Recognising Emotions in Chatbot–Learner Dialogues in Educational Robotics

Recognising Emotions in Chatbot–Learner Dialogues in Educational Robotics

Nicolai B. Plintz
Learning, Design & Technology
University of Mannheim
Mannheim, Germany
nicolai.plintz@uni-mannheim.de

Martina Morano
Dipartimento di Ingegneria
dell'Informazione
Università Politecnica delle Marche
Ancona, Italy,
TALENT, s.r.l.,
Osimo, Italy
m.morano@pm.univpm.it

Lorenzo Cesaretti
TALENT, s.r.l.,
Osimo, Italy
lorenzo.cesaretti@weturtle.org

David Scaradozzi
Dipartimento di Ingegneria
dell'Informazione
Università Politecnica delle Marche
Ancona, Italy,
d.scaradozzi@univpm.it

Dirk Ifenthaler
Learning, Design & Technology
University of Mannheim &
Curtin University
Mannheim, Germany & Perth,
Australia
dirk@ifenthaler.info

***Abstract*— Emotions significantly influence learning processes, yet their detection and classification in authentic educational chatbot interactions remain underexplored, particularly for non-English contexts. This study investigates whether a BERT-based model using few-shot learning can reliably classify emotions according to Plutchik's eight-emotion framework in Italian educational robotics dialogues. We employed the Italian language model umberto-commoncrawl-cased-v1 with 120 manually created and expert-validated few-shot examples (15 per emotion category). Following proof-of-concept validation on MultiEmotions-IT (Top-1 accuracy: 54.6%, F1: 53.8%), we analyzed 958 authentic student-chatbot-tutor interactions from M5Stack/robotics programming contexts. Results revealed distinct role-based emotion patterns: student messages were dominated by fear (69.3%), indicating technical uncertainty and learning struggles, while tutor responses exhibited anticipation (65.3%), reflecting guidance-oriented scaffolding. Despite achieving 85.4% Top-3 accuracy on benchmarks, consistently low confidence scores (10-19%) in educational dialogues revealed a fundamental domain adaptation gap between social media-trained models and formal educational discourse. The findings align with Pekrun's Control-Value Theory and suggest practical implications for designing emotion-adaptive chatbots that respond to learner affective states. However, operational deployment remains constrained by confidence limitations, highlighting the need for domain-specific training data and multimodal emotion detection approaches.**

Keywords— ***educational chatbots, emotion recognition, educational robotics, few-shot learning, BERT, Italian NLP, affective computing***

I. Introduction

Educational Robotics (ER) typically involves hands-on activities through which students engage with broader STEM domains and develop robotics-related competencies, while simultaneously nurturing transversal skills such as problem-solving, critical thinking, and creativity [1],[2],[3],[4]. To meet these objectives, students engage in hands-on activities with programmable robots and are required to program them to perform specific tasks. During these activities, educators act as facilitators, providing minimal guidance and feedback while encouraging autonomous exploration. Within this framework, the chatbot in this study supports students during ER activities as an external, on-demand resource integrated with the programming environment. Its non-intrusive design enables autonomous work while acting as a collaborative peer, fostering self-regulation, reflection, and problem-solving [5]. The effectiveness of such chatbot-mediated support depends critically on understanding learners' emotional states during interactions. Artificial intelligence (AI) applications have become promising tools to support STEM learning. Among these, chatbots -especially LLM-based systems- extend beyond factual knowledge delivery by offering flexible, context-sensitive information and, in principle, responding to learners' affective states [6]. As emotions play a central role in learning [7], understanding their role in human-AI interaction is crucial. Emotions can be seen as constructed experiences that arise when the brain makes sense of sensory input from the body and world, guided by learned emotion concepts [8]. Different theoretical frameworks exist to categorize emotions. Dimensional models describe emotions along axes such as valence and arousal, while discrete emotion models propose distinct basic emotions. Plutchik's [9],[10] Wheel of Emotions, for instance, identifies eight basic emotions (joy, trust, fear, surprise, sadness, disgust, anger, anticipation) that can be represented in varying intensities and combinations. Such categorical frameworks are particularly useful for analyzing and designing human-computer interactions, as they allow for systematic identification of emotional responses during interaction. While contemporary constructivist views [8] challenge basic emotion theories, Plutchik's categorical framework remains practically useful for systematic HCI analysis and enables interpretable emotion-adaptive system design.

