

# Exploring a Real-time Feedback Display of Non-verbal Cues in Online Work Meetings to Support Self-Presentation

KEVIN CHOW\*, Department of Computer Science, University of British Columbia, Canada

ROY RUTISHAUSER\*, Department of Informatics, University of Zürich, Switzerland

ANDRÉ N. MEYER, Department of Informatics, University of Zürich, Switzerland

JOANNA MCGRENERE, Department of Computer Science, University of British Columbia, Canada

THOMAS FRITZ, Department of Informatics, University of Zürich, Switzerland

Expressing oneself appropriately in online meetings through non-verbal cues can be challenging for knowledge workers. Automatic non-verbal cue detection technologies have the potential to support workers' self-presentation efforts through real-time feedback, but little is known about workers' reactions to and the implications of doing so. We designed and implemented *Novecs* as a technology probe of a real-time feedback display that automatically detects and signals users' own non-verbal cues – smiling, nodding, gaze, and posture. *Novecs* was deployed in an exploratory field study (n=18) to support knowledge workers' self-presentation in their everyday meetings. Post-study interviews reveal how *Novecs*' real-time feedback helped increase in-the-moment self-awareness, and how neutrally-framed feedback may help navigate tensions between authentic and in-authentic self-presentation. Participants also emphasized the need for natural timing when adjusting non-verbal cues in-meeting. We discuss design opportunities and challenges of real-time, non-verbal cue feedback systems, such as personalizing feedback based on different meeting types.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; *Interactive systems and tools*; *Field studies*.

Additional Key Words and Phrases: videoconferencing, online work meetings, Zoom, self-presentation, knowledge worker, non-verbal cues, real-time feedback

## ACM Reference Format:

Kevin Chow, Roy Rutishauser, André N. Meyer, Joanna McGrenere, and Thomas Fritz. 2025. Exploring a Real-time Feedback Display of Non-verbal Cues in Online Work Meetings to Support Self-Presentation. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW071 (April 2025), 28 pages. <https://doi.org/10.1145/3710969>

## 1 Introduction

Meetings are a regular occurrence for most knowledge workers, but they come with a host of challenges. Some of these challenges are exacerbated in online meetings, which have become a cornerstone of remote and hybrid work. One such challenge is in non-verbal self-presentation [47], namely how workers convey information about themselves to others through non-verbal cues

\*Both authors contributed equally to this research.

---

Authors' Contact Information: [Kevin Chow](mailto:kchowk@cs.ubc.ca), [kchowk@cs.ubc.ca](mailto:kchowk@cs.ubc.ca), Department of Computer Science, University of British Columbia, Vancouver, BC, Canada; [Roy Rutishauser](mailto:rutis@ifi.uzh.ch), [rutis@ifi.uzh.ch](mailto:rutis@ifi.uzh.ch), Department of Informatics, University of Zürich, Zürich, Switzerland; [André N. Meyer](mailto:ameyer@ifi.uzh.ch), [ameyer@ifi.uzh.ch](mailto:ameyer@ifi.uzh.ch), Department of Informatics, University of Zürich, Zürich, Switzerland; [Joanna McGrenere](mailto:joanna@cs.ubc.ca), [joanna@cs.ubc.ca](mailto:joanna@cs.ubc.ca), Department of Computer Science, University of British Columbia, Vancouver, BC, Canada; [Thomas Fritz](mailto:fritz@ifi.uzh.ch), [fritz@ifi.uzh.ch](mailto:fritz@ifi.uzh.ch), Department of Informatics, University of Zürich, Zürich, Switzerland.

---

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

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2573-0142/2025/4-ARTCSCW071

<https://doi.org/10.1145/3710969>

such as facial expressions or body language in order to shape how they are perceived. Non-verbal cues are important for signaling engagement, communicating more effectively, building trust and rapport with colleagues, and resolving conflicts [45, 68, 81, 84, 87, 94, 100].

Online meetings have unique considerations — limited video field-of-view [9, 102], default-on self-view [69], latency [53], and multitasking [73] — that can require increased self-awareness and more intentional user effort, ultimately contributing to “non-verbal overload” [9] and “Zoom fatigue” [15, 40, 41, 101]. Online meeting attendees often need to exaggerate their nodding to ensure their signal of agreement is perceived [9, 24]. The work of knowledge workers typically involves the discussion and exchange of knowledge and ideas through meetings. As a result, they may face specific challenges in meetings, where their focus may be divided between often cognitively demanding work-related discussions and their self-presentation efforts [103]. Consciously monitoring their own self-view to maintain appropriate non-verbal cues in such meetings can be difficult; focusing too much on one’s own self-view may also distract from their work or even contribute to Zoom fatigue [69]. While self-viewing can enhance self-awareness [61, 79], on its own, it may not be able to help those who lack knowledge of but desire to exhibit appropriate non-verbal cue behaviour [3, 4]. Additionally, self-viewing by itself cannot intervene or notify users of their non-verbal cues in real-time, requiring constant monitoring for self-presentation.

There is an opportunity to leverage technological approaches to address the aforementioned challenges of self-awareness and cognitive load when it comes to non-verbal self-presentation in online meetings. Automatic non-verbal cue detection technology continues to mature, achieving robust, real-time processing speeds even on more generalized hardware like laptop webcams [11, 16, 74]. Interactive feedback systems that support the automatic detection and display of real-time feedback on one’s own non-verbal cues could help facilitate self-awareness and self-reflection on self-presentation during online work meetings.

While such real-time, non-verbal cue feedback systems may have the potential to support self-presentation in unique and powerful ways, there are open questions that should first be explored before expending effort into building robust systems [66]. Perspectives around self-presentation may be especially nuanced for knowledge workers in online work meetings, necessitating careful consideration when it comes to the design of the feedback display. Might users feel like they are being judged? Furthermore, the importance and appropriateness of non-verbal cue behaviour may depend on factors related to the type of meeting a knowledge worker is attending (e.g., one-on-one versus group meetings or who the other meeting attendees are) as well as individual differences regarding self-presentation. Real-time feedback systems, which are often designed to be glanceable, have known challenges around distractibility [70, 71, 96], but work meetings may present context-specific considerations and challenges. Privacy is another key concern — some knowledge workers already face intrusive, possibly privacy-impeding, monitoring at work [14, 25, 28], especially those working remotely — how might privacy be preserved in such systems?

Exploring these questions can further inform and inspire the design of future non-verbal cue feedback systems, and improve our understanding of the potential of such technologies for supporting knowledge workers. To do so, we adopted the use of a *technology probe*, which is a technological artifact that “[collects] information about the use and the users of the technology in a real-world setting” [57]. We felt that an in-situ approach, where knowledge workers could experience the technology in the cadence of their everyday online work meetings, was critical for surfacing the nuanced perspectives, opportunities, and challenges of real-time, non-verbal cue feedback systems.

Thus, we designed and implemented *Novecs* (short for **Non-verbal cue system**) as a technology probe [57] that enables the automatic detection and real-time feedback display of users’ own non-verbal cues during their online work meetings. *Novecs* targets four types of kinesic [48, 72] non-verbal cues, including two of facial expressions — smiling and nodding — and two of body

language — posture and gaze. While we are primarily interested in the potential of the real-time, in-meeting feedback, *Novecs* secondarily also includes summative, post-meeting feedback that visualizes users' non-verbal cues after each meeting. *Novecs* accounts for privacy concerns by processing all data *locally* on users' machines. In addition, we designed the feedback to be neutral and avoid being normative so that users would not feel judged by the probe and would be encouraged to reflect on their non-verbal behaviour in the moment. Altogether, these design decisions reflect our intention for *Novecs* to empower and support the user, rather than be used against them.

Our work aimed to answer the following research questions:

- **RQ1:** How do knowledge workers leverage real-time, non-verbal cue feedback in their various online work meetings?
- **RQ2:** How might the use of a real-time, non-verbal cue feedback system reveal and/or influence knowledge workers' perspectives on their non-verbal self-presentation in online work meetings?
- **RQ3:** What are the opportunities and challenges of a real-time, non-verbal cue feedback system for online work meetings to support non-verbal self-presentation?

To answer these research questions, we deployed *Novecs* in a multi-week exploratory field study with 18 knowledge workers, who used it in their real-world online work meetings. In the first, pre-feedback phase of the study, *Novecs* was deployed without signaling any feedback (only showing user presence/absence) for at least five meetings to allow participants to acclimate themselves to having it running with their regular work meetings. Subsequently, during the feedback phase, participants experienced *Novecs* in at least ten online work meetings each, while receiving in- and post-meeting feedback.

Findings from exit interviews show that *Novecs* helped increase participants' awareness of their non-verbal cues and supported their desire to express their non-verbal cues with appropriate and natural timing. *Novecs'* neutral and non-normative feedback design may have supported the surfaced tensions between authentic and in-authentic self-presentation. The perceived utility of *Novecs'* real-time feedback also varied based on different meeting types. We discuss the importance of authenticity and how participants valued being imperfect to be more “human”, as well as balancing noticeability and distractibility through a consideration of meeting types.

Overall, this paper makes the following contributions:

- **Empirical findings** from deploying the *Novecs* technology probe in-situ with knowledge workers' online work meetings that highlight the nuanced opportunities and challenges of a real-time feedback display of non-verbal cues, including tensions in supporting both authentic and in-authentic self-presentation.
- **Design opportunities** for neutral and non-normative real-time feedback of non-verbal cues, such as an increased consideration of meeting types, utilizing personas for customizing non-verbal cue feedback, and accepting imperfection in non-verbal cue behaviour.

## 2 Background and Related Work

In this section, we first provide a brief overview of non-verbal self presentation and the real-time, automatic detection of non-verbal cues. We then delve into related work on interactive non-verbal cue feedback systems for contexts outside of meetings, before concluding with systems designed specifically for online work meetings.

### 2.1 Self-Presentation and Non-verbal Cues

Self-presentation refers to the ways that one attempts to control how they are perceived by others [47]. While individuals engage in self-presentation efforts in everyday social interactions with those

around them, the unique implications of the work context make it particularly interesting – such as in the formation of one's professional identity and its role in organizational and team dynamics [18]. Self-presentation at work has been widely studied, including its impact on leadership, trust, and worker well-being [5, 12, 13, 33, 36, 44, 67, 83, 105]. Self-presentation efforts can be both controlled or automatic as well as authentic or inauthentic [18, 91]. Authentic self-presentation is when one's outer expressions are congruent with their inner state, whereas in-authentic self-presentation is when they are incongruent [47, 83]. An example of controlled, in-authentic self-presentation in a work meeting might be when one consciously attempts to look engaged and alert (such as through non-verbal cues like frequent nodding and sitting up straight) even when feeling tired. Non-verbal self-presentation [29] specifically involves non-verbal cues, such as facial expressions and body language. In this work, we focus on self-presentation in online work meetings as a large proportion of social interactions between work colleagues now occurs on videoconferencing platforms such as Zoom, Teams, or Google Meet. Online meetings can be especially challenging for self-presentation, as “users are forced to consciously monitor non-verbal behaviour and to send cues to others that are intentionally generated” [9], often leading to increased cognitive load and “Zoom fatigue” [15, 40, 41, 92, 101].

## 2.2 Automatic Detection of Non-verbal Cues in Real-time

Researchers have made advancements in enabling automatic, real-time detection of non-verbal cues, using both computer vision techniques as well as other sensing modalities (e.g., specialized sensors [64, 98], mobile phones and wearables, such as smartwatches or earables [24]). In particular, computer vision algorithms are now able to detect non-verbal cues in real-time, while using standard camera hardware like laptop webcams. Studies have explored the detection of the four non-verbal cues we chose to explore through *Novecs* in real-time – posture [30, 32], gaze [31], smiling [16, 56], and nodding [6, 95] – and also others like head shakes [62] and eye blinks [93]. Such “lower-level” non-verbal cues are specific, observable behaviours. They have been used in aggregate to help predict “higher-level constructs”, which broadly try to encapsulate the complexities of an individual's state of being (e.g., emotions). *Novecs*' four “lower-level” non-verbal cues are often associated with desirable “higher-level” constructs in the meeting context, such as engagement (posture, gaze), agreement (nodding), and positivity (smiling). Examples of aggregation to predict “higher-level constructs” include research in emotion recognition [32, 82], for detecting engagement when watching television, advertisements, and in the classroom [8, 52, 60, 78], detecting fatigue [59], and for mediating group discussions [23, 64].

