Mirror Hearts: Exploring the (Mis-)Alignment between AI-Recognized and Self-Reported Emotions

Si Chen
sic3@illiois.edu
Information Sciences, University of Illinois at Urbana–Champaign
Champaign, Illinois, USA

Haocong Cheng
haocong2@illinois.edu
Information Sciences, University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

Jason Situ
junsitu2@illinois.edu
Computer Science, University of Illinois at Urbana-Champaign
Urbana, Illinois, USA

Yun Huang
yunhuang@illinois.edu
Information Sciences, University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

ABSTRACT

Previous research has found that learners' reflections on their own emotions can improve their learning experience, and AI-based technologies can be used to automatically recognize learners' emotions. We conducted user studies involving 32 participants to investigate the relationship between AI-recognized emotions and their self-reported emotions using emojis and text comments. We found that, even though AI recognized a similar amount of positive-high-arousal and negative-low-arousal emotions, participants self-reported more positive-high-arousal emojis. Participants explained the causality and temporality of the self-reported emojis using text comments. Our findings suggest ways of using AI to capture learners' emotions and support their reflections.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; • **Applied computing** → *E-learning*.

KEYWORDS

Facial Recognition, Emoji, Reflection, Self-Regulated Learning

ACM Reference Format:
Si Chen, Haocong Cheng, Jason Situ, and Yun Huang. 2023. Mirror Hearts: Exploring the (Mis-)Alignment between AI-Recognized and Self-Reported Emotions. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3544549.3585607

1 INTRODUCTION

Recent research has explored the use of artificial intelligence (AI) technology called automatic emotion recognition to recognize learners' emotions via faces and promote reflections [1, 4, 6, 12, 15, 28, 31]. Although automatic emotion recognition can be used as observed emotions, it should not replace internally reported emotions [19]. Emojis, as a time-saving approach, are used widely in existing works for self-reporting emotions [14, 32, 34]. However, there is a limited understanding of how learners utilize self-reported emotions, such as emojis, in combination with AI-recognized emotions, during reflection.

This study focused on analyzing how learners use a post-learning dashboard to reflect on their emotions during self-regulated video-based learning. We visualized self-reported emotions (emojis and text) and AI-recognized emotions (facial expressions) to support learners' reflections. Self-regulated learning (SRL) is a conceptual framework to understand the cognitive, motivational, and emotional aspects of learning [29, 35]. Reflection is a process of SRL as it allows individuals to evaluate, analyze and observe their own learning and performance. Our study answered the following research question:

RQ: What is the relationship between self-reported emotions (during post-viewing) & AI-recognized emotions (via real-time facial expressions detection) in self-regulated learning (SRL)?

2 RELATED WORKS

Learners' awareness of their own emotions is important for SRL. Learning analytic tools, such as [13, 29], provide information visualization on the course schedule, student pace, and learning assessments, but they do not provide the capability to monitor one's own emotions using automatic emotion recognition. This study incorporates self-reported emotions with AI-recognized emotions to induce a reasoning process in comparing AI-recognized emotions and self-reported emotions. This research is important for improving SRL in video-based learning, which was found to be associated with higher levels of video engagement and learning performance [20]. Emotions are typically represented in the form of discrete emotion categories [11] or continuous dimensional models such as the circumplex model [30], which represents human emotions in a two-dimensional arousal-valence circular space. Recent automatic emotion recognition technologies enable the prediction of arousal-valence from facial expressions [4, 7, 8, 22, 27]. A state-of-the-art arousal-valence regression algorithm returns time-based arousal-valence pairs [7], which was used in our study.

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585607

Observed and self-reported emotions are distinct signals [19]. Observed emotions are seen as a third-person perspective of monitoring one's own emotions, while self-reported emotions are seen as a first-person perspective of describing internal feelings. Kim et al. [21] found that self-awareness was increased when observing facial expressions differ from what was internally perceived (from a first-person view). On the other hand, self-reported emotions from a first-person perspective in text-based explanation are relatively inexpensive and require minimal resources but may not accurately reflect the actual experience of the person, including exaggeration of symptoms or minimizing them [3].

Emojis are a popular way to express emotions in online communication rather than textual explanations [2, 3]. Ghaban [14] investigated the effect of emoji usage in online learning settings, while Schouteten and Meiselman [32] looked at the potential of using emojis in questionnaires in response to food-related emotions. Toet et al. [34] presented the Emoji grid as an alternative to the Self-Assessment Manikin for self-reporting emotions in virtual reality experiences. However, there is still limited research on how learners use self-reported emotions together with observed AI-recognized emotions for reflections, particularly in the form of emojis and text comments. To address this, we performed a user study using a post-learning dashboard.

