*** can be distracting and interruptive, as it demands cognitive effort to read text (P1, P6) and can disturb the flow of the conversation (P3, P5). ***Metaphor*** is less distracting due to its more subtle nature (P3, P4, P5, P6, P10), while animations of ***Cat*** can distract people from the conversations (P1, P2, P8, P9, P10, P13). Importantly, participants noted that distraction is not necessarily a negative thing. For example, distraction can be beneficial when it interrupts undesirable behavior, such as helping to *"distract the person who is dominating the conversation"* (P2). Additionally, agent-based elements such as the cat can create shared conversational entry points, with P3 describing how it *"gives other people a perfect excuse to enter the conversation,"* such as commenting on the cat itself, thereby facilitating turn-taking. In negative emotion situations, distraction could be *"a kind of distraction and break"* when experiencing negative emotions (P9).

7.3 Perceptions and Extensions of Design Strategies for Reducing Embarrassment

Participants reflected on how design strategies were realized in the proposed designs and contributed to reducing embarrassment caused by public exposure. Using the proposed ***Metaphor*** and ***Cat*** probes as starting points, participants also envisioned additional socially acceptable 3D feedback designs and interaction mechanisms.

7.3.1 Strategy 1: Incorporate Humor. Participants highlighted humor as an important strategy for reducing embarrassment by softening the tone and making it easier to accept. Both metaphor- and cat-based designs were perceived as incorporating humor through their *"funny and interesting"* presentation, which avoids treating feedback as *"a very serious or negative thing"* (P1). P9 expressed that *"Seesaw is humorous and playful because it's an activity which we would conduct in the playground or park in our daily lives"*. Similarly, some participants expressed that cat feedback can remind people in friendly way and are easier to accept because the cat is adorable (P1, P3, P4, P6) and common in life (P2, P5). *"It is cute and adorable and it will serve as a friendly way to make me notice the problem"* (P6). However, some participants also suggested a need of moderate humor in representation, as if it is purely humorous, it will have negative impact on the meeting content (P1).

Participants envisioned that humor can be further realized through more expressive and engaging design elements (Fig. 6). For metaphor-based designs, participants proposed incorporating more imaginative and lighthearted elements (e.g., a sunflower with glasses as shown in Fig. 6(b)) to create a playful tone (P10). For agent-based feedback, they emphasized playful and interactive behaviors, such as exaggerated animations or actions (e.g., a cat carrying an encouraging object as shown in Fig. 6(c)), which make feedback feel lighter and more approachable (P2, P3, P8, P10).

7.3.2 Strategy 2: Reduce Social Attention. Participants discussed reducing attention as a way to mitigate embarrassment by shifting focus away from individuals. While some subtle representations (e.g., small metaphorical changes) were described as *"small and not so distinct sign"* (P6) and *"not a big issue"* (P9), participants noted that current designs of metaphor still carry a pointing effect, making it obvious who the feedback is directed at. For example, visible changes such as position change or noticeable animations can still draw group attention to a specific individual, limiting their ability to reduce social pressure (P1, P6).

Some participants suggested alternative approaches that disconnect feedback from the individual, such as through environmental cues or by redirecting attention to an agent (Fig. 7). For metaphor-based designs, participants proposed feedback as environmental or ambient changes (e.g., color, lighting, weather, or scene transitions), allowing users to perceive feedback without being directly pointed out (P1, P2, P5, P8). For agent-based designs, participants suggested that the agent could

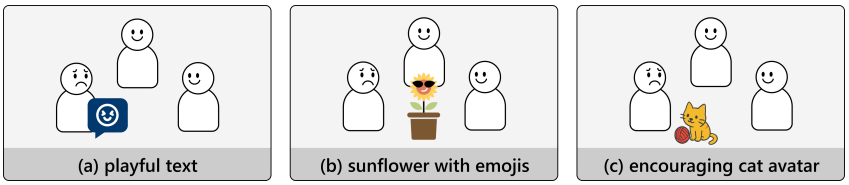


Fig. 6. Participants envisioned how **Strategy 1 (Incorporate humor)** can be realized and extended through other designs, illustrated by coauthor. When someone is in negative emotion during the conversation, (a) a joke is told through text message; (b) a sunflower with glasses and emojis to provide emotional support; (c) a cat carries a small ball next to the person who has negative emotion, drops the ball with encouragement.

delivers feedback while attracting attention to itself. For example, the agent may take on roles such as a social facilitator or an additional group member, coordinating turn-taking and guiding participation (P6). By drawing attention toward the agent rather than any individual, this approach helps reduce social attention directed at individuals.

7.3.3 Strategy 3: Use Indirectness. Participants also discussed indirect feedback as a way to mitigate embarrassment by conveying feedback in a less explicit manner. They described metaphor and cat-based feedback as *"indirect and interesting"* (P1) and *"friendly"* without making users feel *"offended or pressured"* (P2). However, participants also emphasized the need for a balance between indirectness and clarity (P5). As P5 suggested, *"We need something not too vague, but vague enough"* so that it takes other people some time to process what's going on. If feedback is too indirect, as P1 noted, *"it will confuse people"*, and others pointed out that metaphorical cues may require prior knowledge to interpret (P2, P8).