In our work, *Novecs* maintains the use of computer-vision algorithms with standard laptop webcams (hence not requiring any additional hardware) to detect “lower-level” non-verbal cues rather than “higher-level” constructs. Predicting “higher-level” constructs can face additional challenges with accuracy and objectivity, especially across diverse use cases and cultural demographics [1]. Notably, feedback on “higher-level” constructs may be more difficult to act upon [26]. Some may struggle with grasping how they should change their behaviour to improve a “high-level” construct like engagement whereas specific “lower-level” non-verbal cues like posture are more easily acted upon, especially in real-time during an already cognitively demanding meeting [39]. “Lower-level” cues can also face challenges of accuracy and cognitive load in meetings, and we explore their potential for in-meeting feedback through *Novecs*.

## 2.3 Displaying Feedback on Non-verbal Cues Outside of Online Work Meetings

Existing work has utilized automatic detection technologies to enable interactive feedback systems for improving users' non-verbal self-presentation in a variety of contexts such as media or job interviews [51, 54, 84, 85], telehealth [39, 72], or just for general conversational skills in non-work

contexts [3, 4, 55, 86]. These existing systems resulted in some improvements in participants' self-presentation, suggesting the potential of such systems. However, the majority were evaluated in controlled lab settings and only provided *summative feedback* on users' non-verbal cues. Real-time feedback may have unique opportunities and challenges akin to those of glanceable displays, such as enabling in-the-moment awareness and adjustments but also challenges with distraction [70, 71, 96, 97].

Researchers in telehealth are among the few to explore *real-time feedback* of non-verbal cues to support clinician self-presentation in patient-clinician video calls, which can be considered a specific instance of an online work meeting. Hartzler et al. [49] conducted a design study with a Wizard-of-Oz real-time feedback system, demonstrating that an ambient, glanceable approach is acceptable to clinicians for displaying real-time information about their non-verbal cues. Faucett et al. [39] implemented ReflectLive, a functional system for patient-clinician video calls that includes glanceable, real-time, non-verbal cue feedback on clinicians' speaking contribution, eye gaze, conversational overlap, and screen centredness. Through a user evaluation with 10 clinicians, they found that ReflectLive increased self-awareness of eye gaze and was not overly distracting [39].

While *Novecs* also adopts a real-time feedback approach, it was designed for a more general context of online work meetings rather than patient-clinician video calls. There are key differences – in the telehealth context, patient-clinician video calls are fairly structured and well-defined: they are one-on-one, the clinician plays a specific role in the conversation and is the one that receives non-verbal cue feedback, and desirable non-verbal cue behaviour is well-studied (i.e., observational studies have linked non-verbal cues with desirable patient/clinical outcomes [39, 75]). On the other hand, a knowledge worker might attend a wide range of different types of work meetings even within a single work day, including one-on-ones or larger group meetings and with varying other meeting attendees. This variety in knowledge workers' meeting types suggests that desirable non-verbal self-presentation may differ from meeting to meeting, and is far more nuanced than in patient-clinician video calls, where a more normative approach might be appropriate. Knowing this, we deviated from existing work by deliberately taking a more neutral and non-normative approach to presenting feedback, in order to better accommodate the diverse range of possible meetings that workers might engage in.

## 2.4 Interactive Feedback Systems for Online Work Meetings

In the context of online work meetings, prior work has explored a variety of interactive systems for supporting knowledge workers, providing feedback on various factors, including both verbal and non-verbal cues, such as speaker participation [64, 89], attendee experience [7], inclusion and relational affiliation [27, 48], speaker consensus [88], and emotions [38, 43, 89]. Particularly relevant is Samrose et al.'s [88] MeetingCoach, as it includes a subset of the same non-verbal cues as *Novecs*. The tool provides an automated post-meeting dashboard that summarizes relevant information, including speaking turn and consensus, but also non-verbal cues like head nods and facial sentiment. They found that the MeetingCoach dashboard helped improve meeting attendees' awareness of meeting dynamics, and that they desired more “feedback assistance to make meetings more effective and inclusive” [88]. However, MeetingCoach, like much of the existing work on interactive feedback systems for non-verbal cues, only provides feedback in a summative fashion, leaving unknown the potential value of real-time feedback for promoting in-the-moment self-awareness of non-verbal cues. Real-time feedback allows users to not only be aware of and adjust their non-verbal cues while a meeting is taking place, but also supports reflection on their self-presentation *in-the-moment* of the meeting in a more granular manner.

To the best of our knowledge, we are one of the first to explore automatic detection and display of real-time, non-verbal cue feedback in a neutral, non-normative manner with knowledge

workers' real (thus highly varied) online work meetings to both support and better understand self-presentation in this context.

### 3 The *Novecs* Technology Probe

*Novecs* (short for **Non-verbal cue system**) is a technology probe [57] designed to explore automatic, real-time feedback for supporting non-verbal self-presentation in online work meetings. *Novecs* was designed to encourage an individual knowledge worker to personally reflect on and become more aware of their non-verbal cues during the varied range of online meetings they attend for work. Rather than enforcing specific behaviour and to be more inclusive, *Novecs* avoids overly normative design in non-verbal cue feedback and considers the individual's goals for their self-presentation. Our design goals evolved in conjunction with various rounds of piloting that we had conducted with graduate students and other knowledge workers who used multiple iterations of *Novecs* with their own meetings, with some using it up to a week. We also elicited feedback on the design of *Novecs* from an hour-long session with expert HCI researchers from one of the lead authors' institution. Piloting helped us to refine our final set of non-verbal cues, adjust interface and cue feedback design, and make improvements to the detection algorithm for each of the cues (see Section 3.2).

In the following sections, we first describe our rationale behind selecting the four non-verbal cues, how these cues are detected, how meeting sessions are managed, and end with how real-time feedback and summative feedback are displayed through the In-Session Sidebar and the Post-Session Summary window, respectively.

#### 3.1 Selection of Non-verbal Cues

We considered several factors in our selection of non-verbal cues for the *Novecs* technology probe, including visibility in a video feed as well as accuracy and technical feasibility without needing specialized hardware (i.e., can be run locally on laptops and laptop webcams). Overall, *Novecs* supports the sensing of four kinesic [48, 72] non-verbal cues, also known as movements of the head and body, including both facial expressions and body language: posture, gaze, smiling, and nodding. We focused on kinesic behaviours, as they are among the most salient in the teleconferencing context [72] – frequently expressed, changing within the field-of-view of the video feed, can be easily noticed by other meeting attendees, and also be easily changed by the user themselves. In addition, these four cues also all have a range of well-documented and accurate sensing algorithms in literature [6, 30–32, 56, 95], that are feasible with regular laptop webcams, crucial for deployment in the field in a study like ours.

We decided not to include other potentially relevant kinesic cues like eye blinks, eyebrow raises, or hand gestures, as they are either not very visually salient, have challenges with sensing accuracy [48, 72], or can be difficult for an individual to be aware of or change. Existing work has also looked at other relevant non-verbal cues like speaking time or tone, but analyzing those would require additional access to the audio stream of the online work meeting, and thereby potentially capturing meeting content. We chose not to do so to avoid intruding on participants' (and other meeting attendees') privacy and to facilitate recruitment.

#### 3.2 Detection of Selected Non-Verbal Cues

*Novecs* analyzes image frames with pre-trained machine learning models using OpenFace 2.0 [11] to detect non-verbal cues. We use pre-trained models to allow for the image processing to run locally on users' machines, and delete image frames after processing. No specialized hardware is

required, as the processing supports any regular built-in or external webcam. OpenFace [11] has built-in facial landmark detection that enables the calculation of presence and intensity information for action units in the Facial Action Coding System (FACS) [35]. It also outputs head position, rotation, and eye gaze information, all of which we use for detecting non-verbal cues. Occlusion and low illumination conditions present accuracy challenges for any vision-based system, and may be exacerbated in real-world scenarios. Despite these common limitations, OpenFace has been shown to be robust even in non-ideal lighting, when faces are non-frontal or partially occluded. OpenFace has also been evaluated on multiple datasets comprising samples of subjects with diverse ethnic backgrounds and skin colors (e.g., [65, 106]). *Novecs* is implemented as an Electron app with Node.js v16.17.x and supports macOS 12+. Detected non-verbal cue signals are stored in an SQLite database residing on users' local machine.

To later visualize the detected non-verbal cues to the user in the In-Session Sidebar (section 3.4) and Post-Session Summary (section 3.5), the sensed values need to be compared against a user's target threshold, their individualized ideal behavior for that non-verbal cue. Users personalize their target thresholds through two calibration steps, a per-individual and per-session calibration. The *per-individual* calibration is performed once with the initial setup of *Novecs* to capture per-user thresholds. Users can set these thresholds themselves to personalize the sensitivity of each feature (e.g., allowing for more or less mobility on their chairs). The *per-session* calibration is run at the beginning of each meeting to determine "target gaze and posture" for the particular meeting, and can also help to account for changes in different physical meeting environments.

The following list describes the sensing approach of the four non-verbal cues:

- **Posture:** 3D head position is used as a proxy for users' posture as we found from piloting that this works well as a low-cost, accurate estimation of posture. The distance between the target head position (from the per-session calibration) to the detected head position is evaluated against a user-defined threshold. *Novecs* detects non-target posture when the distance exceeds the threshold. The direction of non-target posture (backwards, forwards, sideways) is based on the direction in which the target and detected head positions differ.
- **Gaze:** Gaze is calculated by considering both eye gaze and head rotation in a similar manner as posture. Incorporating head rotation was found in piloting to produce a more robust detection of gaze, as users often turn their heads when looking in a different direction. Both the difference between target and detected eye gaze and head rotation are calculated and each compared against user-defined thresholds to produce an eye gaze and head rotation score. This score is then combined and compared against a system-defined threshold to decide if a non-target gaze signal (directionless) should be detected.
- **Smiling:** A smile is detected and displayed if the sum of the intensity of AU6 (cheek raiser) and intensity of AU12 (lip corner puller) [42, 56] is greater than a threshold for 500 milliseconds or longer. The threshold was defined based on piloting as the minimum time to detect a smile without too many false positives, e.g. when a person is speaking.
- **Nodding:** To detect nodding, our approach examines the pitch (vertical head rotation across the x-axis) of a set of frames in a time window of 1 second (typical duration of a nod according to [63]) to determine if a nod has occurred. Specifically, it checks this set of frames for a "high-low-high" pattern in the frame sequence, first by finding the lowest the head is rotated and subsequently checking for the highest points before and after. If both differences (from the lowest head rotation) exceed the user-defined nod angle threshold, and also a movement score (standard deviations of translation components of the head position) remains below a user-defined movement threshold, a "nod" signal is displayed. We found that accounting for the general movement of the head helped reduce false positives.

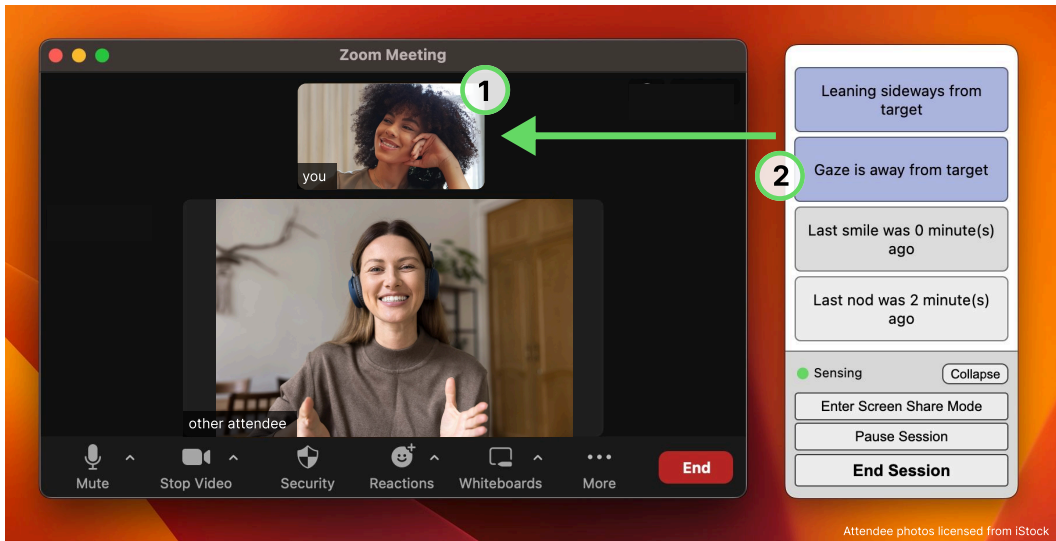


Fig. 1. Screenshot of the *Novecs* In-Session Sidebar (2) used in parallel to a Zoom video meeting. It indicates that the user ((1) denoted as ‘you’) is currently displaying non-target posture (leaning sideways, widget changed to blue), non-target gaze (closed eyes, widget changed to blue), is currently smiling (widget changed to grey, last smile counter reset to 0 minutes), and recently nodded (2 minutes ago).