3 INTERACTION AND USER STUDY DESIGN

3.1 Interaction Design: Aligning Self-Reported Emojis and AI-Recognized Emotions via Facial Expressions to support Reflections

Our design goal of the post-learning dashboard was to enable learners to compare discrepancies between self-reported and AI-recognized emotions on a time scale. Below, we describe in detail how the interface collects and displays AI-recognized emotions and self-reported emotions. Our reason for including self-reports is not to understand the accuracy of the AI but to support the process of post-learning reflections in SRL and understand how AI can support reflections.

3.1.1 Collecting and Displaying AI-recognized Emotions. We obtained AI-recognized emotions by gathering video streams of the learner watching an assigned video, cropping their faces, and feeding the post-processed images into an arousal-valence regression module (Fig. 1 ①). The resulting arousal-valence value pairs were then represented in a temporal chart that displays AI-recognized emotions as data points (Fig. 1 ③(c)). We remapped the pair values in the arousal-valence 2D space into polar space represented by intensity and angle referring to [6, 18, 23], as shown in Fig 1 ③(c). Additionally, we provided a corresponding legend of arousal-valence 2D space on the interface to interpret the emotions, as shown in Fig 1 ③(e). The legend used a continuous rainbow with cool colors representing negative results and warm colors representing positive results to depict the relationship between arousal and valence. The legend has two dimensions that refer to valence (positive/negative) and arousal (high/low). The performance of the state-of-the-art models employed in our interaction was evaluated using Concordance Correlation Coefficient, yielding a score of 0.440 for valence and 0.454 for arousal. To ensure user privacy, the captured images and AI-recognized results from a user could not be seen by any other users. The interface also indicated to the participants when the webcam was used by the AI. Examples of time-based curves from participants are shown in Fig. 2, Fig. 3, and Fig. 4.

3.1.2 Self-Reporting and Displaying Emotions. To understand how AI can support learners in their SRL and the potential of AI-enabled reflections, we included self-reported emotions in our study. We focused on post-learning self-reports to avoid distracting the learning process and in accordance with prior literature that noted post-learning reflection as an important stage in the learning process [29].

To obtain self-reported emotions, the learner was asked to report using emojis and text-based comments that represent the moment they experienced emotions during video-watching (Fig. 1 ②). We collected self-reports in the post-learning phase to avoid distraction while watching the video. We provided ten emoji options: 🥱, 😎, 😍, 😌, 😐, 😭, 😀, 😢, 😕, 😲. We referred to [24, 34] to select these emojis and to locate them on the valence and arousal grid. A color wheel titled "Emotion Legend" (Fig. 1 ③(e)) was provided to users in step 2 and step 3 as a reference for representing the emoji. The interface explicitly explained to the user that the emojis roughly represent nearby emotions. The self-reports were displayed in step 3 (shown in Fig 1 ③(d)) temporally based on the timestamp of each self-report. Users could hover on a tick to see the emoji and text comments in a self-report.

3.2 User Study Design and Data Analysis

We conducted a user study with 32 participants, who were undergraduate and graduate students (18 female, 14 male). Each study consisted of four steps: 1) self-regulated video-based learning with AI interpreting facial expressions, 2) self-reporting learning emotions, 3) thinking aloud with AI-recognized and self-reported emotions (shown in Fig. 1 ③), and 4) end of study interview. All studies were conducted online through Zoom and lasted around one hour. Participants were compensated with 10 USD each hour. During the first step, participants watched a 15-minute video introducing Augmented Reality (AR) technology while the automatic emotion recognition module was activated to detect and recognize their facial expressions. In the second step, an interface was provided to allow them to self-report learning emotions, which took around 5 minutes. In the third step, the participants were presented with a post-learning dashboard which visualized their own facial recognition results and self-reports. Participants were asked to "think-aloud" while reviewing and exploring the results, which took around 10 minutes. The study concluded with an end-of-study interview, which took around 20 minutes.

To answer our RQ, we aggregated self-reports to analyze trends in emoji usage and compared self-reported emojis with AI-recognized emotions and self-reported text comments, and analyzed think-aloud, interview recordings to understand participants' strategies. We grouped the emojis on the arousal-valence grid as positive-high-arousal (😎, 😲, 😍, and 😀) and negative-low-arousal 😭, 😕, and 🥱) in reference to [24, 34], because such emotions are important learning emotions in SRL [10]. We performed paired t-tests