Participants envisioned that indirectness can be achieved by representing feedback through subtle forms rather than explicit indicators (Fig. 8). For metaphor-based designs, participants proposed using visual metaphors such as water levels, writing amount, hourglasses, or microphones to guide inclusive participation (P4, P5, P10, P13). For emotional states, participants suggested affective metaphors (e.g., temperature, color, lighting, or environmental scenes) to subtly convey feelings and create a more supportive atmosphere (P1, P2, P5, P8, P14). For agent-based designs, the agent's behavior or status may reflect participation or emotional conditions (P4, P9, P10, P14).

7.3.4 Additional Interaction Mechanisms. Participants also proposed additional mechanisms for delivering the feedback. **Temporal Delivery and Transition.** Participants mentioned that there should be a transition for delivering public feedback (Fig. 9). P14 proposed that feedback could be delivered gradually, allowing time for self-correction or increasing feedback intensity if the behavior continues. **Personalization.** Participants emphasized the importance of personalizing appearance to improve comfort and engagement. For instance, they suggested offering a range of alternative agents (e.g., different animals or characters) and allowing users to select their preferred option, which can make feedback feel more positive and less intrusive (P1, P3, P4, P8, P10).

8 Discussion

Based on the findings, we further discuss design implications for future socially acceptable feedback design in XR-based small-group conversations.

8.1 Constructing Embodied Agents and Metaphors as XR Feedback Forms

Previous work has explored using embodied agents and metaphors to shape users' perceptions and behaviors in XR environments [29, 52–54, 63, 82]. Our work leveraged this potential and

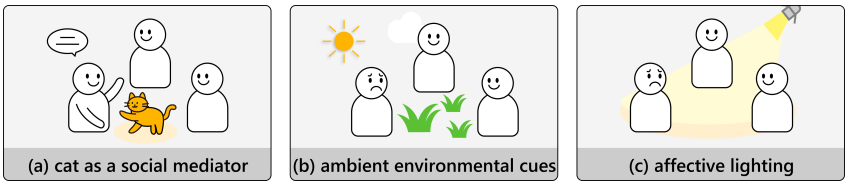


Fig. 7. Participants envisioned how **Strategy 2 (Reduce social attention)** can be realized and extended through other designs, illustrated by coauthor. (a) When someone is dominating the conversation, a cat acts as a social facilitator in the group, coordinating turn-taking and guiding participation. (b) When someone is in negative emotion, use environmental cues such as sunlight and grassland to create a more positive atmosphere. (c) When someone is in negative emotion, use color to help participants feel better.

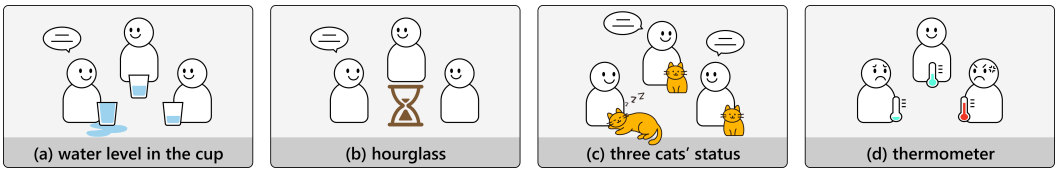


Fig. 8. Other designs envisioned by participants about how **Strategy 3 (Use indirectness)** can be realized and extended. (a) Use the water level in a cup to represent each member's engagement. (b) An hourglass to signal that a dominant participant's speaking time is running out. (c) There are three cats in the group, and each cat's status reflects the corresponding group member's participation level. (d) Use a thermometer to represent each member's emotional status.

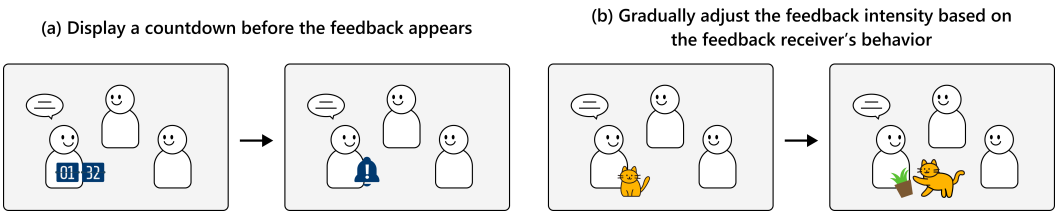


Fig. 9. Participants also mentioned that there should be a transition for delivering public feedback. Here we take Scenario 1 (someone is dominating the conversation) as an example. (a) A countdown is shown before the public feedback appears, allowing the feedback receiver time to make changes or get prepared for the upcoming reminder. (b) The feedback begins soft and becomes stronger if the feedback receiver continues the dominant behavior. Specifically, the cat simply wags its tail to remind the person. If the dominant person is still talking after a while, the cat will exhibit more aggressive behaviors like knocking something over.

constructed embodied agents and metaphors as novel feedback forms in XR. Participants responded positively to this. They described metaphor- and cat-based feedback as useful for encouraging behavioral adjustment, less confrontational, and more supportive (Section 7.2). These responses suggest that feedback forms may shape how a reminder is socially interpreted. For example, participants interpreted the cat's behavior as an invitation to participate or a source of emotional support, giving the feedback a more approachable tone. Participants also envisioned a broader range of metaphorical representations, agent behaviors, and environmental cues (Section 7.3). Overall, these findings suggest that designing XR feedback involves considering both what a cue communicates and how its presentation shapes the experience of receiving it in front of others. The

same conversational state could be expressed through different metaphors, agents, or environmental cues. Future work could benefit from examining how different feedback forms affect both perceived social acceptability and actual conversational outcomes in live, multi-user XR settings.