### 3.3 Managing *Novecs* Sessions

For each meeting, users can decide if they want to manually start *Novecs*, either through the menubar or through a small reminder notification that briefly appears when Zoom or Microsoft Teams are active (i.e., window is in focus). When triggered, a pop-up window appears prompting users to estimate the length of the meeting (used for an end-of-meeting reminder notification) and choose the webcam for the meeting. In this paper, we use “session” to refer to when the *Novecs* probe is enabled for a meeting. As *Novecs* does not ask for admin access (to avoid further restricting potential participants), it does not have access to whether or not a webcam is being used and by which applications, meaning that *Novecs* cannot automatically start or stop a session based on the status of an online meeting and whether or not the user’s webcam is on or off.

### 3.4 In-Session Sidebar

Once a *Novecs* session is started, the In-Session Sidebar appears and initiates the *per-session* calibration step after a short 3-second countdown. This calibration step, which prompts users to maintain their “target posture and gaze”, lasts until ten calibration frames are captured (the duration of which will depend on the performance of the user’s machine; calibration times typically range from 2-5 seconds). In the study, we explained to participants that they themselves are free to define what their own “target” posture and gaze was (i.e., their own ideal or preference).

The In-Session Sidebar (see Figure 1) consists of four vertically stacked feedback widgets, one for each of the four non-verbal cues, along with controls for collapsing the sidebar (collapsing only makes it smaller to save space), entering/exiting screen sharing mode for users to toggle when sharing their screen (turns sidebar into miniature version to avoid sharing the feedback with other meeting attendees), pausing the session (pauses non-verbal cue detection), and ending the session. The In-Session Sidebar is also implemented to remain always-on-top while a session is started to display glanceable feedback. While an ideal implementation would have the widgets



embedded into one's own self-view on video-conferencing platforms, we did not pursue this option for technical reasons and to allow participants to use their own choice of video-conferencing platforms for work. Nonetheless, we encouraged study participants to position the sidebar close to their video-conferencing platform window where their self-view would be located, like in Figure 1. A status indicator displays if the non-verbal cue detection is currently active, paused, or does not detect a user in frame.

The feedback widgets provide glanceable, real-time feedback to users on each of the four non-verbal cues via changes in colour and text. We outline how the widgets display feedback for each non-verbal cue below:

- **Posture:** Colour changes from grey (target posture) to blue and text changes to indicate the direction in which *Novecs* detected out-of-bound posture (backwards, forwards, sideways).
- **Gaze:** Colour changes from grey (target gaze) to blue and text changes to signify that the user has been looking away from the target gaze area for a significant amount of time.
- **Smiling/Nodding:** Colour switches from grey to a darker grey for the duration of the sensed smile/nod. The text is updated to display the number of minutes since the last sensed smile/nod.

Core to the design of the feedback widgets was our design goal of encouraging nuanced self-reflection by avoiding an overly normative framing. To ameliorate concerns about feeling judged from piloting early iterations of *Novecs*, we decided to use more neutrally-associated colours, blue/grey instead of red/green, and less normative language, using a self-defined, per-user "target" instead of language such as "(not) ideal", which may imply one ideal standard of non-verbal cue expression. We additionally ruled out red/green to be more inclusive towards colourblind individuals. Displaying the number of minutes since last smile/nod (rather than directly prompting users to smile/nod) was another deliberate design decision to avoid artificially pushing users to smile/nod, while still facilitating self-awareness of existing smiling/nodding behaviour. We chose to use a less salient colour change for smiling and nodding (dark grey) compared to posture and gaze (blue) as users may benefit from being alerted of non-target posture/gaze in real-time during their meeting, whereas they likely do not need to be alerted of when they smiled or nodded.

### 3.5 Post-Session Summary

When a *Novecs* session ends, the Post-Session Summary window (see Figure 2) displays a timeline visualization and summary statistics (proportion of non-target gaze and posture & total smile and nod counts) for all four non-verbal cues. While the real-time feedback of non-verbal cues through the In-Session Sidebar is the primary focus of *Novecs* as a technology probe, *Novecs* also includes intentionally minimal summative feedback through the Post-Session Summary to explore how both types of feedback might be experienced by users in tangent.

The visualization contains four horizontal timelines, one for each of the non-verbal cues, with the x-axis representing the spent meeting time (in minutes). Posture and gaze timelines are accompanied by a legend that defines the colour scheme; grey shaded parts of the bar represent target posture/gaze, while non-grey shading represents non-target posture/gaze. The length of the shaded area corresponds to the duration of the detected non-verbal cue. Each detected smile and nod event is visualized as a black dot in their respective timelines. Shaded sections of the visualization that cut across all timelines represent time periods when *Novecs* has been paused (light blue) or when the user is absent (light orange).

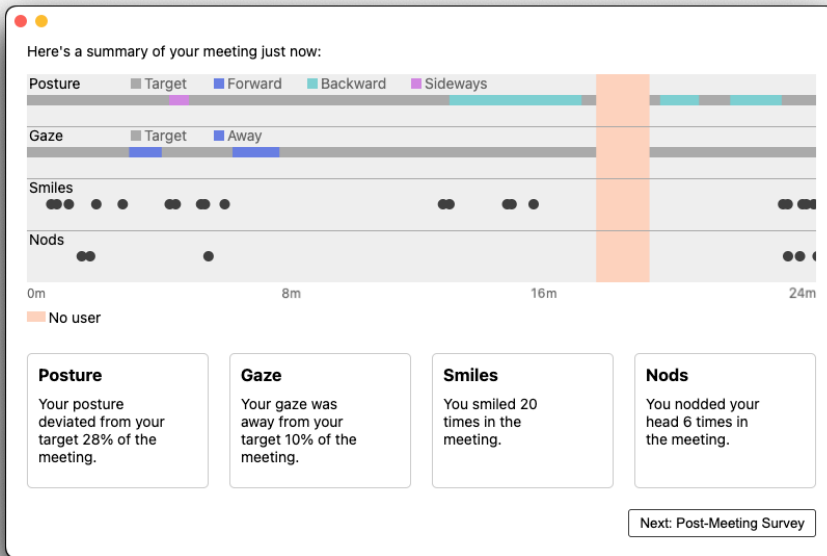


Fig. 2. The Post-Session Summary window, which contains a timeline visualization and summary statistics of the four non-verbal cues. This example indicates that the user had non-target gaze near the beginning, leaned backwards near the end, was out of frame for a short duration (shaded orange area), and smiled and nodded mostly at the beginning and end of the session.

## 4 Method

We conducted an exploratory field study with the *Novecs* technology probe to elicit knowledge workers' reflections on the system in-situ, and to investigate the potential opportunities and challenges of displaying real-time non-verbal cue feedback for supporting self-presentation during online work meetings. The study was conducted entirely remotely, but involved researchers and participants based in two countries on separate continents [countries omitted for anonymity].

### 4.1 Participants

Participants were recruited through public study recruitment platforms, social media posts, online study mailing lists, and convenience sampling through the research team's personal networks. To be eligible for the study, participants needed to: (a) be knowledge workers that attend at least 2-4 online meetings for work weekly with video on; (b) be using a computer running macOS 12.0 or higher for work, with at least 2 GB of free/unused RAM on average; and (c) use a webcam for video calls (including both built-in webcams, such as those on MacBooks, and external webcams).

We recruited 30 participants who met our eligibility requirements through a screener survey. 12 participants dropped out, citing busyness, lack of qualifying meetings, or other unresolved technical issues. 18 participants (women = 9, men = 9, non-binary = 0; see supplementary materials for full demographics table) completed the study. They were on average 32 (SD: 8.83, min: 23, max: 51) years old at the time of the study. 9 participants were recruited from Switzerland and 9 from Canada. We used a broad definition of knowledge workers who came from a number of industries and backgrounds, including: Information Technology, Education, Administration, Engineering, and Medical Technology. 2/18 participants were PhD students working on their research (i.e., not taking courses), the remainder were professionals with roles including: Communications Lead,

Account Manager, Marketing Associate, Data Analyst, and Product Designer. Participants were compensated with a 40 CHF (or local currency equivalent) gift card. The study was approved by the research ethics boards at the researchers' academic institutions in both countries.

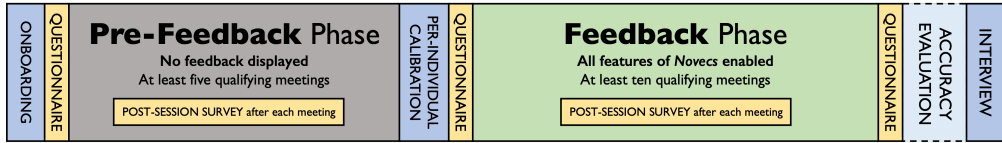


Fig. 3. Diagram visualizing the study procedure.

## 4.2 Procedure

In this section, we describe the study procedure for advancing participants through the two phases of our field study: **Pre-Feedback** and **Feedback** (see Figure 3 for an overview). We began by onboarding participants to the study through an initial 30-minute Zoom call, conducted by one of the lead researchers. In the onboarding session, participants were given an overview of the study procedure and goals, installed the Pre-Feedback phase version of *Novecs*, and familiarized themselves with the tool, including how to start, pause, stop a meeting session, and guidelines for positioning the In-Session Sidebar appropriately, relative to video-conferencing window and their self-view. Participants completed the pre-study questionnaire as soon as possible after the onboarding session and before starting the Pre-Feedback phase of the study.

During the **Pre-Feedback phase** of the study, participants continued attending video meetings as usual with *Novecs* until they completed five or more qualifying sessions. Qualifying sessions needed to meet the following criteria: (a) the participant's video was enabled for more than 50% of the meeting; (b) most meeting attendees were remote; (c) the participant self-reported that their overall level of interaction in the meeting was more than just passively listening; and (d) the meeting was at least 15 minutes long. These criteria were selected so that participants experienced *Novecs* in meetings where it could potentially be relevant.

Whenever participants started a meeting session in *Novecs*, a *per-session* calibration step would be initiated (to determine "target gaze and posture" for that meeting, see Section 3.4 for details). For the Pre-Feedback phase, *Novecs* did not display any feedback to the participant, aside from detected user presence or absence. After finishing a meeting and ending the session, participants answered the post-session survey to report meeting context information, which was then used by *Novecs* to determine if the past meeting was qualifying or not. The primary intention behind the Pre-Feedback phase was to give participants time to acclimate to manually starting *Novecs* with their regular work meetings.

After the Pre-Feedback phase, participants completed the mid-study questionnaire and participated in another 30-minute Zoom call with a lead researcher to be onboarded to the **Feedback phase**, where both types of feedback, the In-Session Sidebar and Post-Session Summary, were enabled. In addition, participants were guided to complete the *per-individual* calibration step, to tailor thresholds for each non-verbal cue specifically to the participant. The per-individual calibration only occurred at this point in the study to avoid influencing participants from focusing on the study's four non-verbal cues at the study outset; participants could not adjust these thresholds without researcher intervention; threshold changes were made only if significant issues arose with the non-verbal cue detection. New in the Feedback phase was that after completing the per-session calibration, participants would see the In-Session Sidebar. After finishing a meeting and ending the

session, participants were first shown the Post-Session Summary before they were prompted to complete the post-session survey. Participants were part of the Feedback phase until completing ten or more qualifying (same criteria as in the Pre-Feedback phase) sessions, upon which they completed the post-study questionnaire and were invited to a final 60-minute Zoom call, consisting of a semi-structured interview, the accuracy evaluation step (for participants who opted-in, see Section 4.3.5), data transfer, debriefing and uninstallation of *Novecs*.

### 4.3 Data Collection

Throughout the study, we collected both qualitative and quantitative data via (a) the *pre-*, *mid-*, and *post-*study questionnaires, (b) the post-session surveys, and (c) the semi-structured exit interview. Our primary focus was on the *qualitative* data, due to the exploratory nature of a technology probe [57] study; we include quantitative data analysis where appropriate to help contextualize our qualitative findings. We conclude this section by describing the procedure for quantitatively assessing the sensing accuracy of *Novecs*.

