Journal of Inclusive Methodology and Technology in Learning and Teaching
ISSN 2785-5104
Anno6 n.3 (2026)

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

From Natural Language to Reflective Practice in AI-Supported Conflict Management Training

Elena Dell'Aquila [1], Davide Marocco [2] , Raffaele Di Fuccio [1] , Pierpaolo Limone and Umberto Barbieri [2]

[1] Pegaso University, Naples, Italy; elena.dellaquila@unipegaso.it; raffaele.difuccio@unipegaso.it; pierpaolo.limone@unipegaso.it; umberto.barbieri@unipegaso.it
[2] University of Naples Federico II, Naples, Italy; davide.marocco@unina.it
* Correspondence: elena.dellaquila@unipegaso.it

Abstract: The use of language models in teacher training raises questions about how situated responses become data and how their interpretation can support learning. This paper presents a conceptual and methodological analysis of a proposed extension of ACCORD, a single-player serious game for training teachers to manage intercultural conflicts in school settings. The existing game combines predefined responses based on Rahim's conflict-management framework with tutoring and debriefing focused on situational appropriateness. A future integration of a small language model (SLM) could allow teachers to formulate their own responses, creating greater expressive freedom while requiring careful interpretation of their language. The paper examines the conditions under which these interpretations could support self-reflection and self-regulation. Ongoing work compares SLM classifications with expert judgements to guide refinement before a possible integration into the game. The analysis distinguishes classification consistency, situational appropriateness and formative value, addressing ambiguity, possible biases and transparency in the use of language data. Teachers' opportunities to examine and question feedback are central to the proposal, whose contribution to reflective engagement remains to be evaluated.

Keywords: Educational datafication; Interpretability; Small language models; Conflict management; Teacher training

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Learning to manage classroom conflict involves examining how a response takes shape within a particular interaction and what it may mean for those involved. For teachers, this includes considering how their intentions relate to the ways students may understand their words and actions. Reflection on these exchanges connects professional learning to the capacity to evaluate one's own behaviour and adjust subsequent choices, which Bandura (2001) identifies as central to human agency. Training environments can provide opportunities to explore this

relationship, although their educational value depends on how the experience supports the examination of one's decisions.

ACCORD was developed to support teachers in managing intercultural and interethnic conflicts through a single-player virtual role-play (Dell'Aquila et al., 2022). Teachers interact with a computer-controlled character by selecting responses designed to represent different conflict-management styles. This provides a structured setting for exploring alternative approaches, within the range of expressions offered by the designers.

Allowing teachers to formulate their own responses could make the exchange more flexible and closer to how they would express themselves in a classroom encounter. The conflict-management style would then need to be inferred from their language. Work is currently underway to compare classifications produced by a small language model (SLM) with those assigned by human experts to the same teacher responses. This comparison is intended to guide refinement before a possible integration into the game.

The present paper examines the psychological and ethical implications of turning teachers' responses into conflict-management categories within a proposed extension of ACCORD. It considers what these categories can represent, where interpretation may be uncertain and how feedback can help teachers examine the meaning assigned to their words. Particular attention is given to possible biases in classification and to teachers' opportunities to question an interpretation. The paper also distinguishes the classification of conflict-management styles from evidence of engagement in reflective learning. Through this analysis, it explores the conditions under which the use of language data could support professional reflection while keeping teachers actively involved in interpreting their own experience.

2. Psychological Foundations and Their Operationalisation in ACCORD

Conflict management in ACCORD is approached as a situated process in which the meaning of a response depends on the relationship and the circumstances of the exchange. In a classroom, a teacher may need to maintain a rule while recognising a student's concerns, or decide whether an issue requires immediate intervention or further discussion. These considerations make the purposes of a response important to its interpretation: similar words may serve different functions, depending on what has happened and what the teacher is trying to achieve.

Rahim's (2002) model provides a framework for examining these orientations through two dimensions: concern for self and concern for others. Their combination defines five conflict-management styles. Integrating involves working towards a solution that addresses both parties' concerns, whereas obliging gives priority to the other party's concerns and dominating prioritises one's own. Avoiding reflects limited engagement with either set of concerns, while compromising involves reciprocal concessions. The appropriateness of a style depends on the demands of the situation. For example, collaborative exploration may be valuable when there is time to understand the disagreement, whereas an immediate decision may require a more directive response. In applying this framework to teacher training, the categories can help participants examine the strategy expressed in an exchange without treating it as a fixed characteristic of the person.

