

# Criticize my Code, not me: Using AI-Generated Feedback in Computer Science Teaching

Sibylle Kunz and Adrienne Steffen

IU International University of Applied Sciences, Erfurt, Germany

[sibylle.kunz@iu.org](mailto:sibylle.kunz@iu.org)

[adrienne.steffen@iu.org](mailto:adrienne.steffen@iu.org)

**Abstract:** Large Language Models (LLMs) like ChatGPT can help teachers to tailor learning tasks for their students, combining learning objectives and storytelling to raise interest in the subject. AI-based learning task design can help to support competency-based learning, especially for girls in STEM courses like computer science, where otherwise the “Leaky STEM pipeline” (Speer 2023) leads to a constant loss of female students over school time. LLMs support many steps of the creation cycle of learning tasks. One important step is the feedback process between teachers and students during and after solving the tasks. Students need person-related as well as process-related feedback to make progress. Sometimes problems occur when teachers give feedback in a way that embarrasses or hurts the students. Especially female students often need more confirmation to make them aware of their progress, but studies show that boys demand and get more attention by teachers in this situation. This is one of the many reasons why girls lose motivation and interest in STEM courses over time. Since male and female teachers differ in expressing feedback without being aware of it, it is necessary to raise their consciousness. LLMs like ChatGPT can be used in two scenarios here. The first scenario is helping teachers to formulate objective feedback in a way that is adequate and understandable for the target group – e.g., young girls or boys - in a specific situation. The second scenario is training the teacher in a Socratic way, where the LLM simulates a student receiving the feedback and reacting to it according to established communication models like the Four Ears-model by Schulz von Thun (Schulz von Thun 1981) or Berne’s Transactional Analysis (Berne, 1964). This case study provides examples and prompting schemes for both scenarios and discusses the fragile balance between avoiding gender stereotypes in LLMs and giving more helpful and sustainable feedback for female students to foster self-esteem and competency-awareness.

**Keywords:** Female in STEM, Large Language Models, ChatGPT, Feedback Process, AI-based Feedback

---

## 1. Introduction

The phenomenon called “the leaky STEM pipeline” (Speer 2023) describes the fact that many girls start into STEM (Science, Technology, Engineering, Math) subjects successfully but lose interest over the years which leads to an underrepresentation of women in STEM vocational training and STEM-related courses of study at universities, especially computer science. The phenomenon of horizontal segregation has been going on for many years and many aspects of it have been studied. Amongst others, the following have been found to be the main reasons:

A lack of confidence and self-esteem (Weber 2012, Kollmayer et al. 2018) makes girls think they are not as competent and successful as boys in these subjects and are not as capable of mastering tasks. This is also caused by the lack of female role models. And it leads to another reason: a sense of inadequacy in meeting the requirements of universities (Thurner et al. 2014, Zehetmeier et al. 2014). Many girls think taking up a course of study, e.g. in computer science, is very difficult and bears the risk of failure.

A lack of interest due to unsuitable learning tasks (IU Internationale Hochschule 2022, Hess et al. 2023) is another motivational factor: very often the learning tasks, especially in computer science, deal with subjects from the real world that girls are less interested in, e.g., combat games, cars, mathematical problems from engineering etc., instead of using fields like ecology and environmental protection, medicine, or societal aspects where girls often show a higher interest.

It has been found that a general lack of knowledge about IT as an area of work also contributes to the leaky STEM pipeline (Hur et al. 2017). Many girls associate computer science mainly with coding software, not knowing about the many work activities and use cases in almost every professional field, e.g., process management in the industry or user experience design for interfaces, AI in medicine or robotics in healthcare.

In addition to this there is a general lack of female role-models in many STEM fields, especially computer science (Corbett & Hill 2015). This relates to lower visibility of female scientists or managers in IT companies in the public, and in public media as well as a lack of female role models amongst family members (e. g., mothers, grandmothers, aunts, sisters) or friends in the girls’ personal environment (IU Internationale Hochschule 2022).