Academic emotions are emotions experienced by students in relation to their academic learning, classroom instruction, and achievement [11],[12]. According to Control-Value Theory [7], these emotions arise from learners' appraisals of control over and value of achievement activities. In the context of educational chatbot interactions, however, emotions arise not only from achievement-related activities but also from the chatbot's responses, user behavior, and the overall experience of the human-chatbot interaction [13]. Research indicates that interactions between learners and educational chatbot evoke a spectrum of positive, negative, and ambivalent emotions, which in turn influence learning motivation, interest, and perceived competence [14]. Negative emotions such as frustration are particularly common in human-chatbot interactions [15], though evidence also suggests that well-designed chatbots with embodied agents such as robots can enhance engagement and motivation [16]. Understanding these induced emotions is essential for designing chatbots that can adapt to learners' affective states and provide appropriate support [14].

The developed infrastructure integrates an ER kit and an AI-enhanced conversational agent designed to assist students in solving programming-related problems when dealing with ER activities (Figure 1) [5]. This is accomplished through a three-part system comprising a human-robotics kit layer, the dataset and backend layer, and the AskLEA interface (developed by TALENT srl, https://asklea.ai/functions-ext) layer. Given this collaborative role, understanding the emotional landscape of student-chatbot-tutor interactions is crucial for optimizing the chatbot's support strategies and informing the design of future affect-sensitive educational systems. Despite the recognized importance of emotions in learning, reliably detecting and classifying emotions in authentic educational chatbot dialogues remains a significant challenge [17]. Most emotion recognition models in online learning situations are trained on general-purpose datasets, not dedicated on academic emotions [18]. This gap is particularly pronounced for non-English educational contexts, where annotated corpora of authentic student-chatbot dialogues are scarce. Furthermore, the role-specific nature of emotions in educational interactions, where students, tutors, and systems may exhibit distinct affective patterns, remains largely unexplored in literature.

Traditional supervised learning approaches require large, labelled datasets [19], which are costly and time-intensive to produce for domain- and language-specific applications. Few-shot learning methods offer a potential solution by adapting pre-trained language models to specialized domains with minimal training examples (n=15 per class), allowing proof-of-concept validation before investing in large-scale annotation efforts [e.g. 20]). However, their effectiveness for emotion classification in authentic educational dialogues remains unexplored.. The present study addresses this gap by examining the following research question:

Can a BERT-based model using few-shot learning reliably classify emotions according to Plutchik's framework in authentic Italian educational robotics dialogues?

To address this question, we first validate our few-shot approach on the established MultiEmotions-IT [21] benchmark before applying it to 958 authentic student-chatbot-tutor interactions. By analyzing these dialogues, this study provides empirical evidence on both the potential and fundamental limitations of current emotion detection approaches for affect-sensitive educational system design.

II. METHODOLOGY

A. Infrastructure and Data Collection

The infrastructure integrates an ER kit and an AI-enhanced conversational agent to assist students in solving programming-related problems during ER activities (Figure 1). The system consists of three layers [5]: (i) a human-robotics kit layer, (ii) a dataset and backend layer, and (iii) the AskLEA interface layer. The human layer (in blue, Figure 1) encompasses interactions among humans (i.e., students and teachers) as well as interactions between students and the ER kit. The ER kit is based on M5Stack products (https://m5stack.com/) and, during the experimental sessions, included an M5StickC Plus based on the ESP32 microcontroller, an ultrasonic sensor, and a pair of motors. Students were also required to program the robots. The programming process relied on the UIFlow IDE, an online platform that facilitates coding through a block-based interface. The backend layer (in orange, Figure 1) includes the database in which both the programming sequences and the AI interactions were stored for subsequent analysis, before being forwarded to the AskLEA backend. This backend ensures that the data are properly processed and prepared to support the LLM-based generation of responses when students interact with the system. Finally, the AskLEA interface (in grey, Figure 1) represents the primary point of interaction between students and the system. Through this interface, students were able to request support via the AI-enhanced conversational chatbot. Guidance and assistance were provided through tailored, on-demand responses, activated exclusively upon students' requests and without interrupting or guiding the programming process proactively. This design choice fostered students' ability to solve problems autonomously, while ensuring that the AI support remained non-intrusive.

