



Erasmus+



ROLEPL-ai

RESEARCH PLAN

ROLEPL-AI

*Project funded by the European Commission within the ERASMUS+ programme
under the agreement n° 2023-1-FR01-KA220-VET-000157570*

Deliverable 5.1 – Version 1

Type of Activity		
IO	Intellectual Output	X
A	Project Management and Implementation	
M	Transnational Project Meeting	
E	Multiplier Event	

Nature of the deliverable		
	Feedback from participants	
	Direct effect on participants and project partners	
	Practical & reusable resources for the practitioners	
	Research material bringing forward the reflexion in the sector	
	Community building tools	
	Partnerships and Cooperation	
	Dissemination material	
	Organizational and working documents	X

Dissemination Level		
PU	Public	X
CO	Confidential, only for members of the consortium (including the Commission Services)	

ACKNOWLEDGEMENT

This report forms part of the deliverables from a project called "ROLEPL-AI" which has received funding from the European Union's ERASMUS+ programme under grant agreement No. 2023-1-FR01-KA220-VET-000157570. The Community is not responsible for any use that might be made of the content of this publication.

This project aims at training soft skills remotely, by pushing the practice through the implementation of AI-based simulation.

The project runs from September 1st, 2023, to August 31st, 2025 (24 months), it involves 5 partners (Manzalab and Inceptive, France; VUC Storstrøm, Denmark; Fachhochschule Dresden, Germany) and is coordinated by Manzalab.

List of participants

Participant No.	Participant organisation name	Acronym	Country
1 (coord)	Manzalab	MZL	France
2	Inceptive	ICV	France
3	VUC Storstrøm	VUC	Denmark
4	Fachhochschule Dresden	FHD	Germany

CONTENT

1	Introduction	4
1.1	Overview	4
1.2	Deliverable positioning	4
1.3	Presentation.....	4
2	Hypothesis.....	5
3	Methodology	5
3.1	Participants	5
3.2	Experimentation Locations	6
3.3	Material.....	6
3.3.1	Hypothesis H1: Perceived Self-Efficacy in Soft Skills	7
3.3.2	Hypothesis H2: Presence	8
3.3.3	Hypothesis H3: Flow	8
3.3.4	Hypothesis H4: User Experience	9
3.3.5	Hypothesis H5: Open-mindedness about AI.....	10
3.4	Experimental Workflow	11
3.5	Use Case	13
4	Conclusion.....	16
5	Bibliography	17
6	Annexe.....	18
6.1	Consent Form	18
6.2	Tutorial.....	19
6.2.1	Paper Version – First Pilot	19
6.2.2	In-app Version – Second Pilot	20

Abbreviations

[AI] Artificial Intelligence
[MPS] Multimodal Presence Scale
[FSS] Flow Short Scale
[UEQ] User Experience Questionnaire
[UX] User Experience

1 INTRODUCTION

1.1 OVERVIEW

This report outlines the strategic approach for evaluating both the suitability and effectiveness of the AI-based simulation developed within the project. To achieve this, a research methodology will be employed, including the development of a comprehensive measurement protocol. The research will integrate quantitative and qualitative methods, ensuring a robust and holistic understanding of user interaction, learning progression, and pedagogical outcomes.

The evaluation of the AI-based simulation will focus on user-centred assessment tools—including Likert-scale questionnaires, open-ended surveys—grounded in established scientific frameworks. The collection of user feedback, both educators and learners, will provide a rich dataset to inform the refinement of the tool and the validation of the pedagogical model. This comprehensive research approach will ensure that the developed solution is not only technologically sound but also pedagogically impactful.

1.2 DELIVERABLE POSITIONING

D5.1 is developed as part of Work Package 5, at the transition point between the development and evaluation phases of the ROLEPL-AI project. It builds upon earlier deliverables, particularly the insights from D2.1 (Report of the literature review about research on AI and Learning), D2.3 (Recommendations for use of AI in education and ALTAI self-assessment), and the use case requirements defined in D3.1.

The research plan sets the methodological framework for the project's experimentation phase and is directly linked to Tasks 5.2 and 5.3, which involve piloting and analysing the AI-based simulation in educational settings.

This deliverable is an essential component of the evaluation process, ensuring consistency and coherence in the collection, interpretation, and application of results throughout the experimentation phase.

1.3 PRESENTATION

The objective of the project is to evaluate the relevance and potential of the ROLEPL-AI prototype for training in the field of tourism and services. The experiment aims to examine the impact of AI-based practical training in a 3D virtual environment on trainees' development of soft skills.