**4.3.1 Pre-, Mid-, and Post-Questionnaires.** The three questionnaires each contain the same set of questions (except demographic questions only asked in pre-study). In the questionnaires, participants rated their awareness of their own non-verbal cues during video meetings, the importance of non-verbal cues during video meetings for meeting effectiveness, and whether their own non-verbal cues matched those of an effective meeting participant. They also included Likert-scale questions to better understand participants' habits during video calls as well as select questions of the Zoom Exhaustion & Fatigue scale [41] and Emotional Intelligence Scale (WLEIS) [104]; specific wording and analysis for this data can be found in the supplementary materials.

**4.3.2 Post-session Surveys.** After each session, the probe asked participants a few questions about the past meeting, such as: the number of meeting attendees, modality (hybrid or fully remote), video conferencing software used, whether screensharing was used, whether participants' shared their video feed, and their level of interaction (ranging from passive listening to active participation). Qualifying sessions included two additional 5-point Likert-scale questions on awareness of and satisfaction with their non-verbal cues during the past meeting.

**4.3.3 Semi-structured Exit Interview.** We conducted 18 semi-structured interviews at the end of the study, starting by asking about their overall reactions to and general usage patterns with *Novecs*, such as when and how often they noticed or glanced at the In-Session Sidebar. We also asked about potential learnings and impact from using *Novecs*, to better understand the usefulness and usability of the In-Session Sidebar and Post-Session Summary. Participants were further asked to share their perspectives on each of the four non-verbal cues and use cases where *Novecs* might be more or less valuable. We concluded by discussing *Novecs*' design, including any suggested improvements or missing functionalities. Interviews were audio-recorded and transcribed for qualitative analysis.

**4.3.4 Novecs Usage and Meeting Session Data.** Whenever participants manually started a meeting session, *Novecs* collected usage and meeting session data in the background. The collected data includes information on session length, raw data on when each of the four non-verbal cues was detected during a session, and usage data, such as when the participant paused/resumed the session. No data specific to the meetings, such as meeting content, was collected. Note that the quantitative data analysis is from 17 participants (P8's data aside from their interview was corrupted).

Overall, the quantitative usage data showed that participants varied in the duration of the time they participated in the study. Participants took 1.43 weeks on average (SD: 1.41, max: 6.03, min: 0.29) in the Pre-Feedback phase (five qualifying meetings) and 3.71 weeks (SD: 2.93, max: 11.86, min: 0.86) in the Feedback phase (ten qualifying meetings). Participants used *Novecs* for a total of

279 qualifying meetings, 106 in the Pre-Feedback phase (6.2 meetings on average per participant) and 173 in the Feedback phase (10.2 meetings on average per participant). Meeting session lengths were similar in both phases, with averages of 41.49 minutes (SD: 21.02, min: 15.98, max: 114.44) in the Pre-Feedback and 42.11 minutes (SD: 23.07, min: 15.15, max: 142.71) in the Feedback phase. 176 meetings were held in Zoom, 51 in Microsoft Teams, 17 in Google Meet, 17 in DingTalk, 14 in Slack Huddles, and 4 in other video conferencing platforms. There were 130 meetings with 2 participants, 36 with 3, 50 with 4-5, 42 with 6-10, and 21 with 10+. Additional usage statistics can be found in the supplementary materials.

*4.3.5 Accuracy Evaluation.* We quantitatively evaluated *Novecs*' accuracy by asking participants to self-assess 64 images, sampled from the set of webcam images stored during meeting sessions.

*Image Data Collection:* To collect image data, *Novecs* persisted images taken from participants' webcams during sessions only if they opted-in to the accuracy evaluation step, which all participants did (except one participant with a corrupted dataset). Normally, for privacy reasons, all images are immediately discarded after real-time processing. As it is too intrusive and resource intensive to store and label all image frames from a session, our approach was to store a specific, pre-defined subset of 60 image frames for each meeting and participant. 24 of these stored frames were based on random moments during the meeting. The remaining 36 frames (9 per feature) were selected based on 3 different "moments of interest", as calculated by *Novecs*. Since the goal was to assess accuracy, "moments of interest" were based on when maxima, minima, or challenging frames were detected, according to a strength score that was calculated by our detection approach and was specific to each of the four non-verbal cues. Using "smiling" as an example, 3 image frames were each selected from the session's pool of image frames, where the calculated "smiling" score was highest (maxima), lowest (minima), and closest to the threshold for differentiating between smile/non-smile (challenging).

*Participant Assessment:* As part of the exit interview, participants were instructed to rate a subset of their own collected image frames (of themselves; 64 in total) to produce labelled data that could be used as the quantitative accuracy evaluation. We used a stratified random sampling method where 16 frames (4 random, 4 challenging, 4 min, 4 max) were randomly sampled each for all four non-verbal cues, across all meetings. For each of the sampled image frames, we presented the frame (either as a still image or GIF for depicting the nodding motion), together with a positive or negative statement describing the state of the non-verbal cue. For example, the statement would say "You are smiling." for a maxima, "You are not smiling." for a minima, and either statement randomly picked, for random and challenging frames. When reviewing these 64 image frames, participants could exclude images they were not comfortable sharing with the researchers. In total, we collected 1087 accuracy ratings (one participant's data only had 63 ratings).

#### 4.4 Data Analysis

We analyzed interview transcripts using the Braun and Clarke approach to reflexive thematic analysis [19], identifying themes with a primarily inductive approach, while being theoretically informed by existing literature in self-presentation and glanceable display design. Our codes and subsequent themes were constructed from the data and were not predetermined by an existing theoretical framework. Three members of the research team first independently coded a subset of the interviews (3/18) to generate preliminary codes. Then, a lead researcher, paired with one of the two other team members, each coded separate subsets (3/18 each), meaning that half of our dataset (9/18) was coded by at least two members of the research team. One of the lead researchers then coded the remaining interviews (9/18). Our data analysis process was marked by frequent meetings to review our codes and discuss a range of interpretations and perspectives in analyzing the data.

In line with best practice in qualitative analysis, we did not calculate inter-rater reliability, deciding instead to use multiple coders to achieve ‘crystallization’ [37, 99]. We took a recursive approach to our data collection and analysis process, moving back and forth between analyzing transcripts and iterating on the semi-structured exit interview questions. Although we reported participant counts per theme, we acknowledge (in line with Braun and Clarke [19]) the key limitations to this convention for representing prevalence. Counts mask the depth to which a participant might be articulating a theme and is often rendered problematic due to the nature of qualitative data collection [20]. For quantitative analysis, we calculated and reported descriptive statistics from the Likert-scale questions in the pre-, mid-, and post-questionnaires and the post-session surveys to provide context to the qualitative findings, where appropriate.

**Researcher positionality and reflexivity.** We describe our positional stance as researchers to help contextualize the lenses by which we viewed the data and subsequent analysis. We reflected, discussed, and shared our own experiences with online meetings using various video conferencing platforms and our own challenges with expressing and interpreting non-verbal cues. In particular, three members of the research team work at an academic institution in Country1, and two in Country2; members from both countries were involved in coding, enabling us to share and contemplate on our own cultural perspectives. Both of the lead researchers of the study also experienced using *Novecs* regularly in their own online work meetings. The research team consists of two graduate students, one senior researcher, and two professors. All members of the team are knowledge workers who regularly participate in online work meetings.

## 5 Findings

To contextualize our findings, we first summarize participant perceptions of *Novecs*' accuracy and calculated accuracy scores, suggesting that the technology probe was functional with participants' real meetings. We then delve into each of our four identified themes, beginning by highlighting participants' increased awareness of their non-verbal cues (**Theme 1, RQ3**). Next, we describe the specific meeting types where non-verbal cue feedback was perceived to be most useful (**Theme 2, RQ1**). We also identified tensions between authentic and in-authentic self-presentation, and how critical natural timing is when changing non-verbal cues (**Theme 3, RQ2 & RQ3**). While *Novecs*' neutral feedback design helped with authenticity and natural timing, it also surfaced additional tensions in noticeability, distractibility, and actionability (**Theme 4, RQ3**). We conclude with participants' overall willingness to continue using systems like *Novecs* in the future.

### 5.1 Perceived Accuracy and Accuracy Scores

When asked about accuracy, 7/15 (1 was not asked, and 2 were not confident in their awareness to provide a conclusive response) participants shared that they perceived *Novecs* to be fairly accurate for all four non-verbal cues and 7/15 shared that it was accurate for some of the cues but not all.

1/15 (P3) felt like it was “not quite accurate”, but also said that “even though they have some false positives, I still think it is useful” (P3), suggesting that they may have still been able to effectively experience the probe during the study. Inaccuracies around smiling were mentioned by 5/15 participants, nodding 4/15, gaze 3/15, and posture 2/15. Participants experienced occasional false positives with smiling and nodding, and gaze had issues when participants had wide screen monitor setups. Unsurprisingly, when there were too many false positives, a few participants shared that they began to doubt or ignore some of the feedback for problematic cues, and selectively focused on cues that they believed to be more accurate, which would have shaped their experience of *Novecs*. We took note of these cases, checked with participants during the interview to ensure that their reactions were not due to excessive inaccuracies, and also reviewed our selected quotes to provide additional context around accuracy if-needed. The quantitative accuracy evaluation of

*Novecs'* sensing accuracy at the end of the study (see section 4.3.5) revealed an average overall accuracy of 74% across all four types of cues. The individual average accuracy scores for each non-verbal cue were 67% for gaze, 72% for nodding, 71% for posture, and 83% for smiling. While there are differences in detection algorithms along with how and where (i.e., in-situ versus lab) accuracy is evaluated, these scores generally fall in line with existing literature on automatic non-verbal cue detection, which mostly hover between 65-85% [2, 11, 46].

## 5.2 Theme 1: Increased Awareness of Non-Verbal Cues during Video Meetings

Overall, 13/18 participants shared that *Novecs* helped increase their awareness of their non-verbal cues. Participants attributed these awareness increases to both the real-time feedback, which helped with *in-the-moment* awareness, and also the post-session feedback, which helped with *recapping* the meeting. While the impact was mainly on awareness, a few participants also reported short-term behaviour changes in their non-verbal cues during meetings.

**5.2.1 Real-time Feedback Facilitated In-the-Moment Awareness.** We begin by summarizing participants' different approaches to becoming aware of the real-time feedback in meetings. 10/18 participants self-reported a *proactive* approach of glancing at the sidebar regularly to check on the feedback: "Maybe like every five to ten minutes or so, I would probably just take a look or maybe if I wasn't actively talking for a while, I would probably take a look at it" (P13). Others (11/18) *reactively* noticed in their peripheral vision when the sidebar's colours changed as a signal, without actively glancing at it: "[The sidebar] would just kind of catch my eye every time one of those widgets (...) turned grey, but otherwise, yeah I was kind of looking at the counter for like the last time I smiled or nodded. Basically, every time the widget changed, I'd kind of glance over at it" (P2).

9/18 participants in the study reported how the real-time feedback helped them to think about their own non-verbal cues "in-the-moment" (P17). Specifically, the *Novecs* sidebar allowed them to become aware of, and reflect on, the **moment-specific appropriateness of their non-verbal cues during the meeting**. Rather than simply being aware of their non-verbal cues in general, participants were able to connect their behaviours to that moment in the meeting, and reflect on if it was appropriate or not. P10 used the real-time feedback to make sure that they were "showing the right reaction to something" and if their expressions were "matching what's going on". P10 added, describing a specific example where they changed the amount of times they smiled to be less to match that moment in the meeting: "I felt like I reduced that a bit, and like where it's more appropriate to do that, just kind of thinking about matching more towards the kind of overall atmosphere of the meeting." Similarly, when P9 was asked about how they reacted when they noticed the real-time feedback about smiling, they said: "Like my [role] was sort of a shadow figure in the meeting, it was actively listening, but not actively participating, so probably it makes a bit of sense [to not smile]. The tool helped me realize that, I think I would not have noticed that by myself probably."

**5.2.2 Post-Session Feedback Helped Participants Recap and Reflect.** 10/18 participants liked the post-session feedback as a way to recap the meeting that had just occurred. Specifically, the post-session feedback helped participants become **more aware of and reflect on their participation and engagement** in the meeting by looking at the non-verbal cues that were visualized on the timeline. P2, for example, used the empty spaces in the timeline to become aware of when they might have been less engaged: "It showed me when I was kind of zoning out in a meeting or something when there were some sections that were like, completely bare." Similarly, P10 reported the summary helping to reflect on the meeting outcome: "I thought it was pretty cool to like see over time (...) and also think of how the discussion had unfolded and how that matched up with the different kinds of movements like nodding and stuff like that."