B. Emotion Recognition Approach

For emotion recognition in chatbot interactions, we employed a few-shot learning approach using a specialized Italian BERT-based model (Musixmatch/umberto-commoncrawl-cased-v1) with Plutchik's eight basic emotions (joy, trust, fear, surprise, sadness, disgust, anger, anticipation). To establish the reliability of this approach, we first validated our few-shot model on the established Italian emotion benchmark MultiEmotions-IT before applying it to authentic educational dialogues. The few-shot model was subsequently evaluated on three different datasets: (i) 202 conversations consisting of student messages, (ii) 479 conversations containing mixed interactions, and (iii) 277 conversations with tutor responses.

The 120 few-shot examples (15 per emotion category) were generated using a chatbot primed with examples from the benchmark datasets, then cross-validated by an Italian-speaking education domain expert. The educational corpus comprises 958 authentic conversations collected during M5Stack/robotics programming activities, categorized into three datasets: (i) student messages contain only learner inputs to the chatbot, (ii) mixed conversations include the full dialogue exchanges between students, chatbot, and tutors, and (iii) tutor responses consist of human tutor replies within the dialogues.

This study adopts an exploratory approach, as ground truth emotion labels for the educational dialogues are unavailable. The reported emotion distributions reflect model predictions validated against benchmark performance, rather than human-annotated gold standards. This design allows for initial pattern detection while acknowledging that operational deployment requires future validation with expert-annotated educational corpora.

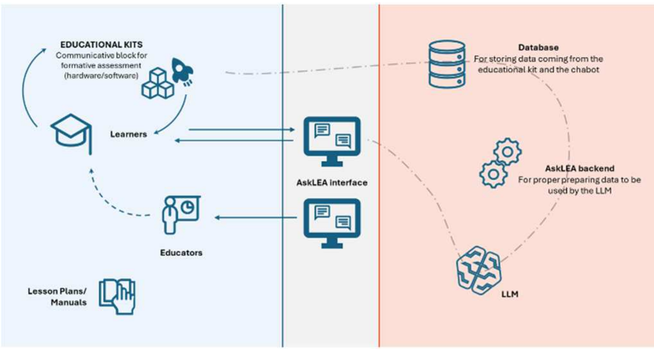

Fig. 1 - The three-layer infrastructure ER kit (in blue), AskLEA interface (in grey), and backend–database (in orange) employed during student activities.

The preprocessing pipeline included an extended token length (256) and an optimized implementation based on PyTorch and HuggingFace Transformers. The model was evaluated using F1-score, precision, recall, and top-k metrics (k = 1–3). On MultiEmotions-IT, the model achieved moderate classification performance with a Top-1 accuracy of 54.6% and an F1-score of 53.8%. High recognition rates were observed for surprise and anticipation, while trust and disgust were more challenging. Nevertheless, the top-k analysis revealed that in over 85% of cases the correct emotion was among the top three predictions.

III. Results

On MultiEmotions-IT, the few-shot model achieved moderate classification performance with a Top-1 accuracy of 54.6%, Top-2 accuracy of 76.3%, and Top-3 accuracy of 85.4% (F1: 53.8%, Precision: 57.7%, Recall: 54.6%). Class-specific performance varied substantially: surprise (F1: 0.90) and anticipation (F1: 0.89) showed high recognition rates, while trust (F1: 0.27) and disgust (F1: 0.20) were notably more challenging. Top-3 error analysis revealed that trust (26.7%) and fear (23.3%) had the highest error rates, while disgust (6.7%) showed the lowest.

The validated model was applied to three datasets comprising 958 authentic Italian educational conversations in M5Stack/robotics contexts: 202 student messages, 479 mixed conversations, and 277 tutor responses. Distinct role-based emotion patterns emerged: Student messages were dominated by fear (69.3%), indicating technical uncertainty and learning struggles, with notably low joy (10.9%). This pattern manifests in expressions of technical helplessness such as "mi si e bloccato il led" (User 007), persistent uncertainty despite scaffolding ("perché non si muove", User 002), and skepticism toward AI guidance ("sei sicura di questa risposta?", User 003). Mixed conversations showed predominantly anticipation (92.1%), suggesting collaborative problem-solving discourse. Tutor responses exhibited anticipation (65.3%) and fear (21.3%), reflecting guidance-oriented communication addressing student concerns. A representative tutor message ("Se Wally non si connette, facciamo un controllo...") was classified as anticipation with confidence 0.14, demonstrating the forward-looking guidance typical of tutoring interactions. A more detailed overview is provided in Table I.

Across all educational conversations, confidence scores remained consistently low, indicating a domain adaptation gap between general emotion datasets and specialized educational discourse.