To raise the female students’ interest in the field of computer science, teachers need to improve lessons and learning tasks. For example, tasks could be more appealing to girls by using storytelling approaches to frame the

setting of a task and to focus on topics that girls generally show more interest in (Kunz & Heß 2023a). So, instead of programming a classical shooter game, the task could be changed to developing a game where the player saves the environment, manages a social institution, etc.

Another important step in the process of learning is the feedback given by the teacher (Paechter et al., 2020). The way feedback is given and received depends on many factors and can be examined using different communication models. The effect the teacher's reaction has on error correction, further learning and understanding and the emotional status of the students concerning self-esteem and self-assurance can be significant. But carefully preparing deliberate feedback instead of just criticizing faults is time-consuming and is something not all teachers are adequately trained in.

Large Language Models (LLMs) are Artificial Intelligence (AI) systems that can analyze and generate text, because they were trained on large amounts of language data. LLMs are used in AI-chatbot systems like ChatGPT. They can be integrated in many phases of a learning cycle (Kunz & Hess 2023b). This article shows how the feedback phase can be improved by either optimizing textual feedback or training the teachers in simulated dialogues, where the chatbot acts in the role of a student receiving feedback and reacts according to different communication models. By doing so, difficult conversational situations with adolescent students can be explored and repeated to sensitize teachers for the special requirements of their students.

Section 2 of this article gives an overview of theoretical concepts in the field of feedback and its different manifestations as well as research findings on the quantitative and qualitative aspects of teachers' feedback towards male and female students. It also refers to a framework for designing and evaluating learning tasks and the role of critical incidents that might occur during assessments. A short overview of two communication models is given. Section 3 briefly introduces the methodology used for the case study based on ChatGPT. In section 4, example prompts and the AI-generated answers are presented. Section 5 discusses whether using this approach can lead to more helpful and effective feedback and gives an outlook on future research.

## **2. Theoretical Foundation**

### **2.1 Feedback Types and Feedback Levels**

Feedback is "information provided by an agent regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p.81). It can be defined as the answer of the receiver of a message to the sender. From this answer, the sender can evaluate if and how his message has been understood and interpreted. There are different types of feedback. Studies show that task-related feedback with hints how to work more effectively had the highest impact on learning outcomes in contrast to praise or punishment (Hattie & Timperley, 2007, p.84).

Hattie and Timperley distinguish between "feed back", "feed forward" and "feed up". "Feed back" refers to answering the questions "how am I going? what progress is being made toward the goal?" whereas "feed forward" raises the question "where do I go from here? what activities are needed to make better progress?" "Feed up" focuses on the main goal, comparing the current to the target status. Feedback, according to their framework, relates to four different levels: task level, process level, self-regulation level and self level. Feedback on the task level shows the student whether his or her answer is correct or what is missing. The process level refers to the methods used. The self-regulation level is addressed whenever the student's confidence, commitment, and self-control are triggered. The self-level focuses on the character of person receiving the feedback. It often contains too little task or process-related information to change the way the student is working and therefore is less effective (Hattie & Timperley 2007).

A similar approach by Paechter et al. (2020, p. 2) distinguishes between person-related feedback and process-related feedback. Person-related feedback aims at a student's characteristics, e.g., having a talent for math or music. Used in an appropriate way, it can foster a "fixed mindset" – or otherwise convince a person that abilities are pre-determined and cannot be altered. Process-related feedback concentrates on the learning process itself, on the activities performed and the methods used to solve a task. Therefore, it supports developing a "growth mindset" (Paechter et al., 2020, p.3) and a positive attitude towards learning.