The evaluation will focus on participants' performance while using ROLEPL-AI, identifying both the strengths and limitations of this new AI-driven training approach in asynchronous learning contexts.

We will investigate how the prototype influences the learning process, the organisation of training, and the development of learners' skills. This will involve

assessing participant performance before and after using ROLEPL-AI, with a focus on intra-group comparisons to measure change over time.

2 HYPOTHESIS

To address this research question, we developed several hypotheses based on prior needs analysis and a literature review:

H1 – The perceived self-efficacy in soft skills competencies is positively impacted by the use of AI-based training.

H2 – The immersive environment generates a high level of presence, which positively influences the user experience.

H3 – The ROLEPL-AI training produces a high level of flow during the training phase.

H4 – AI-based training provides a positive user experience.

Following the first round of experimentation, new insights emerged, highlighting the need to also evaluate how the application influences users' perception of AI technology itself.

H5 – The perception of AI-based applications is positively influenced by the use of the ROLEPL-AI application.

3 METHODOLOGY

3.1 PARTICIPANTS

The evaluation will be conducted in two distinct phases, aligned with key stages in the development and refinement of the ROLEPL-AI prototype.

The first evaluation phase is scheduled for November 2024, followed by the second phase in April 2025. Both phases aim to test the hypotheses outlined in the research plan, applying the same methodological approach, questionnaires, and evaluation tools to ensure consistency and comparability across time. This repeated-measures design allows us to observe how users interact with the evolving technology and assess the impact of improvements made between the two phases.

Each partner school will involve 15 participants, selected from their vocational training programs in tourism and services. During each phase, participants will engage with at least two ROLEPL-AI simulation modules, representing key scenarios relevant to their training context.

To preserve the validity of the comparison and minimise learning bias, a different group of trainees will be recruited for each experimental phase. This ensures that all participants are novice users encountering the ROLEPL-AI tool for the first time, allowing for a clearer analysis of the initial impact and perceived usefulness of the system as it develops.

In addition to trainee feedback, data collection from teachers was included during the second experimentation phase and the Final Multiplier Event. Questionnaires were adapted and extended to gather insights from educators involved in preparing, observing, or facilitating the simulations. This addition helps to broaden the evaluation perspective, particularly regarding the usability, pedagogical value, and integration of ROLEPL-AI from a teaching standpoint.

3.2 EXPERIMENTATION LOCATIONS

The experimentation will take place at the two partner institutions: VUC and FHD, during regular class hours. Conducting the evaluations within the school setting provides a controlled environment, reducing external variables that could affect the learning experience—particularly technical issues such as unstable internet connections. This setup ensures that the focus remains on assessing the pedagogical impact of the ROLEPL-AI prototype, rather than testing the technological infrastructure or students' access to it.

Each institution will allocate a dedicated room for the sessions, where the experiment will be carried out under consistent and standardized conditions. The sessions will last approximately two hours, during which trainees will engage with a minimum of two ROLEPL-AI modules. Every participant will have access to an individual desk and computer, ensuring personal space and uninterrupted focus during the simulation and evaluation activities.

To minimize the risk of technical disruptions, the Wi-Fi connection will be tested in advance in each location. This preparation is crucial to support the seamless operation of the ROLEPL-AI application and to preserve the flow and immersion necessary for reliable evaluation of the training experience.

By conducting the experiments in this manner, the project ensures that the collected data reflects genuine user interaction and learning outcomes, rather than being influenced by avoidable logistical or technical constraints.

3.3 MATERIAL

Each participant may bring a personal laptop to the experimentation session. To ensure smooth execution and address potential technical issues, each school will provide additional laptops as backup, in case of connectivity problems or if a participant is unable to bring their own device. Headphones with microphone are extremely important if interaction with AI is oral.

For optimal interaction with the ROLEPL-AI application—particularly when the simulation involves oral communication with the AI—participants are strongly

encouraged to use headphones with a built-in microphone. This equipment is essential for ensuring clear audio input and output, which significantly improves the quality of the interaction and the overall learning experience.

To evaluate the proposed hypotheses, data will be collected through online surveys, which consolidate various validated tools and custom-designed questions. These surveys will be accessible via a secure internet link and administered before and after the ROLEPL-AI training sessions to assess changes in perception and experience.

On the day of the evaluation, each participant will be assigned a numbered test account, which will be used to access the ROLEPL-AI application. At the same time, participants will receive the links to the pre-test and post-test surveys. The account number will serve as an anonymous identifier, allowing us to match each participant's survey responses without collecting personal information.