TABLE I. Excerpt from Student Inquiries

User ID	Message	Predicted	Conf.	Context
001	"mi si è bloccato il led"	Fear	0.12	Technical blockage
002	"perché non si muove"	Fear	0.15	Robot malfunction
003	"sei sicura di questa risposta?"	Anger	0.18	Bot skepticism
004	"ciao! perché il nostro robot va storto?"	Fear	0.13	Polite inquiry
005	"signora lea noi non sappiamo far tenere acceso..."	Fear	0.16	Extended problem
006	"perché ci lampeggia una volta?"	Fear	0.14	Repeated issue

This suggests that synthetic training examples may not fully capture the complexity of authentic chatbot interactions, where language tends to be more formal, task-oriented, and less overtly affective than benchmark datasets. Despite low confidence values, the emotional patterns align with constructivist learning theory and Pekrun's control-value theory: learners express uncertainty while tutors provide scaffolding through forward-looking guidance.

IV. Discussion and Outlook

A. *Discussion of the Emotion Patterns*

The dominance of fear in student messages reflects authentic learning struggles rather than chatbot failure, aligning with existing literature on educational frustration. Frustration can be triggered when chatbots fail to provide adequate support [22], and misinformation could potentially induce frustration [23]. The present analysis appears accurate and consistent with these findings. Students question the chatbot's reliability, as exemplified by User 005's message "sei sicura?" (are you sure?). Simultaneously, such interrogative patterns may indicate a degree of critical thinking. This observation aligns with current research suggesting that cautious AI usage can enhance critical thinking capabilities [24].

From a theoretical perspective, the prevalence of fear aligns with Pekrun's Control-Value Theory [7], [11], where students expressing fear demonstrate low perceived control over the learning situation. The identification of such emotional states could enable adaptive chatbot responses, where detected fear might trigger more detailed explanations or additional scaffolding.

However, the low confidence values, ranging between 10% and 19% across all classifications, warrant critical consideration. One explanation may be that both benchmark datasets were constructed from social media posts rather than educational chatbot interactions, despite the inclusion of such examples in training. The low confidence reveals a benchmark-to-application gap, suggesting that future work should prioritize the creation of domain-specific datasets. It appears that the 120 few-shot examples were not fully sufficient for the repurposed benchmark datasets. Educational discourse differs fundamentally from social media language in its formal and technical characteristics. Moreover, the synthetic few-shot examples may miss the complexity of authentic chatbot interactions, contributing to the observed confidence limitations.

B. Practical Implications

Given current confidence limitations (10-19%), emotion detection should inform human tutors rather than trigger automated responses. Nevertheless, several design considerations emerge from the observed patterns. First, emotion-adaptive response systems should be implemented, where detected fear could trigger more detailed scaffolding, additional explanations, or simplified instructions. The prevalence of student uncertainty (69.3% fear) indicates that current chatbot responses may not adequately address cognitive challenges. Second, proactive assessment of user needs could help identify struggling learners early. Rather than waiting for explicit help requests, chatbots might periodically query students about their understanding or emotional state. Third, the low joy levels (10.9%) and instances of skepticism ("sei sicura?") suggest a need for trust-building mechanisms. Chatbots should emphasize response accuracy, provide transparent explanations of their reasoning, and acknowledge their limitations when uncertain. This could potentially increase user confidence in AI-generated guidance. Finally, motivational design elements, such as encouragement, progress acknowledgment, or gamification, might counterbalance the predominantly fear-driven interactions and foster more positive learning experiences.

C. Technical Implications & Model Performance

The moderate Top-1 accuracy of 54.6% on MultiEmotions-IT, while increasing to 85.4% in Top-3 predictions, reveals important insights about few-shot learning in specialized domains. The substantial gap between Top-1 and Top-3 performance suggests that the model captures relevant emotional features but struggles with fine-grained classification boundaries, particularly for conceptually similar emotions. This pattern indicates that confidence thresholds could be strategically employed: rather than relying solely on the highest-probability prediction, systems might flag ambiguous cases (confidence <0.20) for human review or request user clarification. The class-specific performance disparities, with surprise (F1: 0.90) and anticipation (F1: 0.89) substantially outperforming trust (F1: 0.27) and disgust (F1: 0.20), reflect both the inherent difficulty of detecting subtle emotions in text and potential class imbalance in the training data. The consistently low confidence scores (10-19%) across the educational dataset, despite acceptable accuracy on benchmarks, underscore a fundamental domain adaptation challenge. This suggests that few-shot learning with 120 synthetic examples, while sufficient for proof-of-concept validation, cannot fully bridge the linguistic and contextual gap between social media discourse and formal educational interactions. Future implementations should consider domain-specific pre-training or active learning approaches to incrementally improve confidence in specialized contexts.