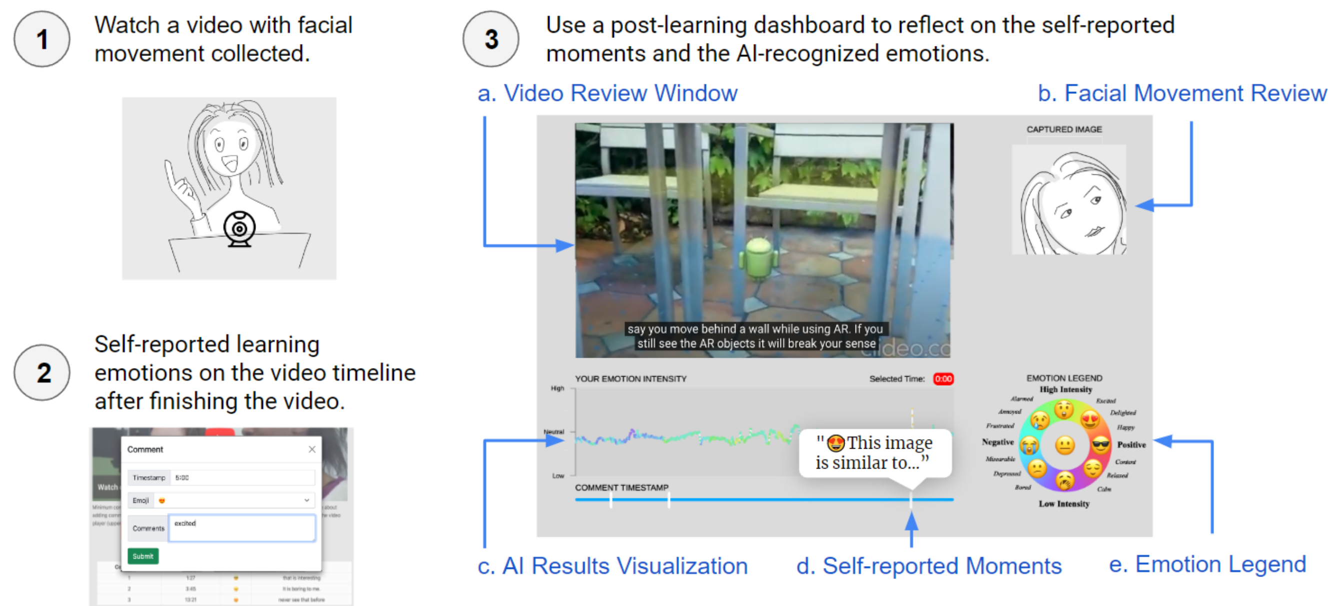

Figure 1: The SRL process in our learning analytic tool consists of three steps. Learners follow the order of each step in our learning analytic tool. ① A user watches an online video alone with their webcam turned on. The AI captures the user's facial expression when watching the video. ② After the video ends, the user uses this interface to make self-reports. The user can select a timestamp on the video player, then select an emoji and write text comments. However, the user may choose not to select an emoji or write any text comments. ③ Post-Learning Dashboard. The user can review their own AI-recognized emotions and self-reports on this interface for reflections. The user is able to use the video player (a) to review different sections of the same video. When a timestamp is selected, the user is able to see their captured facial expression image in (b). The AI-recognized emotions are visualized over the length of the video in (c). The height of the data points demonstrates the intensity of the specific states, and the color shows the recognized emotions. The timestamps of self-reports are visualized in (d). The user can hover on a tick to see the selected emoji and text comments at that timestamp. A color wheel titled "Emotion Legend" is provided to help the user interpret the represented emotions of the color in AI-recognized emotions and the self-reported emojis, as shown in (e).

to understand the differences between self-reported emotions and AI-recognized emotions. All tests performed in the analysis met the assumptions of the t-test, including that the variability of the data in each group was similar and the distribution was approximately normal. We also conducted thematic analysis [5] on self-report, think-aloud, and interview transcripts. To distinguish among the different transcripts, we denote quotes from text comments in self-reports as italic (e.g., *self-reports*), double-quoted for think-aloud transcripts (e.g., "think-aloud"), and double-quoted with italic for interview transcripts (e.g., "*interview*") in the rest of the paper.

4 FINDINGS

In total, 259 self-reports were created by 32 participants (mean = 8.09, std = 5.57); and 248 (95.75 %) of these self-reports included emojis. All participants used at least one emoji in their self-reports.

4.1 (Mis-)alignment between AI-Recognized and Self-reported Emotions using Emojis

4.1.1 Different Distributions between AI-Recognized Emotion and Self-reported Emojis. AI and self-reported emojis captured different emotions. We conducted a paired t-test to compare the distributions of AI and self-reported emojis to determine how they captured positive-high-arousal and negative-low-arousal emotions. The test results showed that, on average, AI detected an equal percentage of positive-high-arousal and negative-low-arousal emotions, while each positive-high-arousal emoji was 1.47 times more likely to be self-reported than a negative-low-arousal emoji. Specifically, the ratio of positive-high-arousal emojis to negative-low-arousal emojis (Ratio-Emoji) was significantly higher (M=2.47, SD=2.31) than the ratio of frames indicating positive-high-arousal to those indicating negative-low-arousal (Ratio-AI) (M=0.96, SD=2.18) ($p = .003$, $<.05$).