### **2.2 The Impact of Feedback on Girls' Attitudes Towards STEM**

Paechter et al. (2020) state that there are three key variables fostering girls' decision for a STEM career: STEM self-efficacy, outcome expectations and interest in the subject. Self-efficacy means the girls believe in their competency to solve STEM-related tasks successfully. Outcome expectations link the personal behavior (e. g.,

trying very intensely to master a learning challenge) to the immediate consequence of success. Apart from personal experiences, aspects like grading, the experience of flow during the work or pride in the case of having solved an issue also influence the outcome expectations. Teacher feedback plays a big role in influencing self-efficacy and outcome expectations. It has been shown that girls tend to believe in teachers' feedback more intensely than boys, although they often get less (Skipper and Lehmann 2017). Person-related feedback towards girls in STEM-subjects carries the danger of consolidating a lack of self-efficacy in connection with a belief in gender stereotypes about women's inferiority in STEM. Many people believe technical or scientific subjects depend more on a naturally given talent than on successful learning (Hattie & Timperley 2007).

These findings were also supported by a recent study with 777 schoolgirls aged between 14 and 21 who were asked if they were interested in STEM subjects. Although 70 % stated their interest in STEM fields, 44% said that they found it too difficult and 42% felt overstrained with the topics (IU Internationale Hochschule 2022).

Based on data from this study, Steffen et al. (2023) examine the influence of critical incidents – positive or negative ones – on the female students' interest in STEM topics and their willingness to engage in a technical or scientific field during schooltime and later in university. It shows that being embarrassed or inadequately criticized by a teacher leads to shyness in class, fear of making errors, or rejection to further engage in the course, but also that experiencing pride after understanding a difficult topic and being able to explain it to others in class, can raise the girls' self-esteem.

### **2.3 Using an AI-based Framework to Create Learning Tasks and Provide Feedback**

In the framework developed by Kunz & Hess (2023a), a set of dimensions for describing learning tasks for computer science lessons like group size, gender pairing, duration, assessment, grading, etc. is given. The way learning results are being assessed and/or graded should therefore depend on the task content as well as on the intended outcome and should consider the requirements of the group constellation. For example, an app for organizing neighborhood help developed by an all-female team of students over several lessons could undergo a formative assessment by the teacher during development and a summative assessment after completion. The concept, the software architecture, the correctness or efficiency of the code or the design of the user interface could be graded separately, but for the whole team, or the contribution of each team member could be measured separately. Each variant will lead to different forms of feedback – and different ways of reception among the team and by each student herself.

As suggested in Kunz & Hess (2023b), generative AI in the form of LLMs like ChatGPT or other chatbots can be used to help in most of the learning cycle's actions – from setting a topic, creating the task description, additional material, and sample solutions to grading criteria. Integrating the feedback phase will add another element to the list of AI-supported activities.

### **2.4 Feedback as Part of a Communication Process**

Feedback is a part of a communication process. A student has worked on or finished a learning task, and a teacher provides a reaction, either person-related, task-related, process-related or as a mixture of all dimensions. This process can be described using different communication models. Two very suitable models for our case study are the Four-Ears model by Schulz von Thun (1981) and Berne's Transactional Analysis (1964).

The Four-Ears model is built on two prior concepts: one is Bühler's "Organon Model" (Bühler 1999), where a linguistic sign has an expressive function, an appealing function, and a representational function. The other is Watzlawick's second axiom of communication (Watzlawick 1967), stating that every communication has a content and a relationship aspect, and the latter determines the former. The Four-Ears model expands the concept. Here, each message has four facets: the factual information the sender wants to transmit, a self-revelation statement, an appeal to the receiver and a statement about the relationship between sender and receiver. Since the receiver of the message can "hear" it through four different "ears" corresponding to the facets, there are four ways to understand – or misunderstand - the message. Section 4.2 provides an example.