8.2 Calibrating Humor, Indirectness, and Attention Redirection

Humor Should Soften Feedback Without Trivializing the Situation. Previous research has identified humor as a remedial strategy for reducing social embarrassment [11, 23, 36, 62]. Consistent with prior work, our findings show that humor may provide group members with a socially acceptable way to acknowledge the corrective message. However, participants also mentioned that excessive humor could undermine the seriousness of the situation (Section 7.3.1). This concern aligns with prior work about understanding the appropriateness of humor in group collaboration, which demonstrates that humor during serious conversation or goal-directed talk like a debate, lesson or business meeting could be perceived as intrusive, interruptive, or a waste of time [55, 81]. Future work should adopt humor with care and only after considering the conversation context. Subtle playfulness may be sufficient for routine participation reminders [7, 8], whereas explicitly humorous behavior may be excessive and distracting during serious contexts [55, 81].

Decoupling Feedback from Individuals to Reduce Social Exposure. Prior work on the role of explicit social attention in amplifying embarrassment [11, 27, 66] informed our Strategy 2: Reduce social attention. Participants, however, noted that a new object could still expose a person if other members could easily infer whom it targeted. This finding adds an important nuance to the strategy: reducing social attention may require decoupling the feedback from a specific individual, rather than merely introducing a new element to the shared space. Building on participants' proposals for environmental cues and agents acting as social facilitators (Section 7.3.2), future work could examine whether group-level environmental changes or agent behaviors can redirect attention without singling out an individual, while keeping the intended feedback understandable.

Indirectness should preserve face without obscuring the intended action. Previous HCI research has found subtle directives to be more socially acceptable than overt text-based instructions [72], while linguistic research suggests that indirect expressions can protect people's public self-image [12, 31]. Our findings similarly indicate that indirect metaphors and agent behaviors are perceived to reduce the confrontational quality of public feedback. However, indirectness introduced a tension between social comfort and interpretability: if the mapping between a cue and the expected behavioral adjustment is too weak, users may be uncertain about whether the feedback applies to them or how they should respond. Research on indirect communication further suggests that interpreting such cues often depends on shared cultural knowledge and communicative conventions [34]. Future work should carefully calibrate the degree of indirectness. One possible way to reduce confusion is to familiarize users with the meanings of feedback cues in advance, for example through onboarding, a shared legend, or optional explanations.

8.3 From Fixed to Flexible Feedback Delivery Mechanisms

Our findings suggest that appropriate feedback delivery depends on the social context of the conversation. Participants considered public feedback more acceptable for positive reinforcement and in groups with greater familiarity and trust, while preferring private feedback for more sensitive situations, such as low participation associated with stress or lack of confidence (Section 7.1). These preferences suggest a tension between supporting shared awareness and protecting individual comfort. Making feedback visible to the group may help coordinate participation, but it may also expose a person whose behavior calls for a more sensitive response. Participants' suggestions for gradual delivery and personalized feedback forms offer possible ways to accommodate these different needs (Section 7.3.4). This aligns with prior work on customizable and context-relevant

feedback [67, 68], as well as user control over how non-verbal cues are shared [79]. Future work could explore how feedback visibility, form, timing, and intensity can adapt to conversational contexts while allowing users to configure these decisions. Such systems should consider how to support group coordination while respecting recipients' preferences about public exposure.

9 Limitations and Future Work

One limitation is that our study assessed participants' perceptions of the 3D feedback probes rather than during actual group conversations. Although the survey and follow-up interviews provided valuable insights into our work, participants evaluated visual sketches combined with text descriptions rather than experiencing these probes in a real-time interactive setting. Moreover, the study focused on the perspective of the feedback receiver and did not examine how other group members might perceive feedback directed at someone else. Future work could evaluate both receiver and spectator perspectives and investigate how these feedback designs influence participants' real-time conversational behavior and their perceptions in XR environments.

Second, as our study focused on a relatively limited set of scenarios and design probes, the generalizability of the results may be limited by the study context. We examined three conversational scenarios and a small number of feedback styles, which does not cover all the possible situations and feedback designs that may arise in small-group conversations. Additionally, co-located and distributed groups may respond differently to the feedback design due to differences in social presence and communication cues. Future work could extend this work by exploring a wider range of conversational scenarios, feedback styles, and interaction contexts.

Third, racial and ethnic diversity does not fully capture cultural factors relevant to the social acceptance of conversational feedback. We did not systematically examine participants' country of residence/origin and cultural background, which limits the cross-cultural generalizability of findings. Future work should investigate how these factors shape feedback preferences and acceptance.

Finally, there was s time gap of approximately one month between the survey and interviews, and some participants developed different views on the relevant design in the later session. While this provided more nuanced insights for our work, it also suggests that participants' perceptions were not entirely stable across different sessions. Future work could shorten this time gap or more systematically explore how participants' perceptions evolve over long-term use.