We further explored the difference in self-reported emojis. On average, each participant used each of the positive-high-arousal emojis in 1.18 (SD = 0.67) self-reports, which was significantly more than the 0.59 (SD = 0.35) for each negative-low-arousal emoji according to paired t-test, $t(31)=3.76$, $p = .0007$, $< .001$. Two participants did not report any positive-high-arousal emoji, and eight participants did not report any negative-low-arousal emoji.

4.1.2 Participant's Reflection Using Temporally (Mis-)aligned Self-Report and AI-Recognized Emotions (Think-aloud). During think-aloud, we observed that participants tended to explore parts of the AI-recognized results when they made self-reports or when they

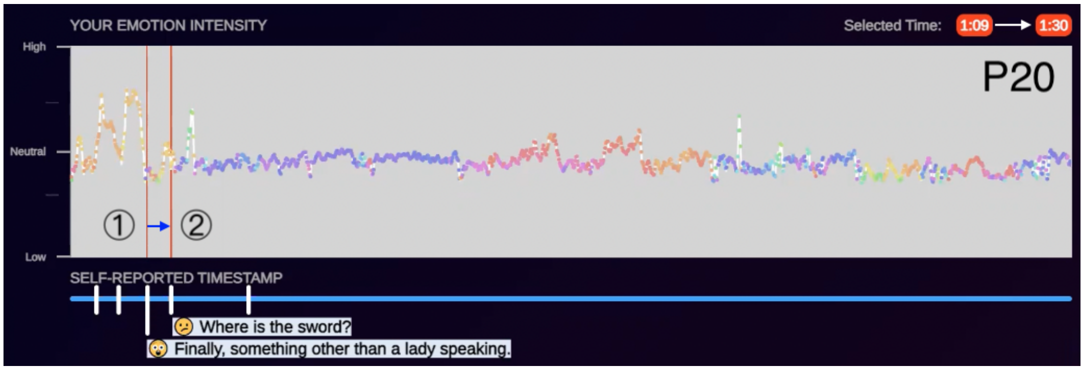

Figure 2: Case study of AI-recognized emotions and self-report timestamps for P20, who reflected his two self-reports and AI-recognized emotions from timestamp 1 (1:09) to timestamp 2 (1:30). He was able to recall his feelings when watching the video.

noticed a spike or color change. They were able to review the video clips around these timestamps and reflect on why they had such emotions at that moment. A total of 401 reflection timestamps were collected in the think-aloud sessions. A reflection timestamp refers to participants clicking a self-report and/or on the AI-recognized graph (red line pops up on the interface as shown in Fig. 3).

Reflecting on Temporally Aligned Emotions. In two-thirds of the reflection timestamps, participants were able to find a temporal alignment between their self-reported emotions and AI-recognized emotions. For example, P20 was able to start a reflection from his self-report at 1:11 (😲 *Finally, something other than a lady speaking.*). Since he was "waiting for a figure to show up" rather than seeing only the presenter speaking, he could immediately locate this timestamp and thought-aloud, "there is that figure showing up." At the same time, the AI was able to capture the change in facial expression, as shown in Fig 2. He then noticed that his AI-recognized emotions changed at around 1:14. He thought-aloud, "I'm really attracted to the figure, and I'm thinking 'Oh, what is this here,' but I think my attention has been drawn too much into the figure that I kind of ignored what the narrator was saying." He continued to review his next self-report (😕 *Where is the sword?*) at 1:32, accompanied by more changes in AI-recognized emotions when a new term was introduced on the screen. He reviewed the video and further reflected that this is when he focused on the narrator's speech again after the visual demonstration ended, and he was confused by the new terms introduced on the screen.

Reflecting on Emotions with No Temporal Alignment. In the remaining one-third of the reflections, there was no alignment between self-reports and AI-recognized emotions; therefore, they explored these mismatched cases to start reflection. In some cases, they noticed a change in AI-recognized emotions but did not see self-reports at that timestamp. For example, P24 noticed that her emotion changed around 11:51, but she did not make a self-report, as shown in Fig 3. She thought-aloud, "Let me see what happened here. My emotions changed from some green color to purple and pink color. What happened?" She then reviewed the video clip and was able to recall what she saw, "Ah, I saw a very, very cute pizza standing in front of ... the room. I've never thought AR could be so adapted in real [world]. So cute." In other cases, they found self-reports when there was no obvious change in AI-recognized emotions. For example, as shown in Fig 4, P21 reflected on his self-report at 3:59 (😀 *This is something that a feel a positive emotion about because it's something I haven't heard about with google. This is new information to me that makes me feel pretty happy to learn about because I think it's cool when companies work on new projects involving future technology.*). He thought-aloud, "A lot of these captured images are really just me watching the video. I don't really show much emotion when I'm watching videos, I guess. But seeing the captured images kind of looks like I'm just sitting there bored, but I was actually pretty intrigued learning that there is a new thing that Google was working on, and it's something I've never heard before, which is cool."