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

Translating these distinctions into the game required the construction of observable communicative behaviours. The design combined Rahim's framework with models of communication and emotional expression, through which psychological dimensions were represented in the characters' actions (Dell'Aquila et al., 2022). Each response option was developed to express an intended conflict-management style. Its verbal content was coordinated with cues conveying how the message was delivered, including punctuation, gestures, eye movements and facial expressions. The shapes of the characters' speech bubbles provided an additional means of suggesting vocal tone and emphasis. These features formed part of the operationalisation through which the player encountered the psychological model.

This arrangement gives the interaction a degree of theoretical control. The designers know which style an alternative is intended to represent, and the teacher's selection can be related to that predefined interpretation. At the same time, choosing an available response does not necessarily mean that the teacher endorses every aspect of its wording or accompanying expression. A participant may select the closest available option while preferring to phrase it differently. The choice therefore needs to be understood within the possibilities offered by the interface.

With freely written responses, a teacher might combine elements of several styles or leave their position partly implicit. Applying Rahim's framework would therefore require attention to the preceding exchange and to the limits of assigning a single category. The challenge is to preserve the link with the psychological model while accommodating formulations that were not anticipated during design.

3. Self-Reflection and Self-Regulation in the Game Experience

The educational significance of a communicative choice depends partly on what the teacher can learn by returning to it. Recognising a conflict-management style provides a conceptual reference, but reflection also involves examining why a response seemed appropriate at that moment and whether it expressed the intended purpose. Bandura's (2001) account of agency places this capacity to examine one's conduct alongside the capacity to guide subsequent action. Within the framework proposed here, psychological elaboration refers to the work through which teachers relate an exchange to their intentions and develop a more considered understanding of how they might respond.

The existing ACCORD game offers a basis for this process through opportunities to review the interaction and reconsider the choices made during a scenario (Dell'Aquila et al., 2022). Revisiting an exchange allows the teacher to consider aspects that may have received little attention while responding. A decision initially understood as necessary to maintain classroom order, for example, might also have been shaped by the interpretation of a student's behaviour as a challenge to authority. Examining that interpretation could open other ways of understanding the disagreement. The reflective task concerns the relationship between the situation as the teacher experienced it and the response they considered available.

The proposed introduction of freely formulated responses would make the teacher's own wording part of this examination. An SLM-generated category could offer an external interpretation against which the teacher considers what they intended to communicate. If the two differ, the discrepancy could prompt closer attention to how an intention was expressed, or reveal aspects of the response that the classification does not adequately capture. The teacher's judgement remains

necessary in either case. Accepting the category is not itself evidence of reflection; explaining why it seems appropriate, incomplete or unconvincing would provide a more informative account of how the participant engages with it.

This examination could also involve considering the exchange from the student's position. A teacher might recognise that an attempt to reassure could be understood as dismissing a concern, or that a request for an explanation could be experienced as an accusation. These are possible readings to explore within the scenario. Neither the bot's reaction nor the model's classification establishes how a real student would respond. Their proposed role is to make alternative interpretations available for discussion, including the emotional and relational implications of the teacher's choice.

Reflection becomes connected to self-regulation when this understanding informs a subsequent intention or action. Zimmerman's (2002) cyclical account of self-regulated learning describes how judgements about a previous attempt can shape the goals and strategies adopted for the next. Applied here, this suggests that revisiting an exchange could help a teacher identify what they would preserve in their response and what they would change. A further attempt would then provide an opportunity to examine that revision, with attention to its purpose and its consequences within the simulation.

These considerations suggest that the formative value of SLM interpretations should be examined through teachers' reasoning about their choices. How feedback and debriefing support this process is central to the proposed extension.

4. Feedback and Debriefing as Supports for Reflective Learning

In the existing ACCORD game, tutoring and debriefing connect the styles adopted by teachers to their appropriateness within the scenario. Feedback supports the review of previous choices and offers guidance on approaches that could make the interaction more effective (Dell'Aquila et al., 2022). This guidance draws on the psychological framework and considers the issue being negotiated, its importance to each person, the time available and the circumstances of those involved.