D. Limitation

Several limitations constrain the generalizability and operational deployment of the present approach. First, the dataset comprises only 958 conversations from a single educational context (M5Stack/robotics programming in Italian), limiting transferability to other subjects, languages, or educational settings. Second, the absence of ground truth emotion labels for the educational dialogues prevents direct validation of classification accuracy in the target domain, the emotion distributions reflect model predictions rather than verified human annotations. Third, the reliance on 120 synthetic training examples, while validated by domain experts, may not capture the full complexity and variability of authentic student-tutor interactions. Fourth, the temporal scope of data collection may not represent seasonal variations in student engagement or emotional patterns across longer learning trajectories. Fifth, the consistently low confidence scores (10-19%) fundamentally limit operational deployment without human oversight, as automated emotion-triggered interventions require higher certainty thresholds to avoid inappropriate responses. Finally, the study examines only text-based interactions, ignoring potentially informative behavioral signals such as response latency, typing patterns, or session duration that could enhance emotion detection in multimodal frameworks.

V. CONCLUSION

This study investigated whether few-shot learning with a BERT-based model can reliably classify emotions in authentic Italian educational robotics dialogues, revealing both the promise and current limitations of this approach. We demonstrated that few-shot learning achieves acceptable benchmark performance with a Top-3 accuracy of 85.4% and successfully identifies theoretically meaningful emotion patterns in educational contexts. The dominance of fear in student messages at 69.3% and anticipation in tutor responses at 65.3% aligns with Control-Value Theory and reflects authentic learning dynamics, where students express uncertainty while tutors provide forward-looking scaffolding. However, the consistently low confidence scores ranging from 10% to 19% across all educational classifications reveal a fundamental domain adaptation gap. Models trained on social media corpora cannot be directly transferred to formal educational discourse without substantial domain-specific adaptation, which has significant implications for practical deployment. Current emotion detection systems are unsuitable for autonomous operation in educational chatbots but may serve as decision-support tools for human tutors who can interpret flagged emotional states within their pedagogical expertise. Advancing emotion-adaptive educational chatbots will require curated, expert-annotated corpora of authentic student-chatbot interactions, multimodal approaches integrating behavioral signals such as response latency and typing patterns, and active learning frameworks that incrementally improve from tutor feedback. Only through such domain-specific investments can emotion recognition move from proof-of-concept to operational deployment in educational AI systems.

REFERENCES

[1] D. Scaradozzi, L. Screpanti, and L. Cesaretti, “Towards a definition of educational robotics: a classification of tools, experiences and assessments,” Smart learning with educational robotics: Using robots to scaffold learning outcomes, pp. 63–92, 2019

[2] D. Alimisis, "Educational robotics: Open questions and new challenges," Themes in Science and Technology Education, vol. 6, no. 1, pp. 63–71, 2013.

[3] E. Mangina, G. Psyrra, L. Screpanti, and D. Scaradozzi, “Robotics in the context of primary and preschool education: A scoping review,”

IEEE Transactions on Learning Technologies, vol. 17, pp. 342–363, 2023

[4] D. P. Miller and I. Nourbakhsh, "Robotics for education," in Springer Handbook of Robotics, B. Siciliano and O. Khatib, Eds. Cham, Switzerland: Springer International Publishing, 2016, pp. 2115–2134.

[5] M. Morano, L. Screpanti, B. Castagna, L. Cesaretti, and D. Scaradozzi, "AI-enhanced Chatbot for Student Support in Robotics and Control Education," in *Proc. 33rd Mediterranean Conf. Control and Automation (MED)*, Jun. 2025, pp. 393–398, doi: 10.1109/MED64031.2025.11073475.

[6] G. Bilquise, S. Ibrahim, and K. Shaalan, "Emotionally intelligent chatbots: A systematic literature review," Human Behavior and Emerging Technologies, vol. 2022, no. 1, p. 9601630, 2022, doi: 10.1155/2022/9601630.

[7] R. Pekrun, "The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice," Educational Psychology Review, vol. 18, no. 4, pp. 315–341, 2006, doi: 10.1007/s10648-006-9029-9.