Participants further stated that the post-session feedback allowed them to **better understand their behaviour depending on the meeting and sort-out misconceptions** they had about their own expressions. For example, P9 became more aware of how their nodding and smiling changed in a programming session compared to other meetings: “I was nodding way more often, I was nodding a lot because he was explaining to me a lot of things... And it was like a bit less smiling, it was a technical session so it was less smiling. While at other meetings that are a bit more relaxed, I was smiling more and nodding less. So it was mostly for me a confirmation of the type of meeting I'm in and what's my reactions to it.” Similarly, P15 realized that they were more expressive during meetings than they had previously considered: “When I have to rate myself, I would always be more on the negative side. If you asked me how my facial expressions are during meetings before this study, I would have said it's probably really bad because communication is not my strength. But then seeing that I actually nodded and smiled quite a lot made me feel that maybe it's not as bad as I thought.”

*5.2.3 Impact on Short-term Behaviour Change.* 8/18 participants reported short-term behaviour change during meetings, such as how the real-time feedback helped to “adjust and correct behaviours” (P18) and “course correct in the middle of [meetings]” (P17). P6 also hinted at longer-term behaviour change, saying: “But like for sure the leaning back thing, I keep coming back to that, it's like ingrained in me now. Like honestly I do it myself or if I go too far back, I see it now on the screen and it comes forward.” However, 6/18 participants (P1, P2, P7, P9, P12, P17) also shared that it didn't “change much [of] my behaviour during the meetings” (P9). This was consistent with the quantitative data from the post-session surveys and study questionnaires that measured whether or not participants' assessment of their non-verbal self-presentation changed after the study. Our t-test ( $\alpha = .05$ ) showed no significant differences in participants' self-reported satisfaction (5-point Likert scale, 5 denotes extremely satisfied) with their non-verbal cues when comparing the Pre-Feedback (mean: 3.46, SD: 0.74) and Feedback phase (mean: 3.43, SD: 0.75) from their post-session surveys. Similarly, no differences were found from participants' ratings in the study questionnaires (5-point Likert scale, 5 denotes increased agreement) on whether their non-verbal cues matched that of an effective meeting participant (means for pre: 3.72, mid: 3.94, post: 3.94). See supplementary materials for further details of this analysis. While we did not expect to see measurable, long-term behaviour change in this study, as it is ultimately a long-term process [66], that almost half of the participants self-reported some behaviour change hints at the potential of a real-time, non-verbal cue feedback system.

### 5.3 Theme 2: Different Meeting Types Impact the Utility of Non-verbal Cue Feedback

We identified how different types of meetings can influence participants' perceived utility of real-time non-verbal cue feedback. Meeting types differed in factors such as meeting size, the importance of the meeting, and familiarity with the other meeting attendees.

*5.3.1 Number of Meeting Attendees.* The number of meeting attendees influences how active and engaged participants are or feel they need to be, and thus, **impacts their cognitive capacity to pay attention to, and act on, the feedback** provided through the sidebar. 12/18 participants reported that the real-time feedback was generally most useful in small-to-medium sized group meetings when they were not in a constant state of active participation, but still need to be engaged: “In meetings, where there's more people and my role is not that active, I think it's useful. Especially there are some meetings when I'm really, really trying to take as many notes as possible and I don't say much. And then I concentrate a lot on writing. And then from time to time seeing ‘ohh, I should actually show that I'm still here.’” (P15).



To the contrary, feedback in one-on-one meetings was generally considered as less appropriate, since participants were talking for most of the time and had little-to-no cognitive capacity to pay attention to the sidebar: “Generally for like one-on-one meetings where like I’m talking 50% of the time, then I’m not really paying attention to the widget... maybe not so much when I’m talking, but when other people are talking. Obviously I have more capacity to take a look at the tool then.” (P2). Some participants (5/18: P2, P5, P10, P13, P16) also shared that the rapid, back-and-forth conversational pace in one-on-one meetings sufficiently signaled engagement.

Unsurprisingly, for meetings with a large number of attendees, the feedback from *Novecs* was less useful as their non-verbal self-presentation didn’t matter: “If it’s a session with 70 participants and I’m one of many small pictures on the screen... then I don’t care how I come across” (P7).

**5.3.2 Importance and Formality of the Meeting.** 7/18 participants described that the non-verbal cue feedback was particularly useful in **meetings that had more at stake in terms of importance or were more formal**. Participants described how examples of such meetings were ones with external clients or supervisors, where there were greater potential job repercussions, compared to, say, a meeting with their regular team. Participants felt like such meetings were ones where they had to be especially aware of and better manage their non-verbal cues. For example, P6 said: “It’s like what is the level of professionalism needed? That’s kind of what I used it [*Novecs*] for – to help me.” P3 discussed how their perception of how important the other attendees in the meeting affected how conscious they were about their non-verbal cues: “When I have the meeting with important roles then I have to adjust my visual language and body language based on the tool [*Novecs*].” In addition, when asking participants about how they would see themselves using *Novecs* in the future, a few participants (3/18: P7, P8, P17) identified that they would find the tool particularly useful when training or onboarding people as well as when interviewing for jobs, which emphasizes the utility of *Novecs* in more consequential settings, where impressions might matter more.

**5.3.3 Familiarity with Other Meeting Attendees.** Related to formality is participants’ familiarity with the other meeting attendees. 8/18 participants felt that it was **less important to manage their non-verbal cues with meeting attendees they were already familiar with**. For example, P4 said: “The people I’m working with and my team, I already know them really, really well, so I don’t feel the need to be super polite and nice and whatever. So I think that I usually nod more or smile more if it’s a new person.” Although they were aware of “the whole idea of being forward to show interest”, P18 was comfortable with leaning back only because they were with “people that I really felt super comfortable with, and [were] less focused on (...) the presentation of myself.” In these meetings, familiarity allowed participants to feel like they could be themselves and not have to manage their self-presentation, which is discussed further in the next section.

## 5.4 Theme 3: Tensions between Authenticity, In-Authenticity, and Natural Timing in Self-Presentation

Participants emphasized the importance of being authentic in how they expressed themselves in work meetings, but recognized that there are some situations where they had to externally present themselves in a manner different from how they were feeling internally (i.e., in-authentic self-presentation [18, 83]). They further wanted their self-presentation efforts to appear natural and with appropriate timing. *Novecs*’ neutrally-framed feedback was generally viewed as compatible with navigating this tension and the need for natural timing.

**5.4.1 Importance of Authentic Self-Presentation.** When it comes to self-presentation in work meetings, participants **stressed the importance of being “authentic”** (P8), “genuine” (P2), “natural” (P2, P3, P11, P17), and “organic” (P6). A few participants (4/18: P2, P3, P4, P17) even alluded to

the importance of embracing “humanness” (P2) and imperfections. P17, for example, emphasized how being authentic sometimes means displaying imperfect non-verbal cues: “When you have a lot of meetings like me in a day, I mean, I’m not like a robot. I am going to move. I am going to look around. I am going to get bored. I am gonna (...) sometimes not be in the mood (...). And so I don’t believe there’s a perfect meeting behaviour. Everybody has their own personalities, their own characteristics, and I think my philosophy is like you can make faces, you don’t have to have the perfect [expression] like you could have a resting whatever face, but as long as you’re answering the questions and being respectful and you’re actively listening, engaging, et cetera [sic], then (...) you can overlook some of the human things like that.”

P8 explained how the In-Session Sidebar of *Novecs* was compatible with their emphasis on authentic non-verbal behaviour, both of themselves and of others: “There’s still a way to balance being your authentic self while also using the sidebar for cues... I just want them [others] to be how they are. They [others] don’t have to go out of their way to do all these extra things very frequently. Especially, if, that’s just not who they are.” They continued, saying that the sidebar was a reminder of their current behaviour, but never forced them to change from their authentic self: “I wouldn’t correct it unless it felt right to do it. And that’s why I just kept my body language and my gaze the same because that was authentic to me (...) It didn’t feel like I had to go out of my way to smile extra or not [smile] extra. It was just a little reminder of ‘ohh, you haven’t [smiled] in a while’” (P8).

**5.4.2 In-Authentic Self-Presentation is Occasionally Necessary.** Despite the importance of authenticity, a few participants (5/18: P3, P8, P12, P13, P14) also explicitly mentioned how they needed to engage in in-authentic self-presentation, where they had to **outwardly express something different from what they were feeling internally**. Different meeting types may help explain situations where authentic or in-authentic self-presentation is needed (e.g., important or formal meetings may require more in-authentic self-presentation; see section 5.3).

In situations that call for in-authentic self-presentation, participants like P3 and P8 explicitly describe how *Novecs* was a useful reminder. P3 mentioned a few meetings where they needed to smile more to express positivity and agreement more than they felt internally, finding *Novecs* to be “very useful”, saying: “[When] my supervisor asks me to do something even though I’m not happy, but [sic] I pretend to be happy. I pretend, to hide my actual emotion (...) when I have a meeting with the master’s student who helped me perform experiments, even though I’m not so satisfied, I should still smile a lot.”

P8 also shared a story of when they were having a bad day but still had to facilitate a session at work, another example of an important meeting where in-authenticity might be necessary: “When you are facilitating you can’t really be your authentic self, you have to put on that little customer service vibe... I wasn’t feeling it, but I also couldn’t completely be my authentic self... It was like a good reminder — I haven’t done that in a while, and I’m talking to 30 people, so I should smile a little, or look a little more approachable, which is something I found it [*Novecs*] really helpful for actually, because I struggle with looking approachable, and looking friendly, and showing some of those indicators in larger groups.” While not explicitly discussed by participants in this manner, we note that in-authentic self-presentation may also occur due to a lack of self-awareness of their non-verbal cues (i.e., unconsciously [18]). For example, if one feels agreeable internally, but forgets to express it externally through nodding or smiling; our study design could not capture this.

**5.4.3 Natural and Appropriate Timing is Critical.** A few participants (4/18: P6, P8, P11, P13) stressed the importance of **natural and appropriate timing for changing their non-verbal cues**, as opposed to simply doing so whenever they noticed feedback signals. For example, P6 mentioned: “Nodding is like, if someone makes a point, you want to give them an affirmation of that point. But

I wanted to make sure that it was organically done and not me just being like I haven't nodded in two minutes... I do try and do that and this study did bring that to light." P13 described their thought process when seeing the feedback as waiting for the next opportune moment, saying: "if it says I didn't smile for [some time], I would try to smile at the next opportunity."

6/18 participants (P2, P5, P8, P11, P13, P16) also called out the intentionally neutral, non-normative feedback that *Novecs* provided as compatible with supporting natural timing for changing their non-verbal cues in their work meetings. For example, P11 observed how the "prompt doesn't ask you to smile, it just says that you haven't smiled for these many minutes", saying that they should smile according to the "natural flow [of conversation]". Similarly, P6 shared how the In-Session Sidebar's last smiled/nodded minute counter did not pressure them to change their non-verbal behaviour "without it feeling like the appropriate time" despite the "timer getting high." However, P2 and P17 felt like the reminder could sometimes influence them to change in unnatural ways. Although P2 said that the tool was "not so much pressure" and was "not like I was being forced to do that", they did describe how it felt "a little ingenuine at times". When referring to the last smiled minute counter, P2 described: "I was trying to keep that number down, so I would kind of smile or nod more frequently, even if it didn't feel natural to the conversation that was being had."

#### 5.5 Theme 4: Tensions between Noticeability, Distractibility and Actionability

While non-normative and neutral feedback was one of the key design goals of *Novecs*, and could be important for supporting natural timing of changing non-verbal cues (as suggested by Theme 3), it appears to have also surfaced tensions around balancing noticeability, distractibility and actionability for real-time displays of non-verbal cues.