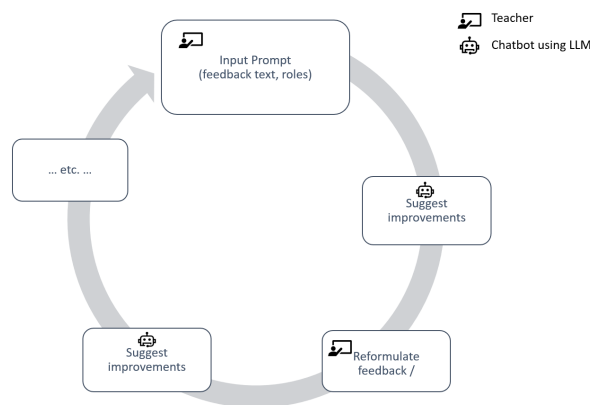
Berne's Transactional Analysis is a psychoanalytic theory that explores human communication, relationships, and personality. Berne defined interactions as transactions between individuals, each playing three specific roles that are called "ego states." These are "parent", "adult", and "child". The parent represents learned behaviors and attitudes from authority figures, the adult is the rational and objective state, and the child reflects emotions and spontaneous reactions. Interactions involve exchanges between these ego states, influencing the dynamics of relationships. Transactional Analysis identifies communication patterns as transactions, classified into

complementary (matching ego states) and crossed (mismatched ego states). These patterns impact interpersonal dynamics, often leading to misunderstandings or conflicts.

### 3. Methodology

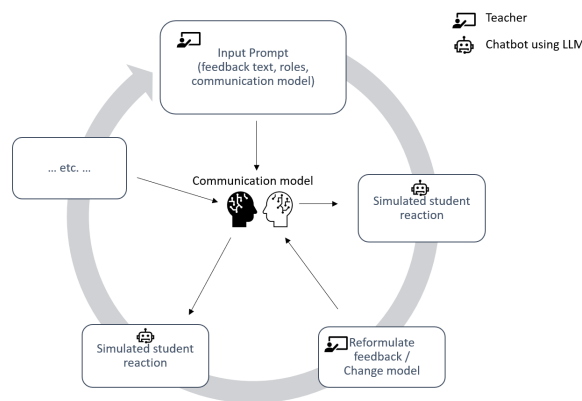
LLMs can be used in two different scenarios to improve the feedback process element: by re-formulating a given feedback text or by engaging the teacher as a user in a Socratic dialogue. Both require prompts containing the feedback text and a description of the two roles. Scenario 2 also requires naming of a communication model.

In the first scenario the objective is to help teachers create objective feedback in a way that is adequate and understandable for the target group – e.g., young girls - in a specific situation. In this case study, the feedback example is a very straight-to-the-point text a teacher might use for commenting a student’s work, so the LLM can find better and “smoother” alternatives. This needs to be done in a way that fosters objectivity and at the same time has an encouraging and motivating effect on the student, avoiding frustration or a depression of self-esteem.



**Figure 1: Training feedback dialogues using different communication models**

In the second scenario, LLM should simulate a student receiving the feedback and reacting to it according to established communication models. Using the Four- Ears model, the teacher can test what a feedback message sounds like and how it is received, adapting to the reaction in his or her next statement. Asking the LLM to act along the framework of Transactional Analysis, the teacher can see difference if a student reacts on Child, Parent or Adult level and find out how to react properly to break up negative communication patterns. Figure 2 shows such an iterative dialogue.



**Figure 2: Training feedback dialogues using different communication models**

Models that include “hidden” parts of the communication like the Iceberg Model are less suitable since – at least for now - a LLM can only answer using text or sometimes images. In the future and using Virtual Reality (VR) environments, it might become possible to integrate an avatar into the training which is able to conduct nonverbal communication or body language in real-time (Rogers et al. 2022).

#### 4. Case study: Example Prompts for Generating Appropriate Feedback to Female Students in Computer Science Tasks

This section provides some case studies using examples and prompting schemes for both scenarios. Version 3.5 of ChatGPT was used. In each case, the teacher explains his own role and the role the chatbot is expected to use, he also adds what the result should look like. Another important part of the prompt is asking for an explanation, forcing the LLM to reveal how the answer was derived.

##### 4.1 Improving Textual Feedback Quality

Table 1 shows a typical prompt and the resulting dialogue for scenario 1: The teacher asks for rephrasing a rather straight-to-the-point feedback.