4.2 Explaining the (Mis-)alignment between AI-Recognized Emotions and Self-Reported Emojis

In this section, we intend to explain the differences between AI-recognized and self-report emotions by looking at the commonly reported emojis and text comments and triangulating them with interview results.

4.2.1 Commonly Self-Reported Emotions. Participants choose positive-high-arousal emojis in 58.69 % of their self-reports to express their excitement or happiness from the video. Fig. 5 shows how each emoji represents the valence and arousal, and the relative frequencies of them being used by the participants. These emotions were more likely to be remembered and reported. As shown in Fig. 6, we observed more self-reports using high arousal emoji (😲) between around 8:00 and 9:00 when the video showed a demo of AR application in classrooms. For example, P22 made a self-report at 8:24 : 😲 *I did not expect AR can be used in Education in a high extent and the scene of students holding phones to explore educational content through AR was interesting to watch.* During the interview, participants perceived that they were more aware of this topic since it was more relevant to them as students. They also mentioned that they tended to report such "*aha*" moments in their self-reports.

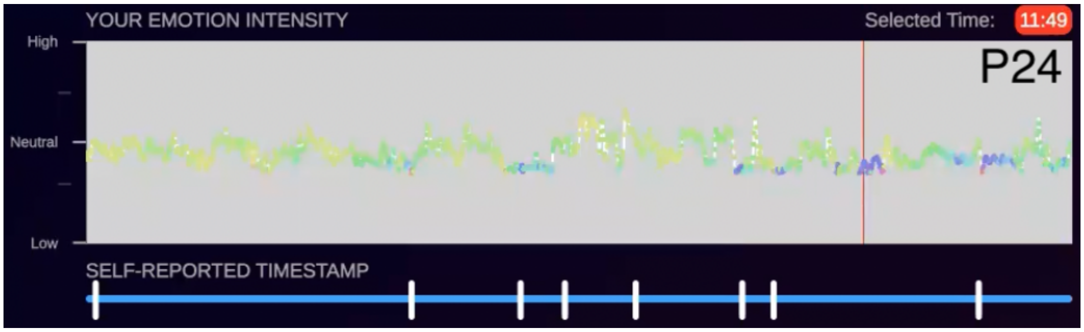

Figure 3: Case study of AI-recognized emotions and self-report timestamps for P24, who reflected at around 11:49 when she did not make self-report earlier but noticed a change in AI-recognized emotions. By comparing AI-recognized emotions, she was able to recall an interesting part of the video that she missed in her self-reports.

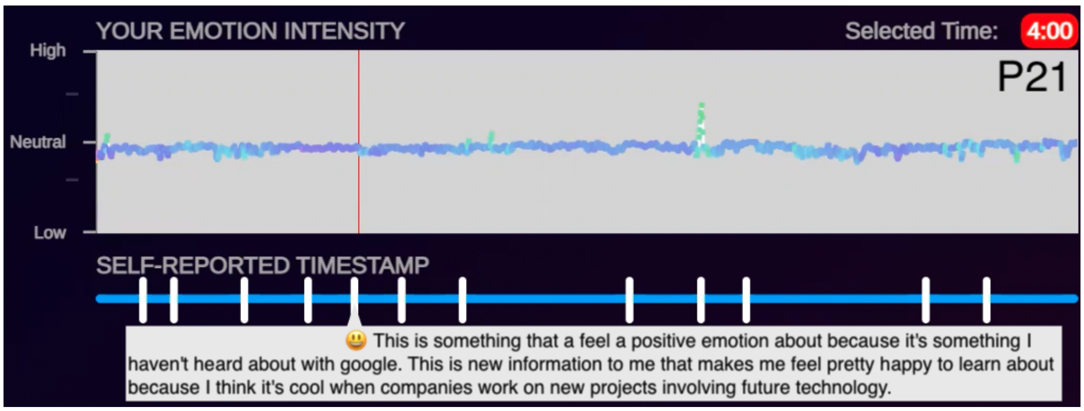

Figure 4: Case study of AI-recognized emotions and self-report timestamps for P21, who reflected at around 4:00 when he made self-report but did not see any change in AI-recognized emotions. He noticed a mismatch between his facial expression and his actual feeling when watching the video.

Participants also reported negative-low-arousal emojis, which made up 22.01 % of the self-reports, significantly less than positive-high-arousal emojis, despite that both emotions were equally recognized by the AI as presented in Section 4.1.1. By reviewing the usage of the low arousal emoji (🥱), we found that it was more frequently used in the first three minutes of the video, as shown in Fig. 6. During the interview, one-third of the participants debriefed that memory and attention impacted their ability to notice and report their own emotions, and low arousal was less noticeable and less likely to be reported when they felt sleepy or bored. For instance, P1 said during the interview: "*I am less aware of my emotions and video content when I get sleepy or bored... you can't pay attention or remember to anything when you're about to fall asleep and that is where I think AI can intervene and alert me.*"

4.2.2 Explaining the Self-Reported Emoji Using Text. Out of 259 self-reports, 122 (47.10 %) used text comments in addition to emojis to report their emotions while watching the video.