10 Conclusion

In this paper, we investigated how public feedback in XR-based small-group conversations can be designed to remain socially acceptable while preserving its effectiveness. Drawing on social science and HCI literature, we identified three design strategies for reducing embarrassment in social contexts: incorporating humor, reducing social attention, and using indirectness. Guided by these strategies and novel digital representations in XR, we developed a set of metaphor- and agent-based 3D feedback probes for three conversational scenarios: dominance, under-participation, and negative emotion. Through a two-phase study combining an online survey and follow-up interviews, our findings suggest that metaphor- and cat-based 3D feedback can mitigate some of the social discomfort associated with public feedback while maintaining perceived usefulness and fairness. Participants explained how design strategies made public feedback easier to accept, while also highlighting the trade-offs involved in applying them. Ultimately, this work contributes design implications for future socially acceptable 3D feedback in small-group conversations.

References

[1] Deborah G Ancona and David F Caldwell. 1992. Bridging the boundary: External activity and performance in organizational teams. *Administrative science quarterly* (1992), 634–665.

[2] W Peter Archibald and Ronald L Cohen. 1971. Self-presentation, embarrassment, and facework as a function of self-evaluation, conditions of self-presentation, and feedback from others. *Journal of Personality and Social Psychology* 20, 3 (1971), 287.

[3] Michael Argyle. 1957. Social pressure in public and private situations. *The Journal of Abnormal and Social Psychology* 54, 2 (1957), 172.

[4] Mark Armstrong, Chi-Lan Yang, Kinga Skiers, Mengzhen Lim, Tamil Selvan Gunasekaran, Ziyue Wang, Takuji Narumi, Kouta Minamizawa, and Yun Suen Pai. 2024. SealMates: Improving Communication in Video Conferencing using a Collective Behavior-Driven Avatar. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–23.

[5] Aadesh Bagmar, Kevin Hogan, Dalia Shalaby, and James Purtilo. 2022. Analyzing the effectiveness of an extensible virtual moderator. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–16.

[6] Robert F Bales, Fred L Strodtbeck, Theodore M Mills, and Mary E Roseborough. 1951. Channels of communication in small groups. *American Sociological Review* 16, 4 (1951), 461–468.

[7] John A Banas, Norah Dunbar, Dariela Rodriguez, and Shr-Jie Liu. 2011. A review of humor in educational settings: Four decades of research. *Communication Education* 60, 1 (2011), 115–144.

[8] Marino Bonaiuto, Elio Castellana, and Antonio Pierro. 2003. Arguing and laughing: The use of humor to negotiate in group discussions. *Humor: International Journal of Humor Research* 16, 2 (2003).

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

[10] John M Carroll. 2003. *HCI models, theories, and frameworks: Toward a multidisciplinary science.* Elsevier.

[11] Joshua W Clegg. 2012. Stranger situations: Examining a self-regulatory model of socially awkward encounters. *Group processes & intergroup relations* 15, 6 (2012), 693–712.

[12] Eliecer Crespo Fernández et al. 2005. Euphemistic strategies in politeness and face concerns. (2005).

[13] Pieternel Dijkstra, Hans Kuyper, Greetje Van der Werf, Abraham P Buunk, and Yvonne G van der Zee. 2008. Social comparison in the classroom: A review. *Review of educational research* 78, 4 (2008), 828–879.

[14] Hyo Jin Do. 2023. *Designing a conversational agent for promoting participation of under-contributing members in a small group chat discussion.* Ph. D. Dissertation. University of Illinois at Urbana-Champaign.

[15] Hyo Jin Do, Ha-Kyung Kong, Jaewook Lee, and Brian P Bailey. 2022. How should the agent communicate to the group? Communication strategies of a conversational agent in group chat discussions. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–23.

[16] Wen Dong, Bruno Lepri, Taemie Kim, Fabio Pianesi, and Alex Sandy Pentland. 2012. Modeling conversational dynamics and performance in a social dilemma task. In *2012 5th International Symposium on Communications, Control and Signal Processing*. IEEE, 1–4.

[17] Amy Edmondson. 1999. Psychological safety and learning behavior in work teams. *Administrative science quarterly* 44, 2 (1999), 350–383.

[18] Mohamed Ez-Zaouia, Aurélien Tabard, and Elise Lavoué. 2020. EMODASH: A dashboard supporting retrospective awareness of emotions in online learning. *International Journal of Human-Computer Studies* 139 (2020), 102411.

[19] Eliecer Crespo Fernández. 2006. The language of death: Euphemism and conceptual metaphorization in Victorian obituaries. *Finnish Journal of Linguistics* 19 (2006), 101–130.

[20] Eliecer Crespo Fernández. 2008. Sex-related euphemism and dysphemism: An analysis in terms of conceptual metaphor theory. *Atlantis* (2008), 95–110.

[21] Rebecca A Ferrer, Jennifer Tehan Stanley, Kaitlin Graff, William MP Klein, Nina Goodman, Wendy L Nelson, and Silvia Salazar. 2016. The effect of emotion on visual attention to information and decision making in the context of informed consent process for clinical trials. *Journal of Behavioral Decision Making* 29, 2-3 (2016), 245–253.

[22] Leon Festinger. 1954. A theory of social comparison processes. *Human relations* 7, 2 (1954), 117–140.