**Table 1: A typical prompt for Scenario 1 – Reformulating a feedback text.**

<b>Teacher</b>	You are a 15-year-old female school student in computer science. I am a teacher giving you feedback on your homework. Analyze my feedback and reformulate it in a way the student might understand or like better. Explain your changes.
<b>ChatGPT</b>	Certainly! Please provide me with the feedback you gave to the student on their computer science homework, and I'll do my best to help you reformulate it in a way that the student might find more understandable or appealing.
<b>Teacher</b>	My feedback is: You have not properly understood the concepts of if-constructs and loops. There are a lot of syntax errors in your code and the code itself does not make sense in relation to the problem given.
<b>ChatGPT</b>	Original Feedback: You have not properly understood the concepts of if-constructs and loops. There are a lot of syntax errors in your code, and the code itself does not make sense in relation to the problem given.  Reformulated Feedback: Hey there! It looks like there might be a bit of confusion with if-constructs and loops in your code. I noticed a few syntax errors, and the code doesn't quite match the problem we're working on. Let's go over these parts together to make sure everything clicks. Keep up the effort!
<b>Teacher</b>	Explain your changes.
<b>ChatGPT</b>	I made several changes to the original feedback to make it more student-friendly: <ul style="list-style-type: none"> <li>• <b>Positive Tone:</b> I started the feedback with a positive and encouraging tone ("Hey there!"), which helps create a supportive atmosphere. This can motivate the student to approach the feedback with a positive mindset.</li> <li>• <b>Softening Language:</b> Instead of stating, "You have not properly understood," I used more gentle language like "there might be a bit of confusion." This helps avoid sounding too critical and gives the student room for improvement without feeling discouraged.</li> <li>• <b>Specific Feedback:</b> I pointed out the specific issues with if-constructs and loops without overwhelming the student. This helps them focus on particular areas that need improvement.</li> <li>• <b>Collaborative Approach:</b> Instead of just highlighting the errors, I suggested working on these parts together. This implies a collaborative effort between the teacher and the student, fostering a sense of teamwork and support.</li> <li>• <b>Encouragement:</b> I concluded with positive encouragement ("Keep up the effort!"), reinforcing the idea that improvement is a continuous process, and that the student is on the right track.</li> </ul> <p>Overall, the goal is to provide constructive feedback in a way that motivates the student to learn and improve, rather than making them feel discouraged or overwhelmed.</p>

It can be seen from the generated answer that the core of the criticism itself remains the same, but a more encouraging and less repellent formulation is used. The feedback is mostly task-related as well as process-related and not person-related. It contains feed-back (e. g., "the code does not match the problem") as well as feed-forward elements ("let's go over these parts together"). Using this approach, teachers can get sensitized for expression that might frustrate their students.

##### 4.2 Training Communication Patterns

Scenario 2: The teacher wants to explore communication patterns he or she might encounter when providing a certain form of feedback. ChatGPT is prompted to react according to some common communication behavior models. Table 2 shows an example for Schulz von Thun's Four-Ears-Model (Schulz von Thun 1981).