Explaining Causality of Emotions. Participants leveraged text comments in their self-reports to explain why they chose the selected emojis. We found that 42.86 % of the comments expressed causality using phrases such as *cause*, *because*, and *make me*. For example, at 9:01, P30 self-reported: 🥱 *I got bored again, i think because it was a long video.* P27 self-reported at 0:45: 😀 *I chose this emoji because I'm intrigued that they brought up vr [virtual reality].* By conducting a paired t-test, we did not find any difference between the percentage of self-reports with comments that had causality explanations for positive-high-arousal emoji and percentage for negative-low-arousal emoji.

Locating the Specific Video Content with the Emotions. Text was leveraged to further explain the temporality of the emotions. By reviewing the text comments, we found 61.65 % of the comments clarified that the self-report was about a section of video rather than only the individual timestamps of the self-reports using phrases such as *this topic*, *this knowledge*, *this tech*, *this segment*. For example, when a demonstration of using AR in classroom was demonstrated between 8:05 and 8:35, P27 self-reported at 8:14: 😍 *This tech looks impactful for students!* During the interview, participants noted that for self-reports of a section of video, the self-report is usually labeled with the timestamp toward the beginning of the section in case there is any chance they'll need to review the video clip. By conducting a paired t-test, we did not find any difference

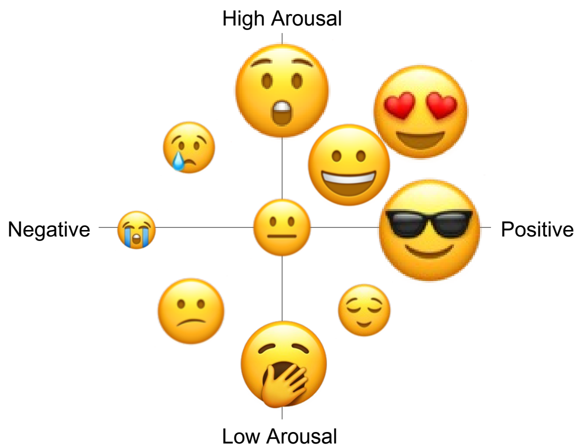

Figure 5: A bubble chart of the frequencies of each emoji being selected by participants in self-reports, with each emoji representing a specific valence (horizontal axis) and arousal (vertical axis) area [34]. Emoji sizes are proportional to the number of emojis used in participants' self-reports. Emojis associated with high arousal and positive sentiment were more commonly used than others.

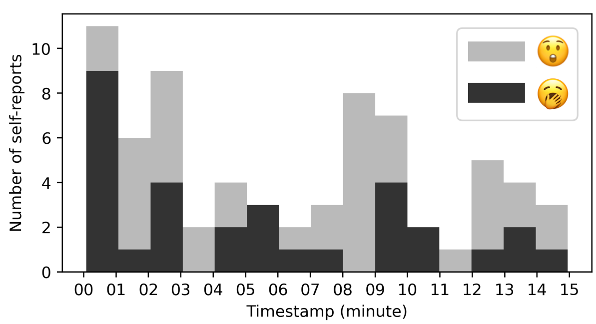

Figure 6: A histogram of self-reported high arousal emoji (😲) and low arousal emoji (🥱) at each minute of the video. Self-reports with low arousal emojis were more frequently made in the first half of the video. Self-reports with high arousal emoji were more frequently made in the middle of the video when a demo on AR applications in classrooms was presented, which our participants strongly related to. Self-reports with low arousal emoji were more frequently made at the beginning of the video as bored and sleepy feelings were less noticeable toward the later stage of the video.

between the percentage of self-reports with comments that had temporal explanations for positive-high-arousal emoji and percentage for negative-low-arousal emoji.

5 DISCUSSIONS

We found discrepancies between AI-recognized and self-reported emotions, with learners reporting a higher proportion of positive-high-arousal emoji than AI-recognized emotions. By analyzing commonly used emoji and labeled text comments, we discovered that learners predominantly reported more positive-high-arousal emojis than negative-low-arousal emojis. We found no differences in the self-reported text comments to explain the causality and temporality between positive-high-arousal emojis and negative-low-arousal emojis. Our findings emphasize the potential of post-learning dashboards to leverage the difference between AI-recognized and self-reported emotions to support reflection on SRL.