[23] Edward L Fink and Barbara A Walker. 1977. Humorous responses to embarrassment. *Psychological Reports* 40, 2 (1977), 475–485.

[24] Guo Freeman, Dane Acena, Nathan J McNeese, and Kelsea Schulenberg. 2022. Working together apart through embodiment: Engaging in everyday collaborative activities in social virtual reality. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–25.

[25] William W Gaver, Andrew Boucher, Sarah Pennington, and Brendan Walker. 2004. Cultural probes and the value of uncertainty. *interactions* 11, 5 (2004), 53–56.

[26] Susanne Geister, Udo Konradt, and Guido Hertel. 2006. Effects of process feedback on motivation, satisfaction, and performance in virtual teams. *Small group research* 37, 5 (2006), 459–489.

[27] Thomas Gilovich, Victoria Husted Medvec, and Kenneth Savitsky. 2000. The spotlight effect in social judgment: an egocentric bias in estimates of the salience of one's own actions and appearance. *Journal of personality and social psychology* 78, 2 (2000), 211.

[28] Connor Graham, Mark Rouncefield, Martin Gibbs, Frank Vetere, and Keith Cheverst. 2007. How probes work. In *Proceedings of the 19th Australasian conference on Computer-Human Interaction: Entertaining User Interfaces*. 29–37.
[29] Florian Grieger, Holger Klapperich, and Marc Hassenzahl. 2021. Trash it, punch it, burn it–using virtual reality to support coping with negative thoughts. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–6.
[30] Bryndis Gunnarsdottir and Amanda Bateman. 2025. Moments of togetherness: humour and playful interactions in the toddler peer group. *Early Years* 45, 3-4 (2025), 655–670.
[31] Zahra Hussain Hadi. [n. d.]. A Pragmatic Study of Metaphorical Euphemistic Expressions in Shakespeare's Hamlet Muhammad Hussain Hamza. ([n. d.]).
[32] Mo Houtti, Moyan Zhou, Loren Terveen, and Stevie Chancellor. 2025. Observe, Ask, Intervene: Designing AI Agents for More Inclusive Meetings. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–18.
[33] Jeroen Janssen, Gijsbert Erkens, and Gellof Kanselaar. 2007. Visualization of agreement and discussion processes during computer-supported collaborative learning. *Computers in human behavior* 23, 3 (2007), 1105–1125.
[34] Cynthia Joyce. 2012. The impact of direct and indirect communication. *The Newsletter of the International Ombudsman Association. The University of Iowa* (2012).
[35] Heekyoung Jung, Heather Wiltse, Mikael Wiberg, and Erik Stolterman. 2017. Metaphors, materialities, and affordances: Hybrid morphologies in the design of interactive artifacts. *Design Studies* 53 (2017), 24–46.
[36] Dacher Keltner and Brenda N Buswell. 1997. Embarrassment: its distinct form and appeasement functions. *Psychological bulletin* 122, 3 (1997), 250.
[37] Soomin Kim, Jinsu Eun, Changhoon Oh, Bongwon Suh, and Joonhwan Lee. 2020. Bot in the bunch: Facilitating group chat discussion by improving efficiency and participation with a chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–13.
[38] Bibb Latané, Kipling Williams, and Stephen Harkins. 1979. Many hands make light the work: The causes and consequences of social loafing. *Journal of personality and social psychology* 37, 6 (1979), 822.
[39] Geonsun Lee, Dae Yeol Lee, Guan-Ming Su, and Dinesh Manocha. 2024. "may i speak?": Multi-modal attention guidance in social vr group conversations. *IEEE Transactions on Visualization and Computer Graphics* 30, 5 (2024), 2287–2297.
[40] Gilly Leshed, Diego Perez, Jeffrey T Hancock, Dan Cosley, Jeremy Birnholtz, Soyoung Lee, Poppy L McLeod, and Geri Gay. 2009. Visualizing real-time language-based feedback on teamwork behavior in computer-mediated groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 537–546.
[41] Aulia Nur Fitri Lestari. 2023. Decoding the Denotative and Connotative Elements in Sony Pictures Animation's Short Film "Hair Love"(2019). *JETLEE: Journal of English Language Teaching, Linguistics, and Literature* 3, 2 (2023), 68–86.
[42] Jialang Victor Li, Max Kreminski, Sean M Fernandes, Anya Osborne, Joshua McVeigh-Schultz, and Katherine Isbister. 2022. Conversation balance: A shared vr visualization to support turn-taking in meetings. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 1–4.
[43] Xiaoyan Li, Naomi Yamashita, Wen Duan, Yoshinari Shirai, and Susan R Fussell. 2023. Improving non-native speakers' participation with an automatic agent in multilingual groups. *Proceedings of the ACM on Human-Computer Interaction* 7, GROUP (2023), 1–28.
[44] 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.
[45] Jiayin Liu, Jean-Marie Burkhardt, and Todd Lubart. 2023. Boosting creativity through users' avatars and contexts in virtual environments—a systematic review of recent research. *Journal of Intelligence* 11, 7 (2023), 144.
[46] Todd I Lubart and Isaac Getz. 1997. Emotion, metaphor, and the creative process. *Creativity research journal* 10, 4 (1997), 285–301.
[47] Gill Marshall and Leon Jonker. 2011. An introduction to inferential statistics: A review and practical guide. *Radiography* 17, 1 (2011), e1–e6.
[48] Meredith Ringel Morris, Anthony Cassanego, Andreas Paepcke, Terry Winograd, Ann Marie Piper, and Anqi Huang. 2006. Mediating group dynamics through tabletop interface design. *IEEE computer graphics and applications* 26, 5 (2006), 65–73.
[49] P Mrstik. 2026. You can't scaffold a see-saw: Epistemic imbalance, role identity, and motivation in AI-mediated collaborative learning.
[50] Jonathan Mumm and Bilge Mutlu. 2011. Designing motivational agents: The role of praise, social comparison, and embodiment in computer feedback. *Computers in Human Behavior* 27, 5 (2011), 1643–1650.
[51] David Murungi, Martin Wiener, and Marco Marabelli. 2019. Control and emotions: Understanding the dynamics of controllee behaviours in a health care information systems project. *Information Systems Journal* 29, 5 (2019), 1058–1082.
[52] Heewon Na, Soyeon Park, and Suh-Yeon Dong. 2022. Mixed reality-based interaction between human and virtual cat for mental stress management. *Sensors* 22, 3 (2022), 1159.