**Table 2: A typical prompt for Scenario 2 – Engaging in a Socratic dialogue**

<b>Teacher</b>	I am a teacher giving feedback on some homework. You are a 15-year-old female student. For each feedback message I give, analyze the message according to the Four-Ear model by Schulz von Thun and explain how it will be understood by you. Then give four different answers according to the four-ear-model and explain why you react in this way.
<b>ChatGPT</b>	Certainly, I'd be happy to help you analyze and respond to feedback using the Four-Ear model by Schulz von Thun. The model consists of four communication aspects: factual information, self-revelation, relationship, and appeal. Please provide the first feedback message, and I'll assist you in breaking it down and providing potential student responses.
<b>Teacher</b>	You have not properly understood the concepts of if-constructs and loops. There are a lot of syntax errors in your code and the code itself does not make sense in relation to the problem given.
<b>ChatGPT</b>	<p>Let's break down the feedback using the Four-Ear model:</p> <ol style="list-style-type: none"> <li>1. Factual information: The teacher is providing information about the specific issues in the student's work - misunderstanding of if-constructs and loops, syntax errors, and the lack of coherence between the code and the given problem.</li> <li>2. Self-revelation: The teacher is expressing their assessment of the student's understanding and coding skills.</li> <li>3. Relationship: The feedback could be perceived as somewhat critical, indicating areas of improvement. It may also convey the expectation that the student needs to work on these concepts.</li> <li>4. Appeal: The teacher may expect the student to review and revise the code, focusing on understanding if-constructs and loops, fixing syntax errors, and ensuring the code aligns with the given problem.</li> </ol> <p>Now, let's consider four potential responses from the student:</p> <ol style="list-style-type: none"> <li>1. Factual response: "Thank you for pointing out the specific issues. I'll go back and review if-constructs and loops, fix the syntax errors, and make sure my code aligns with the problem statement."</li> <li>2. Self-revealing response: "I thought I understood if-constructs and loops better, but I appreciate your feedback. I'll take extra time to grasp these concepts and make sure my code is more accurate."</li> <li>3. Relationship-focused response: "I'm a bit confused by your feedback. Could you please provide more guidance on where I went wrong? I want to improve and ensure my understanding aligns with the expectations."</li> <li>4. Appeal-oriented response: "I see the issues you pointed out. Could you give me some examples or additional explanations to help me understand better? I want to make sure I correct these mistakes and improve my coding skills."</li> </ol> <p>These responses demonstrate how a student might interpret and react to the feedback based on the Four-Ear model's components. Each response addresses different aspects of the feedback, ranging from acknowledging specific issues to seeking clarification and guidance.</p>

This conversation could now be continued delving deeper in each of the four aspects or asking the LLM to reformulate the feedback to make it more coherent with one of the four facets using the techniques described in Scenario 1. Table 3 shows how the dialogue was continued, switching to Berne's Transactional Theory-Model (Berne 1964).

**Table 3: The dialogue is continued using Transactional Analysis**

<b>Teacher</b>	Now use the transactional theory as a basis for your answer and explain it with regard to the level of communication and the position of teacher and student. Use the same teacher feedback from above.
<b>ChatGPT</b>	<p>Transactional analysis involves analyzing social transactions to determine the ego state of the communicator (Parent, Adult, or Child) as a basis for understanding and improving communication. In the given teacher feedback, we can analyze it through the lens of transactional analysis.</p> <ol style="list-style-type: none"> <li>1. Critical Parent to Adult (Teacher to Student): <ul style="list-style-type: none"> <li>- The teacher's feedback can be seen as coming from the Critical Parent ego state, providing a critical analysis of the student's work. The teacher, in the Adult ego state, is offering specific information about the issues in the student's work.</li> </ul> </li> <li>2. Student's Possible Ego States: <ul style="list-style-type: none"> <li>- The student may respond from different ego states: <ul style="list-style-type: none"> <li>- Adapted Child: The student might feel the need to comply and correct the errors as pointed out without questioning or expressing their feelings.</li> <li>- Natural Child: The student might react emotionally, feeling frustrated or upset about the feedback.</li> </ul> </li> </ul> </li> </ol>