6/18 participants (P4, P8, P10, P12, P15, P16) explicitly noted that the real-time feedback was "subtle" (P8, P16) and "minimal" (P12), and thus did **not distract them from focusing on their meetings**. Noticeability and distractibility are a known tradeoff, as increasing the saliency of the feedback would also draw users' attention away from their meetings. The neutral colour scheme of visualizing data in a light blue or grey shade, as opposed to e.g., a bright red shade, likely added to the non-distracting experience. For example, P4 shared: "I never felt interrupted from the tool during a meeting... I was not forced to look at the tool and I really like that because I feel if I had like a red light popping up or something, I would always be distracted because I'm easily distracted during meetings sometimes. So I really liked that it's subtle." Similarly, P16 said: "It helped that the colour is not (...) that interruptive. So it's not like a red signal blinking when I'm leaning a little bit behind. (...) I've never felt judged by the tool." Beyond not being distracting, we also note that P16 brings up not "feeling judged" by the colour scheme, which is consistent with our goal of avoiding overly normative feedback design.

When viewing the In-Session Sidebar, 6/18 participants (P2, P6, P9, P14, P15, P18) also suggested that the real-time feedback by *Novecs* could be **more actionable, such as by being more explicit about what actions to take**. Interestingly, some participants suggested adding "high-level" constructs like engagement, attention, or emotions, despite known challenges with actionability for such constructs [48]. P9 wanted a "facial emotion detector (...) and to see if (...) half of the time you have a pissed off expression, or a happy expression." It may be that "higher-level constructs" help provide a rationale for changing one's non-verbal behaviour, whereas simply seeing, e.g., "non-target posture", may not spur users to action. In addition, P14 suggested wanting the tool tell them if their current posture was "good or bad" and "how it might be interpreted by people seeing me", suggesting some individual desire for more normative feedback.

5/18 participants (P4, P5, P14, P15, P17) also expressed challenges with actionability when viewing the Post-Session Summary. While it was valuable to increase their awareness and reflect on their participation and engagement (see section 5.2.2), many did not know how to interpret and act

on the presented feedback. More specifically, P4 shared that they struggled with **not having a point of reference** of how often one should be expressing non-verbal cues in a meeting: “I don’t really know what the bar should be like. How often should a person be smiling, or how often should a person be nodding in order to look like an engaged participant in a meeting?” (P4). Other participants (P2, P8, P9, P18) also suggested having a way to compare their facial expressions and body language across different types of meetings (Theme 2, see section 5.3), which may help alleviate the lack of a reference point. For example, P2 stated: “It would be helpful to tie the data into what kind of meeting I was in because my one-on-ones are a lot more engaged compared to a meeting that has seven people. So that data is kind of hard to compare. So if I see like side-by-side like ‘[at] my all-hands meeting (...) you were leaning back most of the time’. Whereas like in this [other] meeting you’re engaged the whole time.”

## 5.6 Willingness for Future Use

Finally, the majority of participants (15/17, one did not provide a conclusive response) answered that they would consider using a system akin to *Novecs* in the future. Of these, 10/17 were interested in continuing using *Novecs* regularly as-is, 3/17 occasionally, 2/17 for a short period of time (e.g., for awareness training), and only 2/17 (P2, P7) did not want to use it further. P6 was particularly enthusiastic: “I did really enjoy using it. I think it’s very helpful for being engaged. It would be really good if it was built into some of the video conferencing software like that would be the ideal.” Reasons for not wanting to continue using *Novecs* were that they were already satisfied with their existing awareness (P7) and that being consciously reminded of their own non-verbal cues might lead to “more burnout” (P2).

## 6 Discussion

We begin by summarizing the key findings from the *Novecs* technology probe. Next, we reflect on the importance of authenticity and the imperfect “humanness” in non-verbal self-presentation as well as the tradeoff between noticeability and distractibility, before concluding with design opportunities for future real-time, non-verbal cue feedback systems and study limitations.

### 6.1 Summary of Key Findings

We learned that real-time feedback systems, like *Novecs*, have the potential to foster self-awareness of non-verbal cues. Specifically, it enabled users to connect their behaviour with that current moment in the meeting, encouraging reflection on the in-the-moment appropriateness. When changing their non-verbal cues, participants stressed that this needed to be done with natural timing. The neutrally-framed, real-time feedback encouraged reflection rather than immediate action in-the-moment, surfacing both tensions of authentic and in-authentic self-presentation as well as distractibility, noticeability, and actionability. Notably, we identified meeting types where real-time feedback on non-verbal cues was most useful, such as small-to-medium sized meetings that involve unfamiliar meeting attendees and meetings that are deemed more important.

### 6.2 Importance of Authenticity and Imperfect “Humanness”

For online work meetings, participants shared that in general, they valued authentic non-verbal self-presentation, but acknowledged that there were also times where in-authentic self-presentation [18, 83] was necessary, such as when there was a need to maintain a level of professionalism or when taking on a particular role in a meeting. Our findings extend the known tensions around authenticity and in-authenticity as identified by Goffman [47] through characterizing this tension specifically for knowledge workers in online work meetings. In addition, our findings on meeting types can potentially explain the specific scenarios in which authenticity is valued (e.g., when you

are familiar with the other meeting attendees) or when in-authenticity is needed (e.g., in important or more formal meeting contexts). *Novecs* was compatible with navigating this tension, as well as supporting participants in changing their non-verbal cues with natural timing. While reflection on in-the-moment appropriateness still requires intentional effort even with *Novecs* support, long-term training could shift self-presentation from conscious to automatic [83], eventually opening up the possibility of no longer needing to rely on non-verbal cue feedback.

Notably, there may also be value in imperfect non-verbal cues in online meetings [21, 58, 90], especially as some participants stated that they were comfortable with expressing their natural, imperfect self with familiar coworkers. It may also be reasonable to assume that imperfect non-verbal cues could build connection through expressing vulnerability or signaling one's well-being. For example, if one is feeling unwell or unhappy, authentically reflecting it in one's non-verbal cues through a lack of smiling or a slouched posture could be a signal to other meeting attendees to ask how they are doing, given a culture of trust and psychological safety [34].

### 6.3 Balancing Noticeability and Distractibility

Our findings also replicate the known tensions around distractibility and noticeability from the glanceable display literature [10, 80, 96, 98] but in the context of online work meetings. In online work meetings, non-verbal cues often only take a secondary role to the actual work being conducted or discussed, unlike contexts such as public speaking [96] or social conversations [3, 4], where appropriate non-verbal cue behaviour may be more central. Another challenge in work meetings is managing the dynamics within the meeting. For example, workers may place greater importance on non-verbal cues more at the beginning or end of a meeting when exchanging pleasantries, whereas these cues may be less valued when meeting attendees are all focused on reading a screen-shared document. Distractibility greatly fluctuates even within a particular work meeting. In our findings, while *Novecs* positively impacted participants' awareness on non-verbal cues and, in a few cases, even impacted short-term behaviour, several participants stated that they did not notice the glanceable display during some types of meetings. While it was an explicit, and appreciated, design decision to avoid making the display too intrusive to distract from the actual meeting content, the impact that such a real-time display can have on awareness and behaviour is limited by it being noticed at the specific time of the feedback.

The type of meeting is one such factor to consider in balancing this tension. For example, in one-on-one meetings where participants often failed to notice the sidebar due to rapid back-and-forth discussions, one might consider only disrupting the user to notify them if there is a major issue with their non-verbal self-presentation (what constitutes a major issue may be defined by the user themselves), such as if they were leaning to the side for a significant amount of time in the meeting. Future work could also further consider the role of the user in the meeting (e.g., presenter, speaker, listener) in this specific tension and in the broader utility of a real-time non-verbal cue feedback system. Another way to increase noticeability while keeping distractions low could be to adapt the saliency (i.e., noticeability) of the feedback based on cognitive load. For example, feedback could be displayed only when the user is not speaking and can actually notice and act on the feedback. Designers could further customize the modality, such as feedback in the form of just-in-time notifications inside the video conferencing platform, rather than through a persistent, glanceable display (P2, also in [96]), or even using audio (P4, P6) and haptic (P14) channels.

### 6.4 Design Opportunities for Future Non-verbal Cue Feedback Systems

We propose the following design opportunities for future similar real-time, non-verbal cue systems for online work meetings. This list provides a starting point for future designers' consideration, in line with a technology probe's "design goal of inspiring users and designers to think of new

kinds of technology to support their needs and desires” [57]. We note that they need to be further iterated on and validated in more rigorous, larger-scale, and longitudinal studies.

- **Adapt the feedback design of such systems to meeting types:** Our findings suggest that small-to-medium sized meetings of importance with unfamiliar meeting attendees are meeting types where real-time non-verbal cue feedback are most useful. The saliency, noticeability, and distractability of feedback could adapt to the meeting type. For example, feedback on non-verbal cues could be more proactive or aggressive in notifying users in important meetings where maintaining their ideal non-verbal self-presentation is especially relevant. On the other hand, meetings with a large number of attendees or with close team members could be ones where non-verbal cue feedback has reduced saliency or even larger thresholds for non-target behaviour.
- **Encourage reflection on authenticity and in-authenticity:** Meeting types could also help characterize scenarios where authentic or in-authentic self-presentation might be needed. Systems could help users reflect on this dimension of authenticity in non-verbal self-presentation. For example, if a worker is leading a presentation for a meeting, they might realize that they need to take on a (potentially in-authentic) role of a confident and assertive presenter; Persona-like presets in future systems could be one way to encourage reflection and help target specific non-verbal cue behaviour for training.
- **Leave room for imperfection and avoid overly normative design by default:** Interestingly, a few participants emphasized not wanting to have to always manage their non-verbal cues well, and that it's okay to be imperfect, as it feels more “human”, especially among close colleagues. Future designers should, like in *Novecs*, avoid overly normative design; they could consider features that snooze or dismiss non-verbal cue feedback, or even prevent over-reliance by encouraging tapering use of the system once users are satisfied with their non-verbal self-presentation. Conversely, a few participants desired more normative feedback to help with actionability; customization through a “normativity slider” could be a possible design solution.
- **While low-level cues are more actionable than “high-level” constructs, some users want more guidance:** Participants struggled with not having a point of reference for expressing non-verbal cues. As non-verbal self-presentation is personal, systems like *Novecs* can help knowledge workers begin to become more aware of their non-verbal cues and reflect on it, so that they can come up with personal improvement goals (e.g., smiling more often) that can serve as a personal point of reference to facilitate actionability. Curated personas, as mentioned earlier, that can adapt to both the individual and meeting type, could also help with actionability, by serving as a point of reference for users to emulate in their non-verbal self-presentation. Personas may also be a way to reconcile constructs with cues; individualized targets for non-verbal cues could be tied with personas that are associated with constructs (e.g., the Active Audience persona demonstrates engagement by having target posture for at least 150% more than your typical target posture time, but only for presentation meetings). Future work could also explore the evaluation of one's own non-verbal behaviour from the perspective of the other attendees to inform actionability.

## 6.5 Limitations

**Privacy Considerations & Self-Selection:** Due to the potentially intrusive nature of the approach, privacy was a key consideration. To ameliorate concerns, data processing for *Novecs* was done locally and sessions were started manually. While we did not explicitly prompt participants to think about potential privacy concerns, no participants mentioned any concerns around privacy or

feeling uncomfortable when using it. Although one reason may partially be due to self-selection biases [50], it is encouraging that people seem to be open towards using systems that leverage potentially sensitive facial data so long as they are designed to be privacy-preserving.

**Accuracy and Cultural Considerations:** A key limitation of any vision-based detection system is accuracy and the technologies it relies on [11]. Challenges with accuracy are likely exacerbated in deploying such a system in the field, with a variety of devices and environments where online work meetings would take place. A key issue is the experience of individuals with darker skin tones when it comes to such systems [22]. Unfortunately, while the study included individuals of various racial backgrounds, including Caucasian, East Asian, South Asian descent, we did not have any participants who identified as Black or Hispanic. While our study sampled participants from two countries with distinct cultures – [countries omitted for anonymity] – we did not observe noticeable differences in our findings between participants from either location. Cultural considerations were not a key factor in our study, but should be further explored, as there are known cultural differences in how people express themselves by their non-verbal cues [17, 76].

**Generalizability:** The sampling and number of participants, our choice of non-verbal cues, and especially the types and frequency of the meetings with *Novecs* enabled, might have impacted the generalizability of our findings. Findings on awareness impact may be influenced by the Hawthorne effect [77]. However, the aim of this exploratory work was not to achieve generalizability, but instead learn about participants' openness towards, as well as opportunities and challenges for real-time feedback. Further, the *Novecs* technology probe only supported four specific non-verbal cues. However, they are not necessarily the only cues that future *Novecs*-like systems should adopt, nor are they necessarily the ones that knowledge workers will find the most relevant in the work meeting context. The four cues simply served as a reasonable starting point for the *Novecs* technology probe. We also acknowledge that by including smiling and nodding (but not frowning or head shakes), we focus *Novecs* on capturing positivity and engagement. A future version could also consider negative non-verbal cues to capture confusion, frustration, or disagreement, but additional consideration would be needed in its design.