[53] Huyen Nguyen and Tomasz Bednarz. 2020. User experience in collaborative extended reality: overview study. In *International Conference on Virtual Reality and Augmented Reality*. Springer, 41–70.

[54] Nahal Norouzi, Kangsoo Kim, Myungho Lee, Ryan Schubert, Austin Erickson, Jeremy Bailenson, Gerd Bruder, and Greg Welch. 2019. Walking your virtual dog: Analysis of awareness and proxemics with simulated support animals in augmented reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 157–168.

[55] Neal R Norrick and Alice Spitz. 2008. Humor as a resource for mitigating conflict in interaction. *Journal of pragmatics* 40, 10 (2008), 1661–1686.

[56] Hisashi Ohtsuki, Yoh Iwasa, and Martin A Nowak. 2015. Reputation effects in public and private interactions. *PLoS computational biology* 11, 11 (2015), e1004527.

[57] Kaitlyn M Ouverson, Alec G Ostrander, Jamiahus Walton, Adam Kohl, Stephen B Gilbert, Michael C Dorneich, Eliot Winer, and Anne M Sinatra. 2021. Analysis of communication, team situational awareness, and feedback in a three-person intelligent team tutoring system. *Frontiers in Psychology* 12 (2021), 553015.

[58] Mark T Palmer. 1989. Controlling conversations: Turns, topics and interpersonal control. *Communications Monographs* 56, 1 (1989), 1–18.

[59] Rupa A Patel, Andrea Hartzler, Wanda Pratt, Anthony Back, Mary Czerwinski, and Asta Roseway. 2013. Visual feedback on nonverbal communication: a design exploration with healthcare professionals. In *2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops*. IEEE, 105–112.

[60] Sherry L Piezon and William D Ferree. 2008. Perceptions of social loafing in online learning groups: A study of public university and US Naval War College students. *International Review of Research in Open and Distributed Learning* 9, 2 (2008), 1–17.

[61] Henning Pohl, Andreea Muresan, and Kasper Hornbæk. 2019. Charting subtle interaction in the HCI literature. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–15.

[62] Viva Sarah Press and Hadas Erel. 2023. Humorous robotic behavior as a new approach to mitigating social awkwardness. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.

[63] Julian Rasch, Michelle Johanna Zender, Sophia Sakel, and Nadine Wagener. 2024. Mind mansion: exploring metaphorical interactions to engage with negative thoughts in virtual reality. In *Proceedings of the 2024 ACM Designing Interactive Systems Conference*. 2305–2318.

[64] Willibald Ruch, Tracey Platt, René T Proyer, and Hsueh-Chih Chen. 2019. Humor and laughter, playfulness and cheerfulness: Upsides and downsides to a life of lightness. 730 pages.

[65] Jurgen Ruesch and Weldon Kees. 1956. Nonverbal Communication. Berkeley and Los Angeles. *Cal.: University of California Press* (1956).

[66] John Sabini, Michael Siepmann, Julia Stein, and Marcia Meyerowitz. 2000. Who is embarrassed by what? *Cognition & Emotion* 14, 2 (2000), 213–240.

[67] Samiha Samrose, Daniel McDuff, Robert Sim, Jina Suh, Kael Rowan, Javier Hernandez, Sean Rintel, Kevin Moynihan, and Mary Czerwinski. 2021. Meetingcoach: An intelligent dashboard for supporting effective & inclusive meetings. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–13.

[68] Samiha Samrose, Ru Zhao, Jeffery White, Vivian Li, Luis Nova, Yichen Lu, Mohammad Rafayet Ali, and Mohammed Ehsan Hoque. 2018. Coco: Collaboration coach for understanding team dynamics during video conferencing. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies* 1, 4 (2018), 1–24.

[69] Sanat Sarda, Martin Constable, Justin Dauwels, Shoko Dauwels, Mohamed Elgendi, Zhou Mengyu, Umer Rasheed, Yasir Tahir, Daniel Thalmann, and Nadia Magnenat-Thalmann. 2013. Real-time feedback system for monitoring and facilitating discussions. In *Natural Interaction with Robots, Knowbots and Smartphones: Putting Spoken Dialog Systems into Practice*. Springer, 375–387.