Learner-AI for Complementary Capture of Learning Emotions Our results demonstrate a discrepancy between AI and human self-reported emotions. While self-reports can provide descriptive accounts of emotions, they may overlook certain negative-low-arousal emotions and lack providing further text explanations. To bridge this gap, AI-enabled interventions can be utilized to prompt learners to report more negative emotions. Additionally, AI can be used to remind participants to further provide text explanations, ultimately allowing for a deeper and more comprehensive insight into the learning experience. Building on [19], which found that observed emotions output from an AI tool and self-reported discrete emotions do not align, we empirically found negative-low-arousal to be rather overlooked in self-report. Meanwhile, learner-inputted text comments provide some explanations of the misalignment of self-reported emoji and AI regarding the causality and temporality of emotions in SRL. In the future, affective computing research may focus on developing methods for capturing emotions that leverage the strengths of both humans and AI to provide a more nuanced understanding of emotion. This could include utilizing human narratives to provide further insight into emotions.

First-Person View vs. AI-enable Third-Person View For Reflections Our designed interaction provides an innovative first-person self-report view, as well as an AI-enabled third-person view, in order to facilitate reflection. Through our user study, we found that the comparison of the two views facilitated reflections and awareness of one's own emotions. Results showed that reflections were supported in all three cases: 1) temporal alignment between self-reports and AI-recognized emotions; 2) no changes in AI-recognized emotions at self-report timestamps; 3) changes in AI-recognized emotions for insightful timestamps but no self-report. Building on the importance of reflecting from a third-person view (e.g., distance-reflection diary [16]), our work suggests that automatic emotion recognition could be utilized as an effective approach for third-person view reflections. Additionally, our findings indicate that imperfect automatic emotion recognition does not impede reflections, and the system could suggest both alignments and misalignments between self-report and AI-recognized emotions to support the reflection process.

Limitations and Future Work We recognize a few limitations of our study. First, there are other data that can be used to recognize emotion from learners, such as brain activity through an electroencephalograph [17, 25, 33] and heart rate through electrocardiograph [9, 26]. Second, our participants were limited to current college students, and the results of the study may be different for other populations. Therefore, more studies should be conducted with participants from diverse backgrounds. Third, we only included ten emojis for participants to choose from. Future studies should investigate how other emojis may be used to self-report emotion in SRL. Ethical concerns of apply AI-recognized affects in SRL are not covered on our papers and should be better understood in future research, e.g., learners' agency of handling the uncertainty of AI, reactivity towards outcome-based AI, over-reliance on positive AI results, and fairness of AI informed decision-making [6].

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 2119589. The authors would like to thank the anonymous reviewers for their efforts and valuable comments.

REFERENCES

[1] Roger Azevedo, Michelle Taub, Nicholas V Mudrick, Garrett C Millar, Amanda E Bradbury, and Megan J Price. 2017. Using data visualizations to foster emotion regulation during self-regulated learning with advanced learning technologies. In *Informational environments*. Springer, 225–247.

[2] Qiyu Bai, Qi Dan, Zhe Mu, and Maokun Yang. 2019. A Systematic Review of Emoji: Current Research and Future Perspectives. *Frontiers in Psychology* 10 (2019). https://doi.org/10.3389/fpsyg.2019.02221

[3] Lisa Feldman Barrett. 2004. Feelings or Words? Understanding the Content in Self-Report Ratings of Experienced Emotion. (2004), 266–281 pages. https://psycnet.apa.org/doi/10.1037/0022-3514.87.2.266

[4] Ivo Benke, Maren Schneider, Xuanhui Liu, and Alexander Maedche. 2022. TeamSpiritous-A Retrospective Emotional Competence Development System for Video-Meetings. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–28.

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

[6] Si Chen, Yixin Liu, Risheng Lu, Yuqian Zhou, Yi-Chieh Lee, and Yun Huang. 2022. " Mirror, Mirror, on the Wall"-Promoting Self-Regulated Learning using Affective States Recognition via Facial Movements. In *Designing Interactive Systems Conference*. 1300–1314.

[7] Didan Deng, Zhaokang Chen, and Bertram E. Shi. 2020. Multitask Emotion Recognition with Incomplete Labels. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. 592–599. https://doi.org/10.1109/FG47880.2020.00131

[8] Didan Deng, Zhaokang Chen, Yuqian Zhou, and Bertram Shi. 2020. Mimamo net: Integrating micro-and macro-motion for video emotion recognition. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34. 2621–2628.

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

[10] Anastasia Efklides. 2011. Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational psychologist* 46, 1 (2011), 6–25.

[11] Paul Ed Ekman and Richard J Davidson. 1994. *The nature of emotion: Fundamental questions.* Oxford University Press.

[12] 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.

[13] David García-Solórzano, Eugènia Santamaría, Jose A Moran, Germán Cobo, and Carlos Monzo. 2018. Personal informatics systems for supporting self-regulation in online learning environments. *Computer Applications in Engineering Education* 26, 5 (2018), 1445–1456.

[14] Wad Ghaban. 2022. Do We Speak the Same Language? The Effect of Emojis on Learners in an Online Learning Environment. In *HCI in Games*, Xiaowen Fang (Ed.). Springer International Publishing, Cham, 414–426.