## 7 Conclusion

Our work improves the understanding of how real-time feedback systems might support non-verbal self-presentation, as well as the potential opportunities and challenges of such systems for online meetings. To that purpose, we deployed *Novecs* in an exploratory field study with 18 knowledge workers. *Novecs* is a technology probe of a feedback system that automatically detects and visualizes users' non-verbal cues during online work meetings via both real-time and summative approaches. We learned that participants' awareness of their non-verbal cues generally increased and that there are tensions around authenticity and in-authenticity, which *Novecs*' neutral feedback design was compatible with navigating. We also identified meeting types where real-time feedback on non-verbal cues was most useful. In an era of hybrid work, where online meetings are ubiquitous, the findings and design opportunities surfaced by *Novecs* can both inform and inspire future designers of such systems to empower users in enhancing their non-verbal self-presentation.

## Acknowledgments

The authors thank the anonymous reviewers, members of the UZH HASEL and the UBC eDAPT Lab for their valuable and insightful feedback. In particular, we would like to thank Remy Egloff for their involvement in developing the initial research prototype from which *Novecs* took inspiration. This work was supported by funding from *Designing interactive digital technologies for a post-pandemic world* (RGPIN-2023-045457) and *SNSF TaskFlow Grant* (207916).

## References

- [1] 2022. Microsoft is removing emotion recognition features from its facial recognition tech. <https://www.nbcnews.com/tech/tech-news/microsoft-removing-emotion-recognition-features-facial-recognition-tec-rca35087>
- [2] Sharifa Alghowinem, Roland Goecke, Michael Wagner, Gordon Parker, and Michael Breakspear. 2013. Eye movement analysis for depression detection. In *2013 IEEE International Conference on Image Processing*. IEEE, 4220–4224.
- [3] Mohammad Rafayet Ali, Dev Crasta, Li Jin, Agustin Baretto, Joshua Pachter, Ronald D. Rogge, and Mohammed Ehsan Hoque. 2015. LISSA – Live Interactive Social Skill Assistance. In *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, Xi'an, China, 173–179. doi:10.1109/ACII.2015.7344568
- [4] Mohammad Rafayet Ali and Ehsan Hoque. 2017. Social Skills Training with Virtual Assistant and Real-Time Feedback. In *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*. ACM, Maui Hawaii, 325–329. doi:10.1145/3123024.3123196
- [5] Robert M Arkin, James A Shepperd, R Giacalone, and P Rosenfeld. 1989. Self-presentation styles in organizations. *Impression management in the organization* 1, 1 (1989), 125–139.
- [6] Saygin Artiran, Leanne Chukoskie, Ara Jung, Ian Miller, and Pamela Cosman. 2021. HMM-based detection of head nods to evaluate conversational engagement from head motion data. In *2021 29th European Signal Processing Conference (EUSIPCO)*. IEEE, 1301–1305.
- [7] Bon Adriel Aseniero, Marios Constantinides, Sagar Joglekar, Ke Zhou, and Daniele Quercia. 2020. MeetCues: Supporting Online Meetings Experience. *2020 IEEE Visualization Conference (VIS)* (Oct. 2020), 236–240. doi:10.1109/VIS47514.2020.00054 arXiv:2010.06259
- [8] Ebrahim Babaei, Namrata Srivastava, Joshua Newn, Qiushi Zhou, Tilman Dingler, and Eduardo Velloso. 2020. Faces of Focus: A Study on the Facial Cues of Attentional States. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3313831.3376566
- [9] Jeremy N Bailenson. 2021. Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. (2021).
- [10] Madeline Balaam, Geraldine Fitzpatrick, Judith Good, and Eric Harris. 2011. Enhancing interactional synchrony with an ambient display. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 867–876.
- [11] Tadas Baltrusaitis, Amir Zadeh, Yao Chong Lim, and Louis-Philippe Morency. 2018. Openface 2.0: Facial behavior analysis toolkit. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*. IEEE, 59–66.
- [12] Roy F Baumeister. 1989. Motives and costs of self-presentation in organizations. *Impression management in the organization* 57 (1989), 71.
- [13] Roy F Baumeister and Debra G Hutton. 1987. Self-presentation theory: Self-construction and audience pleasing. In *Theories of group behavior*. Springer, 71–87.
- [14] Chloe Berger. [n.d.]. Workers Are Adopting 'Productivity Theater' to Impress Their Bosses. <https://fortune.com/2023/05/03/workers-adopting-productivity-theater-performative-work/>.
- [15] Rachel Bergmann, Sean Rintel, Nancy Baym, Advait Sarkar, Damian Borowiec, Priscilla Wong, and Abigail Sellen. 2023. Meeting (the) pandemic: Videoconferencing fatigue and evolving tensions of sociality in enterprise video meetings during COVID-19. *Computer Supported Cooperative Work (CSCW)* 32, 2 (2023), 347–383.
- [16] Simone Bianco, Luigi Celona, and Raimondo Schettini. 2016. Robust smile detection using convolutional neural networks. *Journal of Electronic Imaging* 25, 6 (2016), 063002–063002.
- [17] Pio Enrico Ricci Bitti and Pier Luigi Garotti. 2011. Nonverbal communication and cultural differences: Issues for face-to-face communication over the internet. *Face-to-face communication over the internet* (2011), 81–99.
- [18] Mark Bolino, David Long, and William Turnley. 2016. Impression management in organizations: Critical questions, answers, and areas for future research. *Annual Review of Organizational Psychology and Organizational Behavior* 3 (2016), 377–406.
- [19] Virginia Braun and Victoria Clarke. 2006. Using Thematic Analysis in Psychology. *Qualitative Research in Psychology* 3, 2 (Jan. 2006), 77–101. doi:10.1191/1478088706qp063oa
- [20] V. Braun and V. Clarke. 2021. *Thematic Analysis: A Practical Guide*. SAGE Publications. <https://books.google.ca/books?id=eMArEAAAQBAJ>
- [21] Jed R. Brubaker, Gina Venolia, and John C. Tang. 2012. Focusing on Shared Experiences: Moving beyond the Camera in Video Communication. In *Proceedings of the Designing Interactive Systems Conference (DIS '12)*. Association for Computing Machinery, 96–105. doi:10.1145/2317956.2317973
- [22] Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency*. PMLR, 77–91.
- [23] Byungki Byun, Anurag Awasthi, Philip A Chou, Ashish Kapoor, Bongshin Lee, and Mary Czerwinski. 2011. Honest signals in video conferencing. In *2011 IEEE International Conference on Multimedia and Expo*. IEEE, 1–6.



- [24] Jun-Ho Choi, Marios Constantinides, Sagar Joglekar, and Daniele Quercia. 2021. KAIROS: Talking heads and moving bodies for successful meetings. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*. 30–36.
- [25] Marios Constantinides and Daniele Quercia. 2022. Good Intentions, Bad Inventions: How Employees Judge Pervasive Technologies in the Workplace. *IEEE Pervasive Computing* 22, 1 (2022), 69–76.
- [26] Clayton R Cook, Andrew Jordan Thayer, Aria Fiat, and Margaret Sullivan. 2020. Interventions to enhance affective engagement. *Student engagement: Effective academic, behavioral, cognitive, and affective interventions at school* (2020), 203–237.
- [27] Ross Cutler, Yasaman Hosseinkashi, Jamie Pool, Senja Filipi, Robert Aichner, Yuan Tu, and Johannes Gehrke. 2021. Meeting Effectiveness and Inclusiveness in Remote Collaboration. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 173:1–173:29. doi:10.1145/3449247
- [28] Vedant Das Swain, Lan Gao, William A Wood, Srikruthi C Matli, Gregory D Abowd, and Munmun De Choudhury. 2023. Algorithmic power or punishment: Information worker perspectives on passive sensing enabled ai phenotyping of performance and wellbeing. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [29] Bella M DePaulo. 1992. Nonverbal behavior and self-presentation. *Psychological bulletin* 111, 2 (1992), 203.
- [30] Sidney D'Mello, Patrick Chipman, and Art Graesser. [n. d.]. Posture as a Predictor of Learner's Affective Engagement. ([n. d.]), 7.
- [31] Sidney D'Mello, Andrew Olney, Claire Williams, and Patrick Hays. 2012. Gaze Tutor: A Gaze-Reactive Intelligent Tutoring System. *International Journal of Human-Computer Studies* 70, 5 (May 2012), 377–398. doi:10.1016/j.ijhcs.2012.01.004
- [32] Sidney K. D'Mello and Arthur Graesser. 2010. Multimodal Semi-Automated Affect Detection from Conversational Cues, Gross Body Language, and Facial Features. *User Modeling and User-Adapted Interaction* 20, 2 (June 2010), 147–187. doi:10.1007/s11257-010-9074-4
- [33] Andrew J DuBrin. 2010. *Impression management in the workplace: Research, theory and practice*. Routledge.
- [34] Amy Edmondson. 1999. Psychological Safety and Learning Behavior in Work Teams. *Administrative Science Quarterly* 44, 2 (June 1999), 350–383. doi:10.2307/2666999
- [35] Paul Ekman and Wallace V Friesen. 1978. Facial Action Coding System. *Environmental Psychology & Nonverbal Behavior* (1978).
- [36] Naomi Ellemers and Manuela Barreto. 2006. Social identity and self-presentation at work: how attempts to hide a stigmatised identity affect emotional well-being, social inclusion and performance. *Netherlands Journal of Psychology* 62, 1 (2006), 51–57.
- [37] Laura L. Ellingson. 2009. *Engaging Crystallization in Qualitative Research: An Introduction*. SAGE.
- [38] 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. doi:10.1016/j.ijhcs.2020.102411
- [39] Heather A Faucett, Matthew L Lee, and Scott Carter. 2017. I should listen more: real-time sensing and feedback of non-verbal communication in video telehealth. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–19.
- [40] G Fauville, M Luo, ACM Queiroz, A Lee, JN Bailenson, and J Hancock. 2023. Video-conferencing usage dynamics and nonverbal mechanisms exacerbate Zoom Fatigue, particularly for women. *Computers in Human Behavior Reports* 10 (2023), 100271.
- [41] Geraldine Fauville, Mufan Luo, Anna CM Queiroz, Jeremy N Bailenson, and Jeff Hancock. 2021. Zoom exhaustion & fatigue scale. *Computers in Human Behavior Reports* 4 (2021), 100119.
- [42] Wallace V Friesen, Paul Ekman, et al. 1983. EMFACS-7: Emotional Facial Action Coding System. *Unpublished manuscript, University of California at San Francisco* 2, 36 (1983), 1.
- [43] O. Garcia, J. Favela, and R. Machorro. 1999. Emotional awareness in collaborative systems. In *6th International Symposium on String Processing and Information Retrieval. 5th International Workshop on Groupware (Cat. No.PR00268)*. 296–303. doi:10.1109/SPIRE.1999.796607
- [44] William L Gardner and Mark J Martinko. 1988. Impression management in organizations. *Journal of management* 14, 2 (1988), 321–338.
- [45] Jim Gemmell, Kentaro Toyama, C Lawrence Zitnick, Thomas Kang, and Steven Seitz. 2000. Gaze awareness for video-conferencing: A software approach. *IEEE MultiMedia* 7, 4 (2000), 26–35.
- [46] Daniela Girardi, Filippo Lanubile, Nicole Novielli, and Alexander Serebrenik. 2021. Emotions and Perceived Productivity of Software Developers at the Workplace. *IEEE Transactions on Software Engineering* (2021), 1–1. doi:10.1109/TSE.2021.3087906
- [47] Erving Goffman et al. 2002. The presentation of self in everyday life. 1959. *Garden City, NY* 259 (2002).