[70] Edward G Sargis and James R Larson Jr. 2002. Informational centrality and member participation during group decision making. *Group Processes & Intergroup Relations* 5, 4 (2002), 333–347.

[71] Thomas M Scheidel and Laura Crowell. 1966. Feedback in small group communication. *Quarterly Journal of Speech* 52, 3 (1966), 273–278.

[72] Gianluca Schiavo, Alessandro Cappelletti, Eleonora Mencarini, Oliviero Stock, and Massimo Zancanaro. 2014. Overt or subtle? Supporting group conversations with automatically targeted directives. In *Proceedings of the 19th international conference on Intelligent User Interfaces*. 225–234.

[73] Damaris Schmid, Dario Staehelin, Andreas Bucher, Mateusz Dolata, and Gerhard Schwabe. 2022. Does social presence increase perceived competence? Evaluating conversational agents in advice giving through a video-based survey. *Proceedings of the ACM on human-computer interaction* 6, GROUP (2022), 1–22.

[74] Nicklas Schneider, Ignacio Vazquez, and Peter A Gloor. 2024. No Pain, No Gain—Giving Real-Time Emotional Feedback in a Virtual Mirror Improves Collaboration in Virtual Teamwork. *Applied Sciences* 14, 13 (2024), 5659.

[75] Kenzie Scott. 2026. The Flower of Hope: Symbolizing Resilience and Optimism. (2026).

[76] Gabriel Skantze. 2021. Turn-taking in conversational systems and human-robot interaction: a review. *Computer Speech & Language* 67 (2021), 101178.
[77] Mel Slater, Amela Sadagic, Martin Usoh, and Ralph Schroeder. 2000. Small-group behavior in a virtual and real environment: A comparative study. *Presence* 9, 1 (2000), 37–51.
[78] Robin Maria Valeri. 2006. Tails of laughter: A pilot study examining the relationship between companion animal guardianship (pet ownership) and laughter. *Society & Animals* 14, 3 (2006), 275–293.
[79] Carlota Vazquez Gonzalez, Timothy Neate, and Rita Borgo. 2025. Trusting Tracking: Perceptions of Non-Verbal Communication Tracking in Videoconferencing. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–25.
[80] Marie Claire Villeval. 2020. Performance feedback and peer effects. *Handbook of labor, human resources and population economics* (2020), 1–38.
[81] Brian D Vivona. 2014. "To laugh or not to laugh". *The European Journal of Humour Research* 2, 1 (2014), 1–18.
[82] Xinyue Wang, Kelong Lu, Mark A Runco, and Ning Hao. 2018. Break the "wall" and become creative: Enacting embodied metaphors in virtual reality. *Consciousness and cognition* 62 (2018), 102–109.
[83] Qianjie Wei, Xiaoying Wei, Yiqi Liang, Fan Lin, Nuonan Si, and Mingming Fan. 2025. RemoteChess: Enhancing Older Adults' Social Connectedness via Designing a Virtual Reality Chinese Chess (Xiangqi) Community. In *Proceedings of the 2025 CHI conference on human factors in computing systems*. 1–16.
[84] Qianjie Wei, Jingling Zhang, Pengqi Wang, Xiaofu Jin, and Mingming Fan. 2024. Augmented library: toward enriching physical library experience using HMD-based augmented reality. In *Proceedings of the 17th International Symposium on Visual Information Communication and Interaction*. 1–5.
[85] Sinuo Wu, Karen Stendal, and Devinder Thapa. 2024. Emerging trends in XR-mediated virtual team collaboration in digital workspaces: a systematic literature review. *Advances in information systems development: Information systems development, organizational aspects, and societal trends* (2024), 85–108.
[86] Chi-Lan Yang, Keigo Matsumoto, Songping Yu, Leo Sawada, Kiyoaki Arakawa, Daisuke Yamada, and Hideaki Kuzuoka. 2023. Understanding the effect of a virtual moderator on people's perception in remote discussion using social VR. *Frontiers in Virtual Reality* 4 (2023), 1198024.
[87] Steven Yantis. 1993. Stimulus-driven attentional capture. *Current Directions in Psychological Science* 2, 5 (1993), 156–161.
[88] Steven Yantis and John Jonides. 1996. Attentional capture by abrupt onsets: new perceptual objects or visual masking? (1996).

A Survey questions

Metric	Statement
Usefulness	It would help reduce dominant participation/encourage quieter speakers to contribute more/ease negative emotions during the conversation.
Fairness	It would help us participate more openly and fairly.
Comfort	I would feel comfortable receiving it during a conversation.
Embarrassment	Receiving this feedback would make me feel embarrassed.
Distraction	I would be distracted by it.
Incorporate humor	It would be playful and easier to accept.
Reduce social attention	It would avoid putting me at the center of attention.
Use indirectness	It would be indirect and help me save my face.

Table 2. Questionnaire administered after viewing each feedback design probe. All items were rated on a 5-point Likert scale (1 = Strongly disagree, 5 = Strongly agree).

B Open-ended responses

Table 3. Participants' perceived advantages, concerns, and suggested improvements regarding feedback probes. The numbers in parentheses indicate how many participants mentioned each statement.