	<p>- Adult: The student may analyze the feedback objectively, acknowledging the issues and planning a rational response.</p> <p>3. Possible Student Responses:</p> <ul style="list-style-type: none"> <li>- Adapted Child Response: "I'm sorry for the mistakes. I'll fix them right away."</li> <li>- Natural Child Response: "This is so confusing. I don't get it."</li> <li>- Adult Response: "Thank you for highlighting the issues. I'll review the concepts and make the necessary corrections."</li> </ul> <p>4. Improving Communication:</p> <ul style="list-style-type: none"> <li>- For more effective communication, the teacher could also choose to engage the Adult ego state in the student. Instead of a purely Critical Parent message, the teacher might include more guidance or examples to help the student understand the mistakes better, fostering a more collaborative and educational environment.</li> </ul> <p>In summary, the transactional analysis helps to understand the dynamics of communication in terms of ego states. It allows to evaluate not only the content of the message but also the emotional and psychological aspects involved, leading to more effective communication and a constructive learning environment.</p>
<b>Teacher</b>	Make a suggestion how the teacher could engage the Adult ego state in the student here.
<b>ChatGPT</b>	<p>[...] the teacher could provide more specific information and guidance in the feedback. Instead of solely pointing out errors, the teacher can offer constructive suggestions and examples to facilitate the student's understanding. Here's a revised version of the feedback that incorporates a more Adult-to-Adult communication style:</p> <p>"Thank you for submitting your work. I noticed some areas that could use clarification. For instance, in the section related to if-constructs and loops, consider reviewing the following concepts [provide specific concepts]. Additionally, there are a few syntax errors, such as [mention specific errors]. To help you grasp these concepts better, here are some examples and explanations [provide examples]. I believe addressing these points will strengthen your code and align it more closely with the given problem. If you have any questions or need further clarification, feel free to ask. I appreciate your effort in improving your coding skills."</p> <p>This approach encourages a more collaborative and supportive tone, [...]</p>

The dialogue shows how the chatbot can be used to predict possible reactions or patterns from the student's side, so the teacher can use them as training material and get advice on how to continue the feedback process.

## 5. Discussion and Future Research Options

The case study shows that even very short prompts lead to good results for both dialogue scenarios as long as some basic information about the roles of sender and receiver and some requirements for explaining the answer are given. The usage of established communication models leads to detailed explanations of possible dialogue threads. As can be seen from the examples given in section 4, the chatbot's answers do not show gender stereotypes in the answers, although the information about the students' sex was provided. In general, LLMs can be subject to gender stereotypes when it comes to storytelling (Kunz & Hess 2023b), so teachers should always examine generated texts before using them in the feedback process.

LLMs can support teachers during many steps in the creation and evaluation of learning materials. The feedback process is a very sensitive phase, especially for female students, who often misinterpret person-related or process-related feedback which is one influencing factor in the leaky STEM pipeline (Paechter et al. 2020). Teachers can be trained to be more attentive by either improving the quality of their written feedback and having the improvements explained by a chatbot or by engaging themselves in Socratic dialogues with the chatbot taking the role of the young female feedback recipient. Knowing and using different communication models like the Four-Ear-Model or Transactional Theory in this training can highlight different aspects in the resulting communication. Since graphical representations in the form of human avatars become more and more realistic, the future might as well bring the opportunity to interact with using body language or other nonverbal communication means.

Further field research should use the two scenarios in a training for computers science teachers to find out if this approach is perceived useful. In a next step it could be tested in real feedback situations. Again, two settings are possible: Teachers could provide female students with the "original" or "raw" version of their feedback and with the LLM's improved version, asking which they perceive as more helpful. The second scenario can be used to store the simulated reactions from the dialogue and later compare the students' reaction to the answer suggested by the LLM. In addition, the scenarios could be tested with ChatGPT version 4.0 to see, whether even better feedback results can be generated. In any case, LLMs are due to become a helpful tool not only in lessons design, but also in the assessment and feedback phases.