- [48] AL Hartzler, RA Patel, M Czerwinski, W Pratt, A Roseway, N Chandrasekaran, and A Back. 2014. Real-time feedback on nonverbal clinical communication. *Methods of information in medicine* 53, 05 (2014), 389–405.
- [49] A. L. Hartzler, R. A. Patel, M. Czerwinski, W. Pratt, A. Roseway, N. Chandrasekaran, and A. Back. 2014. Real-Time Feedback on Nonverbal Clinical Communication: Theoretical Framework and Clinician Acceptance of Ambient Visual Design. *Methods of Information in Medicine* 53, 05 (2014), 389–405. doi:10.3414/ME13-02-0033
- [50] James J Heckman. 1990. Selection bias and self-selection. In *Econometrics*. Springer, 201–224.
- [51] Leo Hemamou, Ghazi Felhi, Jean-Claude Martin, and Chloe Clavel. 2019. Slices of Attention in Asynchronous Video Job Interviews. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, Cambridge, United Kingdom, 1–7. doi:10.1109/ACII.2019.8925439
- [52] Javier Hernandez, Zicheng Liu, Geoff Hulten, Dave DeBarr, Kyle Krum, and Zhengyou Zhang. 2013. Measuring the Engagement Level of TV Viewers. In *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)*. IEEE, Shanghai, China, 1–7. doi:10.1109/FG.2013.6553742
- [53] Pamela J Hinds. 1999. The cognitive and interpersonal costs of video. *Media Psychology* 1, 4 (1999), 283–311.
- [54] Mohammed (Ehsan) 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*. ACM, Zurich Switzerland, 697–706. doi:10.1145/2493432.2493502
- [55] Mohammed Ehsan Hoque and Rosalind W. Picard. 2014. Rich Nonverbal Sensing Technology for Automated Social Skills Training. *Computer* 47, 4 (April 2014), 28–35. doi:10.1109/MC.2014.98
- [56] Esther Howe, Jina Suh, Mehrab Bin Morshed, Daniel McDuff, Kael Rowan, Javier Hernandez, Marah Ihab Abdin, Gonzalo Ramos, Tracy Tran, and Mary P Czerwinski. 2022. Design of Digital Workplace Stress-Reduction Intervention Systems: Effects of Intervention Type and Timing. In *CHI Conference on Human Factors in Computing Systems*. ACM, New Orleans LA USA, 1–16. doi:10.1145/3491102.3502027
- [57] 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.
- [58] Ellen A Isaacs and John C Tang. 1993. What video can and can't do for collaboration: a case study. In *Proceedings of the first ACM International Conference on Multimedia*. 199–206.
- [59] Rateb Jabbar, Mohammed Shinoy, Mohamed Kharbeche, Khalifa Al-Khalifa, Moez Krichen, and Kamel Barkaoui. 2020. Driver Drowsiness Detection Model Using Convolutional Neural Networks Techniques for Android Application. In *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIOT)*. IEEE, Doha, Qatar, 237–242. doi:10.1109/ICIOT48696.2020.9089484
- [60] Natasha Jaques and B Hon. [n. d.]. Predicting Affect in an Intelligent Tutoring System. ([n. d.]), 137.
- [61] Matthew K. Miller, Martin Johannes Dechant, and Regan L. Mandryk. 2021. Meeting You, Seeing Me: The Role of Social Anxiety, Visual Feedback, and Interface Layout in a Get-to-Know-You Task via Video Chat.. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.
- [62] Ashish Kapoor and Rosalind W Picard. 2001. A real-time head nod and shake detector. In *Proceedings of the 2001 workshop on Perceptive user interfaces*. 1–5.
- [63] S. Kawato and J. Ohya. 2000. Real-time detection of nodding and head-shaking by directly detecting and tracking the "between-eyes". In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*. 40–45. doi:10.1109/AFGR.2000.840610
- [64] Taemie Kim, Agnes Chang, Lindsey Holland, and Alex Sandy Pentland. 2008. Meeting mediator: enhancing group collaboration using sociometric feedback. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 457–466.
- [65] Brendan F Klare, Ben Klein, Emma Taborsky, Austin Blanton, Jordan Cheney, Kristen Allen, Patrick Grother, Alan Mah, and Anil K Jain. 2015. Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 1931–1939.
- [66] Predrag Klasnja, Sunny Consolvo, and Wanda Pratt. 2011. How to Evaluate Technologies for Health Behavior Change in HCI Research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Vancouver BC Canada, 3063–3072. doi:10.1145/1978942.1979396
- [67] Anthony C Klotz, Wei He, Kai Chi Yam, Mark C Bolino, Wu Wei, and Lawrence Houston III. 2018. Good actors but bad apples: Deviant consequences of daily impression management at work. *Journal of Applied Psychology* 103, 10 (2018), 1145.
- [68] Aleksandra Kostic and Derek Chadee. 2014. *The social psychology of nonverbal communication*. Springer.
- [69] Kristine M Kuhn. 2022. The constant mirror: Self-view and attitudes to virtual meetings. *Computers in Human Behavior* 128 (2022), 107110.
- [70] Olga Kulyk, Jimmy Wang, and Jacques Terken. 2006. Real-time feedback on nonverbal behaviour to enhance social dynamics in small group meetings. In *Machine Learning for Multimodal Interaction: Second International Workshop*,

*MLMI 2005, Edinburgh, UK, July 11-13, 2005, Revised Selected Papers 2*. Springer, 150–161.

- [71] 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.
- [72] Chunfeng Liu, Rafael A. Calvo, and Renee Lim. 2016. Improving Medical Students' Awareness of Their Non-Verbal Communication through Automated Non-Verbal Behavior Feedback. *Frontiers in ICT* 3 (2016).
- [73] Jennifer Marlow, Eveline Van Everdingen, and Daniel Avrahami. 2016. Taking notes or playing games? Understanding multitasking in video communication. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. 1726–1737.
- [74] Brais Martinez and Michel F Valstar. 2016. Advances, challenges, and opportunities in automatic facial expression recognition. *Advances in face detection and facial image analysis* (2016), 63–100.
- [75] Marianne Schmid Mast. 2007. On the importance of nonverbal communication in the physician–patient interaction. *Patient education and counseling* 67, 3 (2007), 315–318.
- [76] David Matsumoto and Hysisung C Hwang. 2016. The cultural bases of nonverbal communication. (2016).
- [77] Jim McCambridge, John Witton, and Diana R. Elbourne. 2014. Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology* 67, 3 (2014), 267–277. doi:10.1016/j.jclinepi.2013.08.015
- [78] Daniel McDuff, Rana El Kaliouby, Jeffrey F. Cohn, and Rosalind W. Picard. 2015. Predicting Ad Liking and Purchase Intent: Large-Scale Analysis of Facial Responses to Ads. *IEEE Transactions on Affective Computing* 6, 3 (July 2015), 223–235. doi:10.1109/TAFFC.2014.2384198
- [79] Matthew K Miller, Regan L Mandryk, Max V Birk, Ansgar E Depping, and Tushita Patel. 2017. Through the looking glass: The effects of feedback on self-awareness and conversational behaviour during video chat. In *Proceedings of the 2017 CHI conference on human factors in computing systems*. 5271–5283.
- [80] Sarah Morrison-Smith, Jaime Ruiz, and Lydia Chilton. 2021. AmbiTeam: Providing Team Awareness Through Ambient Displays. In *Graphics Interface 2021*.
- [81] Katja Neureiter, Martin Murer, Verena Fuchsberger, and Manfred Tscheligi. 2013. Hand and eyes: how eye contact is linked to gestures in video conferencing. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. 127–132.
- [82] Nicole Novielli, Daniela Grassi, Filippo Lanubile, and Alexander Serebrenik. [n. d.]. Sensor-Based Emotion Recognition in Software Development: Facial Expressions as Gold Standard. ([n. d.]), 8.
- [83] Jessica A Peck and Mary Hogue. 2018. Acting with the best of intentions... or not: A typology and model of impression management in leadership. *The Leadership Quarterly* 29, 1 (2018), 123–134.
- [84] Monica Pereira and Kate Hone. 2021. Communication Skills Training Intervention Based on Automated Recognition of Nonverbal Signals. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–14. doi:10.1145/3411764.3445324
- [85] Sowmya Rasipuram and Dinesh Babu Jayagopi. 2019. A Comprehensive Evaluation of Audio-Visual Behavior in Various Modes of Interviews in the Wild. In *Proceedings of the 12th ACM International Conference on Pervasive Technologies Related to Assistive Environments*. ACM, Rhodes Greece, 94–100. doi:10.1145/3316782.3321528
- [86] Sowmya Rasipuram and Dinesh Babu Jayagopi. 2020. Automatic Multimodal Assessment of Soft Skills in Social Interactions: A Review. *Multimedia Tools and Applications* 79, 19 (May 2020), 13037–13060. doi:10.1007/s11042-019-08561-6
- [87] Martin S Remland. 2016. *Nonverbal communication in everyday life*. Sage Publications.
- [88] 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 and Inclusive Meetings. In *CHI 2021*.
- [89] Samiha Samrose, Ru Zhao, Jeffery White, Vivian Li, Luis Nova, Yichen Lu, Mohammad Ali, and Ehsan Hoque. 2018. *CoCo: Collaboration Coach for Understanding Team Dynamics during Video Conferencing*. Vol. 1. 24 pages. doi:10.1145/3161186
- [90] Allison Sauppé and Bilge Mutlu. 2014. How Social Cues Shape Task Coordination and Communication. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '14)*. Association for Computing Machinery, 97–108. doi:10.1145/2531602.2531610
- [91] Barry R Schlenker. 2003. Self-presentation. *Handbook of self and identity* 2 (2003), 542–570.
- [92] Kristen M Shockley, Allison S Gabriel, Daron Robertson, Christopher C Rosen, Nitya Chawla, Mahira L Ganster, and Maira E Ezerins. 2021. The fatiguing effects of camera use in virtual meetings: A within-person field experiment. *Journal of Applied Psychology* 106, 8 (2021), 1137.
- [93] Tereza Soukupova. [n. d.]. Real-Time Eye Blink Detection Using Facial Landmarks. ([n. d.]), 8.

- [94] Willem Standaert, Steve Muylle, and Amit Basu. 2021. How shall we meet? Understanding the importance of meeting mode capabilities for different meeting objectives. *Information & Management* 58, 1 (2021), 103393.
- [95] Wenzhao Tan and Gang Rong. 2003. A real-time head nod and shake detector using HMMs. *Expert Systems with Applications* 25, 3 (2003), 461–466.
- [96] M Iftekhar Tanveer, Emy Lin, and Mohammed Hoque. 2015. Rhema: A real-time in-situ intelligent interface to help people with public speaking. In *Proceedings of the 20th international conference on intelligent user interfaces*. 286–295.
- [97] Yla R Tausczik and James W Pennebaker. 2013. Improving teamwork using real-time language feedback. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 459–468.
- [98] Jacques Terken and Janienke Sturm. 2010. Multimodal support for social dynamics in co-located meetings. *Personal and Ubiquitous Computing* 14 (2010), 703–714.
- [99] Sarah J. Tracy. 2010. Qualitative Quality: Eight “Big-Tent” Criteria for Excellent Qualitative Research. *Qualitative Inquiry* 16, 10 (Dec. 2010), 837–851. doi:10.1177/1077800410383121
- [100] Alessandro Vinciarelli, Maja Pantic, Hervé Bourlard, and Alex Pentland. 2008. Social Signals, Their Function, and Automatic Analysis: A Survey. In *Proceedings of the 10th International Conference on Multimodal Interfaces*. ACM, Chania Crete Greece, 61–68. doi:10.1145/1452392.1452405
- [101] Joseph B Walther. 2007. Selective self-presentation in computer-mediated communication: Hyperpersonal dimensions of technology, language, and cognition. *Computers in Human Behavior* 23, 5 (2007), 2538–2557.
- [102] Joseph B Walther, Brandon Van Der Heide, Artemio Ramirez Jr, Judee K Burgoon, and Jorge Peña. 2015. Interpersonal and hyperpersonal dimensions of computer-mediated communication. *The handbook of the psychology of communication technology* (2015), 1–22.
- [103] Robert A. Wicklund. 1979. The Influence of Self-Awareness on Human Behavior: The Person Who Becomes Self-Aware Is More Likely to Act Consistently, Be Faithful to Societal Norms, and Give Accurate Reports about Himself. *American Scientist* 67, 2 (1979), 187–193. jstor:27849150
- [104] Chi-Sum Wong and Kenneth S Law. 2002. Wong and law emotional intelligence scale. *The leadership quarterly* (2002).
- [105] Beste F Yuksel, Penny Collisson, and Mary Czerwinski. 2017. Brains or beauty: How to engender trust in user-agent interactions. *ACM Transactions on Internet Technology (TOIT)* 17, 1 (2017), 1–20.
- [106] Xucong Zhang, Yusuke Sugano, Mario Fritz, and Andreas Bulling. 2015. Appearance-based Gaze Estimation in the Wild. In *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 4511–4520.

Received January 2024; revised July 2024; accepted October 2024