The educational focus is on understanding when and why a particular approach may be useful. Rahim (2002), for example, identifies the need for an immediate decision as a condition that may favour a dominating approach, whereas integrating requires time to explore the problem and work towards a shared solution. The meaning of a choice therefore depends on the demands of the situation. Debriefing helps teachers revisit those demands and consider the reasons for maintaining their approach or choosing an alternative.

A future extension allowing freely written responses would need to preserve this guidance while making room for teachers to examine the interpretation of their own words. Carless and Boud's (2018) account of feedback literacy offers a useful reference by emphasising learners' involvement in making judgements about feedback and using it to inform subsequent action. Applied here, this perspective suggests that teachers should have opportunities to relate the feedback to their intentions and to the features of the scenario on which the advice is based. Predesigned prompts could support this process within self-guided debriefing.

Consider a hypothetical exchange in which a student objects to being asked to speak on behalf of their cultural group. The teacher replies, "Let us finish the activity, and we can discuss this afterwards." A debriefing prompt could invite the

EDIZIONI UNIVERSITARIE ROMANE
editoria scientifica dal 1952

teacher to explain the reason for postponing the discussion. An intention to offer privacy could then be considered alongside the possibility that the student might experience the response as dismissive. The teacher could examine whether the wording communicates that intention and whether postponement fits the situation, taking into account the student's circumstances, the importance of the issue and the time available.

A further prompt could invite the teacher to explain whether they would retain or revise the response. If postponement remained appropriate, for example, a revision might make its purpose clearer while maintaining the decision to discuss the matter later. The reflective work would involve considering both how the response expresses the teacher's position and how well it addresses the situation. This would extend ACCORD's existing focus on appropriate choices to the interpretation and review of teachers' own formulations.

5. Evaluating and Refining Classification for Future SLM Integration

The ongoing work compares how the SLM and human experts classify the same teacher responses. This comparison concerns the identification of conflict-management styles and is intended to guide refinement before a possible integration into ACCORD. A further question is how these classifications could support the existing tutoring and debriefing process, where the appropriateness of a style is considered in relation to the demands of the scenario.

Expert judgements provide a useful reference, although experts may also disagree. Artstein and Poesio (2008) explain that agreement between people assigning categories helps establish the reliability of the coding process. Whether those categories adequately represent the behaviour being studied remains a separate question. A meaningful comparison requires clear definitions of the styles and attention to the information available when interpreting each response. Experts and the SLM should receive the same contextual information, while experts should initially assess responses independently, before discussing their interpretations.

Consider a hypothetical exchange in which a student objects to an activity and the teacher replies, "I understand your point, but you need to finish the activity now." If the SLM classified this response as integrating while experts classified it as dominating, the difference could suggest that the model had given too much weight to the opening expression of understanding. The experts might instead have focused on the teacher's decision to maintain the demand without inviting further discussion. This discrepancy could guide a revision of the model's instructions, clarifying that acknowledging another person's concern does not, by itself, indicate an attempt to find a solution together. Whether maintaining the demand was appropriate would require a separate judgement about the situation.

Other disagreements may arise because a response combines features of different styles or provides too little information for a clear interpretation. Refinement should therefore consider both the categories assigned and the reasons behind different readings. Alongside an overall measure of agreement, it would be useful to identify which styles are most often confused and in which situations. Changes to the classification process should then be assessed using responses that were not used to guide those changes, to examine whether the improvement also applies to new material. Where ambiguity remains, the feedback should acknowledge it.

Future integration would also require attention to how teachers' words are represented within the game. In the existing scenarios, wording, paraverbal features and non-verbal cues are coordinated during design (Dell'Aquila et al., 2022). With freely written responses, the relationship between these elements would need to be reconsidered. Text alone cannot establish the tone, gesture or facial expression a teacher intends. Evaluation would therefore need to examine whether the avatar's delivery and the bot's reaction remain coherent with the exchange.