## References

- Berne, Eric (1964) *Games People Play: The Psychology of Human Relationships*, Grove press.
- Bühler, K. (1999) *Sprachtheorie: Die Darstellungsfunktion der Sprache*, 3rd Edition, Fischer.
- Corbett, C., Hill, C., (2015) *Solving the Equation. The variables for Women's Success in Engineering and Computing*. AAUW, US, Washington D.C.
- Hattie, J. and Timperley, H. (2007) "The power of feedback", *Review of Educational Research*, 77(1), pp. 81-112, Available at: <https://doi.org/10.3102/003465430298487>.
- Hess, C., Dodiya, J., Heinisch, C., Kunz, S., Rebholz, S. Schlömer, I., Steffen, A. and Vaas, S. (2023) "How does school lesson design contribute to girls' perception of their future-readiness for STEM majors – An empirical study using critical incidents technique", *The 4th International Conference on Teaching and Education Management (ICTEM 2023)*, Berlin, Germany, June 2023 (in press).
- IU Internationale Hochschule (2022) *Kurzstudie 2022. MINT-Bildung. Was junge Frauen darüber denken*. Available at: [https://static.iu.de/studies/Junge\\_Frauen\\_in\\_MINT\\_Kurzstudie.pdf](https://static.iu.de/studies/Junge_Frauen_in_MINT_Kurzstudie.pdf).
- Hur, J.-W. et al. (2017), "Girls and computer science: experiences, perceptions, and career aspirations", *Computer Science Education*, 27:2, pp. 100-120. Available at DOI: 10.1080/08993408.2017.1376385
- Kollmayer, M. et al. (2018) "Gender stereotypes in education: Development, consequences, and interventions", *European Journal of Developmental Psychology*, Vol. 15, No. 4, pp. 361–377. Available at <https://doi.org/10.1080/17405629.2016.1193483>
- Kunz, S. and Hess, C. (2023a) "How to catch'em all - Designing attractive learning activities for girls in computer science. A systematic framework", *The 4th International Conference on Teaching and Education Management (ICTEM 2023)*, Berlin, Germany, June 2023 (in press).
- Kunz, S. and Hess, C. (2023) "Teachers, Reload Your Toolboxes: Using a Task Design Framework and ChatGPT to Generate Motivational Exercises for Female Learners in Computer Science", P. Dini (ed.) *Proceedings of Advances on Societal Digital Transformation (DIGITAL 2023)*, pp. 13–18.
- Paechter, M., Luttenberger, S. and Ertl, B. (2020) "Distributing Feedback Wisely to Empower Girls in STEM", *Frontiers in Education*, Vol 5. Available at: <https://doi.org/10.3389/feeduc.2020.00141>.
- Rogers, S. L., Broadbent, R. Brown, J., Fraser, A. and Speelman, C.P. (2022) "Realistic Motion Avatars are the Future for Social Interaction in Virtual Reality", *Frontiers in Virtual Reality*, Vol. 2-2021. Available at <https://doi.org/10.3389/frvir.2021.750729>.
- Schulz von Thun, F. (1981) *Miteinander reden. Band 1: Störungen und Klärungen*. Rowohlt Taschenbuch.
- Skipper, Y. and Leman, P. J. (2017) "The role of feedback in young people's academic choices", *International Journal of Science Education*, 39:4, pp. 453–467. Available at: <https://doi.org/10.1080/09500693.2017.1294783>.
- Speer, D. (2023) "Bye Bye Ms. American Sci: Women and the Leaky STEM Pipeline", *Economics of Education Review* 93:102371.
- Steffen, A. et al. (2023) "An Exploration of Critical Incidents Impacting Female Students' Attitude Towards STEM Subjects", *International Conference on Gender Research*, 6(1), pp. 215–223. Available at: <https://doi.org/10.34190/icgr.6.1.994>.
- Thurner, V. et al. (2014) "Identifying base competencies as prerequisites for software engineering education", *EDUCON: 2014 IEEE Global Engineering Education Conference, Institute of Electrical and Electronics Engineers*, pp. 1069–1076. Available at <https://doi.org/10.1109/EDUCON.2014.6826240>.
- Watzlawick, P., Beavin, J. H. and Jackson, D.D. (1967) *Pragmatics of Human Communication*, New York.
- Weber, K. (2012), "Gender Differences in Interest, Perceived Personal Capacity, and Participation in STEM-Related Activities", *Journal of Technology Education* 24(1), pp. 18–33. Available at <https://eric.ed.gov/?id=ej991238>.
- Zehetmeier, D. et al. (2014) "Self-assessment of freshmen students' base competencies", *EDUCON: 2014 IEEE Global Engineering Education Conference, Institute of Electrical and Electronics Engineers*, pp. 429–438. Available at <https://doi.org/10.1109/EDUCON.2014.6826130>.