[15] Reza GhasemAghaei, Ali Arya, and Robert Biddle. 2016. A dashboard for affective E-learning: Data visualization for monitoring online learner emotions. In *EdMedia+ Innovate Learning*. Association for the Advancement of Computing in Education (AACE), 1536–1543.

[16] Igor Grossmann, Anna Dorfman, Harrison Oakes, Henri C Santos, Kathleen D Vohs, and Abigail A Scholer. 2021. Training for wisdom: The distanced-self-reflection diary method. *Psychological Science* 32, 3 (2021), 381–394.

[17] Mariam Hassib, Max Pfeiffer, Stefan Schneegass, Michael Rohs, and Florian Alt. 2017. Emotion actuator: Embodied emotional feedback through electroencephalography and electrical muscle stimulation. In *Proceedings of the 2017 chi conference on human factors in computing systems*. 6133–6146.

[18] Lun-Kai Hsu, Wen-Sheng Tseng, Li-Wei Kang, and Yu-Chiang Frank Wang. 2013. Seeing through the expression: Bridging the gap between expression and emotion recognition. In *2013 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 1–6.

[19] Harmanpreet Kaur, Daniel McDuff, Alex C Williams, Jaime Teevan, and Shamsi T Iqbal. 2022. "I Didn't Know I Looked Angry": Characterizing Observed Emotion and Reported Affect at Work. In *CHI Conference on Human Factors in Computing Systems*. 1–18.

[20] Dongho Kim, Il-Hyun Jo, Donggil Song, Hua Zheng, Jingwei Li, Jiawen Zhu, Xing Huang, Wei Yan, and Zhen Xu. 2021. Self-regulated learning strategies and student video engagement trajectory in a video-based asynchronous online course: a Bayesian latent growth modeling approach. *Asia Pacific Education Review* (2021), 1–13.

[21] Wonjung Kim, Seungchul Lee, Seonghoon Kim, Sungbin Jo, Chungkuk Yoo, Inseok Hwang, Seungwoo Kang, and Junehwa Song. 2020. Dyadic Mirror: Everyday Second-Person Live-View for Empathetic Reflection upon Parent-Child Interaction. 4, 3, Article 86 (sep 2020), 29 pages. https://doi.org/10.1145/3411815

[22] Dimitrios Kollias, Attila Schulc, Elnar Hajiyev, and Stefanos Zafeiriou. 2020. Analysing affective behavior in the first abaw 2020 competition. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*. IEEE, 637–643.

[23] Kostiantyn Kucher, Carita Paradis, and Andreas Kerren. 2018. The state of the art in sentiment visualization. In *Computer Graphics Forum*, Vol. 37. Wiley Online Library, 71–96.

[24] Gaku Kutsuzawa, Hiroyuki Umemura, Koichiro Eto, and Yoshiyuki Kobayashi. 2022. Classification of 74 facial emoji's emotional states on the valence-arousal axes. *Scientific Reports* 12, 1 (2022), 1–10.

[25] Fannie Liu, Laura Dabbish, and Geoff Kaufman. 2017. Can biosignals be expressive? How visualizations affect impression formation from shared brain activity. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–21.

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

[27] Ali Mollahosseini, Behzad Hasani, and Mohammad H Mahoor. 2017. Affectnet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing* 10, 1 (2017), 18–31.

[28] Omid Noroozi, Héctor J Pijeira-Díaz, Marta Sobocinski, Muhterem Dindar, Sanna Järvelä, and Paul A Kirschner. 2020. Multimodal data indicators for capturing cognitive, motivational, and emotional learning processes: A systematic literature review. *Education and Information Technologies* 25 (2020), 5499–5547.

[29] Ernesto Panadero. 2017. A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology* 8 (2017), 422.

[30] James A Russell. 1980. A circumplex model of affect. *Journal of personality and social psychology* 39, 6 (1980), 1161.

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

[32] Joachim J. Schouteten and Herbert L. Meiselman. 2021. The potential of using emoji instead of words when measuring emotions associated with food. *Food Quality and Preference* 92 (2021), 104182. https://doi.org/10.1016/j.foodqual.2021.104182

[33] Nathan Semertzidis, Michaela Scary, Josh Andres, Brahmi Dwivedi, Yutika Chandrashekhar Kulwe, Fabio Zambetta, and Florian Floyd Mueller. 2020. Neo-Noumena: Augmenting Emotion Communication. In *Proceedings of the 2020 CHI conference on human factors in computing systems*. 1–13.

[34] Alexander Toet, Fabienne Heijn, Anne-Marie Brouwer, Tina Mioch, and Jan B. F. van Erp. 2020. An Immersive Self-Report Tool for the Affective Appraisal of 360° VR Videos. *Frontiers in Virtual Reality* 1 (2020). https://doi.org/10.3389/frvir.2020.552587

[35] Barry J Zimmerman. 1990. Self-regulated learning and academic achievement: An overview. *Educational psychologist* 25, 1 (1990), 3–17.