Probes	Advantages	Concerns	Suggested Improvements
***Text* (Baseline)**	direct (8), clear (5), precise (2), easy to understand (4), efficient (1)	easy to get embarrassed (2), not warm (1), devoid of emotion (1), very serious and less fun (1), too straightforward (2), can come across as blunt (1), judgmental (1)	*Provide flexibility.* Allow users to choose their preferred feedback probes (2). The receiving person may have preferences on what works and what doesn't (1). Make the 3D assets change dynamically depending on the situation, or even make them interactive (1). *Provide multimodal cues.* Voice reminder (1). Incorporate haptic pulses or spatial audio cues to ensure the feedback is accessible to everyone (1).
Metaphor	engaging (1), easy to accept (1), less intrusive (3), reduce hostility (1), not too straightforward (1), less confrontational (2), without lowering self esteem (1)	not informative enough (1), indirect (3), ambiguous (4), sometimes could not receive the hints (1), a bit hard or trickier to understand (2), might not be easily noticeable (1)	
Cat	playful and funny (3), hilarious (1), amusing (1), attractive (1), interesting (1), ease awkwardness (1), gentle (1), emotionally supportive (3), friendly (1), low-stress (1), non-threatening (1)	indirect (5), lack the directness needed in more serious situations (1), sometimes could not receive the hints (1), distracting (1)	

C Follow-up interview questions

C.1 Public vs. private feedback

1. [Interpret open-ended Responses for public feedback] In your responses, you mentioned that public feedback would _____(open-ended responses in survey)_____.

- Can you tell me more about your perceptions?
- What other advantages/disadvantages do you think public feedback has?
- What do you think public feedback would affect shared awareness or coordination in the group? (Ask if not discussed in previous questions)
- In what kinds of situations do you think public feedback would be helpful? And in what situations would public feedback not work?

2. Interpret open-ended Responses for private feedback: In your responses, you mentioned that private feedback would _____(open-ended responses in survey)_____.

- Can you tell me more about your perceptions?
- What other advantages/disadvantages do you think private feedback has?
- What do you think private feedback would affect the social comfort of seeing the feedback? Some people suggested that private feedback may be easier to ignore. What do you think? (Ask if not discussed in previous questions)
- In what kinds of situations do you think private feedback would be helpful? And in what situations would private feedback not work?

C.2 Under public condition, response to Text, Metaphor and Cat feedback designs

1. [Interpret Ratings for usefulness] I noticed that you rated that cat feedback is more useful than other feedback (or other feedback is more useful).

- Can you share your thoughts? Why do you think cat feedback is more useful?
- Why is text style less useful? Why is metaphor style less useful?
- [Scenario] Does that differ in different scenarios?
- [Are they understanding usefulness differently] When you evaluated perceived usefulness, what were you imagining would happen in the group?

2. [Interpret Ratings for fairness] I noticed that you rated that cat feedback would bring more fairness than other feedback (or that other styles would bring more fairness).

- Can you share your thoughts? Why do you think cat feedback brings more fairness?
- Why does the text style bring less fairness? Why does metaphor style bring less fairness?
- [Scenario] Does that differ in different scenarios?
- [Are they understanding fairness differently] When you evaluated perceived fairness, what were you imagining would happen in the group?

3. [Interpret Ratings for distracting] I noticed that you rated that metaphor feedback is more distracting than other feedback (or that other feedback is more distracting).

- Why: Can you share your thoughts?
- Why is the text style less distracting? Why is cat style less distracting?
- [Scenario] Does that differ in different scenarios?

4. [Interpret Ratings for comfort and embarrassment] I noticed that you rated that cat feedback makes you feel more comfortable and less embarrassed (or that other styles make you more comfortable).

- Can you share your thoughts? Why do you think cat feedback makes you feel more comfortable and less embarrassed?

- [Interpret strategy] You rated higher for humor for cat feedback (or other DGs). Do you think it reduces embarrassment because of humor? I wonder if you think that might influence embarrassment. Can you tell us more?
- Ask about indirectness and reduced social attention (feel less "called out")
- Why is the text style less comfortable?
- Why is metaphor style less comfortable? Ask about their ratings for humor, indirectness, and reduced social attention
- [Scenario] Does that differ in different scenarios?

5. Please rank the following six feedback types in order of your overall preference, from your most to least preferred. Briefly explain your ranking.

- Cue the three scenarios

6. [Interpret open-ended Responses for cat/metaphor/text feedback] In your responses, you mentioned that cat/metaphor/text feedback would _____(open-ended responses in survey)_____.

- Can you tell me more about your perceptions?
- In what kinds of situations do you think cat feedback would be helpful? And in what situations would cat feedback not work?

C.3 Additional design suggestions

1. [improvements for each design] Are there other metaphors, characters, or visual styles that might work better?

- e.g., cat behaviors

2. [other designs for each strategy] Beyond metaphor and cat styles, are there other ways to implement humor to reduce embarrassment?

- Any other ways to implement reduced social attention?
- Any other ways to implement indirectness?

3. [other strategies] Beyond the designs shown, are there any other strategies that could reduce embarrassment caused by public feedback?

- What kind of feedback can be implemented to realize this strategy?
- If they cannot think of any, ask in real life, how do you reduce embarrassment in small group conversations?

Received 31 March 2026; revised 17 July 2026; accepted 17 September 2026