[8] L. F. Barrett, "The theory of constructed emotion: an active inference account of interoception and categorization," Social Cognitive and Affective Neuroscience, vol. 12, no. 1, pp. 1–23, 2017, doi: 10.1093/scan/nsw154.

[9] R. Plutchik, Emotion: A Psychoevolutionary Synthesis. New York, NY, USA: Harper & Row, 1980.

[10] R. Plutchik, "The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," American Scientist, vol. 89, no. 4, pp. 344–350, 2001. [Online]. Available: https://www.jstor.org/stable/27857503.

[11] R. Pekrun and R. P. Perry, "Control-value theory of achievement emotions," in International Handbook of Emotions in Education. New York, NY, USA: Routledge, 2014, pp. 120–141.

[12] R. Pekrun, A. Cusack, K. Murayama, A. J. Elliot, and K. Thomas, "The power of anticipated feedback: Effects on students' achievement goals and achievement emotions," Learning and Instruction, vol. 29, pp. 115–124, 2014, doi: 10.1016/j.learninstruc.2013.09.002.

[13] S. V. Jin and S. Youn, "Social presence and imagery processing as predictors of chatbot continuance intention in human-AI-interaction," Int. J. Human–Computer Interaction, vol. 39, no. 9, pp. 1874–1886, 2023, doi: 10.1080/10447318.2022.2129277.

[14] J. Yin, T. T. Goh, and Y. Hu, "Interactions with educational chatbots: the impact of induced emotions and students' learning motivation," Int. J. Educational Technology in Higher Education, vol. 21, no. 1, p. 47, 2024, doi: 10.1186/s41239-024-00480-3.

[15] S. Chhabra, V. Kaushal, and S. Girija, "Determining the causes of user frustration in the case of conversational chatbots," Behaviour & Information Technology, vol. 44, no. 8, pp. 1557–1575, 2024, doi: 10.1080/0144929X.2024.2362956.

[16] E. Voultsiou, E. Vrochidou, L. Moussiades, and G. A. Papakostas, "The potential of large language models for social robots in special education," Progress in Artificial Intelligence, pp. 1–25, 2025, doi: 10.1007/s13748-025-00363-2.

[17] S. Yu, A. Androsov, H. Yan, and Y. Chen, "Bridging computer and education sciences: A systematic review of automated emotion recognition in online learning environments," Computers & Education, vol. 220, p. 105111, 2024, doi: 10.1016/j.compedu.2024.105111.

[18] Plintz, N., & Ifenthaler, D. (2023). Leveraging Emotions to Enhance Learning Success in Online Education: A Systematic Review. International Association for Development of the Information Society.

[19] E. M. Younis, S. Mohsen, E. H. Houssein, and O. A. S. Ibrahim, "Machine learning for human emotion recognition: a comprehensive review," Neural Computing and Applications, vol. 36, no. 16, pp. 8901–8947, 2024, doi: 10.1007/s00521-024-09426-2.

[20] Y. Wang, Q. Yao, J. T. Kwok, and L. M. Ni, "Generalizing from a few examples: A survey on few-shot learning," ACM Computing Surveys, vol. 53, no. 3, pp. 1–34, 2020.

[21] R. Sprugnoli, "MultiEmotions-IT: A new dataset for opinion polarity and emotion analysis for Italian," in Proc. 7th Italian Conf. Computational Linguistics (CLiC-it 2020), Bologna, Italy, Mar. 2021, pp. 402–408. [Online]. Available: https://books.openedition.org/aaccademia/8910.

[22] M. Mimoun, I. Poncin, and M. Garnier, "Case study–embodied virtual agents: An analysis on reasons for failure," J. Retailing and Consumer Services, vol. 19, no. 6, pp. 605–612, 2012, doi: 10.1016/j.jretconser.2012.07.006.

[23] A. Rossner, M. Ambach, and S. Pagel, "Navigating Misinformation: Understanding User Frustration in AI-Driven Chatbot Interactions," in Artificial Intelligence in HCI. HCII 2025. Lecture Notes in Computer Science, vol. 15820, H. Degen and S. Ntoa, Eds. Cham, Switzerland: Springer, 2025, doi: 10.1007/978-3-031-93415-5_16.

[24] R. A. Fabio, A. Plebe, and R. Suriano, "AI-based chatbot interactions and critical thinking skills: an exploratory study," Current Psychology, vol. 44, pp. 8082–8095, 2025, doi: 10.1007/s12144-024-06795-8.