The educational use of the classifications would require further evidence. Following Kane's (2013) approach to validation, evaluation should address both the interpretation of responses and the uses made of those interpretations. Agreement with experts would support a claim about classification consistency under the conditions examined. Future work would then need to assess whether the classifications support appropriate situational guidance and help teachers explain and reconsider their choices during debriefing. Teachers' views on expressive freedom and realism would also inform this evaluation, while effects on classroom practice would require investigation beyond the game.

6. Discussion

The proposed extension of ACCORD illustrates a central issue in educational datafication: how the meaning of an interaction changes when it is represented through categories. Its formative value would depend on how these interpretations are used and on the opportunities teachers have to examine them.

In this context, interpretability concerns both the identification of a conflict-management style and the reasons for judging it appropriate in a particular situation. ACCORD's existing tutoring and debriefing already connect these aspects through criteria drawn from the psychological framework. With freely written responses, the link between language, classification and situational guidance would need to remain clear. Teachers should be able to examine which features of their response support the classification and which features of the scenario support the advice offered. This would help them relate feedback to their intentions and understand why an alternative approach might be more suitable.

Possible biases also need to be considered when moving towards more flexible interaction. Future evaluation should examine whether responses expressing comparable relational positions are classified consistently across different ways of wording them. Differences in directness or formality, for example, would need to be considered in relation to their meaning within the exchange. The comparison with experts can help identify difficulties, although the experts' criteria also require scrutiny. Agreement may reflect a shared interpretation whose assumptions still need to be examined.

These questions concern both the meaning of the data and their use. Slade and Prinsloo (2013) identify context, transparency and the purposes of data collection as central concerns in the ethics of learning analytics. For a future extension of ACCORD, teachers should be able to understand what information is recorded, how it informs feedback and who can access it. Classifications should remain tied to responses within specific situations, with opportunities to question their interpretation. Using these categories to make broader judgements about a teacher's professional competence would require separate justification and evidence.

Maintaining a clear formative purpose is therefore part of the psychological and ethical design of the experience.

The relationship with engagement requires similar care. A conflict-management category describes an interpretation of a response; it does not establish how deeply the teacher has engaged with the experience. Within this proposal, the focus is on engagement in reflection: examining one's wording, considering another reading and explaining a decision to retain or revise a response. Teachers' reasons for accepting or questioning feedback would provide material for studying this process. Counts of interactions or changes in assigned categories would offer only a partial view. The ongoing comparison with experts addresses the interpretation of teachers' language, while future use within the game would allow examination of how that interpretation contributes to reflective learning.

7. Conclusions

The proposed extension of ACCORD connects greater freedom of expression with the challenge of interpreting teachers' responses within a psychological framework. Its development would build on existing tutoring and debriefing, where the appropriateness of a style is considered in relation to the situation.

The ongoing comparison with expert judgements can guide the refinement of SLM classifications before a possible integration into the game. Further work would need to examine how teachers use the resulting feedback. The educational use of language data should make clear how responses are used, why a particular interpretation is offered and where its limits lie. This transparency would support teachers in examining the feedback in relation to their intentions and the demands of the situation.

References

Artstein, R., & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computa-tional Linguistics*, *34*(4), 555–596. https://doi.org/10.1162/coli.07-034-R2

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, *52*, 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Barbieri, U., Di Fuccio, R., Limone, P., Marocco, D., & Dell'Aquila, E. (2026). From scripted to adaptive serious games: On-device language models for conflict training. In *Proceedings of the 18th International Conference on Computer Supported Education*, 734–741). SciTePress. https://doi.org/10.5220/0015090900004021

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Dell'Aquila, E., Vallone, F., Zurlo, M. C., & Marocco, D. (2022). SG-ACCORD: Designing virtual agents for soft skills training in the school context. *Education Sciences*, *12*(3), 174. https://doi.org/10.3390/educsci12030174

Dell'Aquila, E., Marocco, D., Di Fuccio, R., Limone, P., & Barbieri, U. (2026). Natural Teacher Lan-guage for Conflict Style Classification with SLMs. In *Proceedings of HELMeTO 2026* [In press].

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. https://doi.org/10.1111/jedm.12000

Rahim, M. A. (2002). Toward a theory of managing organizational conflict. *International journal of conflict management*, *13*(3), 206-235. https://doi.org/10.1108/eb022874

Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavior-al Scientist, 57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice, 41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2.