Participants will be seated at individual desks. This setup ensures a comfortable and organized environment, allowing participants to complete the entire session—simulation and surveys—without having to move or switch devices.

3.3.1 Hypothesis H1: Perceived Self-Efficacy in Soft Skills

Evaluating soft skills presents a unique challenge, as they are not easily measurable like technical competencies. Direct performance assessments are prone to numerous biases—especially in short-term or artificial settings—which can compromise the reliability of the results. Therefore, this study adopts an indirect evaluation approach by focusing on participants' perception of their ability to manage complex interpersonal situations, a proxy for soft skills development.

This approach is grounded in Bandura's theory of self-efficacy, which posits that an individual's belief in their capacity to execute behaviours necessary to produce specific outcomes is a key predictor of both learning motivation and performance. Numerous studies support the influence of self-efficacy on learning engagement, persistence, and results. In the context of this project, the goal is not to measure actual performance but to examine whether the ROLEPL-AI tool helps students feel more capable and confident when facing situations that demand the use of soft skills.

To capture this, a custom pre- and post-training survey was designed, targeting participants' perceived ability to handle such situations. Since no existing self-efficacy scale directly measures soft skills across diverse domains, we developed a set of three original items, designed to be broadly relevant and applicable to the training scenarios used in ROLEPL-AI. These items are evaluated using a 5-point Likert scale, ranging from 1 - Completely disagree to 5 - Completely agree:

1. I feel insecure in my ability to know how to respond to a client.
2. When unexpected problems occur, I don't handle them well.
3. I feel confident in my capacity to manage an aggressive client.

3.3.2 Hypothesis H2: Presence

The level of presence experienced by participants will be evaluated using the Multimodal Presence Scale (MPS) developed by Makransky et al. This tool will be administered immediately after the ROLEPL-AI session and consists of a set of Likert-scale items designed to capture users' subjective experience of presence in a virtual environment.

The MPS is based on Lee's (2004) theory of presence, which proposes that both real and virtual experiences can be understood through three distinct domains:

- Physical presence – the sense of being in and interacting with a virtual environment;
- Social presence – the sense of interacting with social actors within that environment;
- Self presence – the awareness of one's own identity and reactions within the virtual space.

The Multimodal Presence Scale reflects this framework by including three dimensions, each assessed through five items. Participants respond using a 5-point Likert scale, ranging from 1 (Completely Disagree) to 5 (Completely Agree). This standardised approach allows for a structured and validated assessment of how immersive and realistic the ROLEPL-AI experience feels from the user's perspective.

Figure 1 Multimodal Presence Scale (Makransky & al.)

Table 4
Final selection of items for the MPS and designation of area attribute.

Label	Item	Area attribute
Physical Presence		
PHYS_2	The virtual environment seemed real to me.	PR
PHYS_3	I had a sense of acting in the virtual environment, rather than operating something from outside.	NAPM
PHYS_4	My experience in the virtual environment seemed consistent with my experiences in the real world.	PR
PHYS_5	While I was in the virtual environment, I had a sense of "being there".	SBVE
PHYS_10	I was completely captivated by the virtual world.	NPARE
Social Presence		
SOC_1	I felt like I was in the presence of another person in the virtual environment.	SC
SOC_2	I felt that the people in the virtual environment were aware of my presence.	HR
SOC_3	The people in the virtual environment appeared to be sentient (conscious and alive) to me.	HR
SOC_5	During the simulation there were times where the computer interface seemed to disappear, and I felt like I was working directly with another person.	NASM
SOC_7	I had a sense that I was interacting with other people in the virtual environment, rather than a computer simulation.	NAASI
Self-presence		
SELF_2	I felt like my virtual embodiment was an extension of my real body within the virtual environment.	SBE
SELF_3	When something happened to my virtual embodiment, it felt like it was happening to my real body.	SBC
SELF_4	I felt like my real arm was projected into the virtual environment through my virtual embodiment.	SBE
SELF_6	I felt like my real hand was inside of the virtual environment.	SBC
SELF_7	During the simulation, I felt like my virtual embodiment and my real body became one and the same.	SBC

Note. Physical realism (PR), not paying attention to real environment (NARE), sense of being in the virtual environment (SBVE), not aware of the physical mediation (NAPM), sense of coexistence (SC), human realism (HR) not aware of artificiality of social interaction (NAASI), not aware of the social mediation (NASM), sense of bodily connectivity (SBC), sense of bodily extension (SBE).

3.3.3 Hypothesis H3: Flow

The level of flow experienced by participants during their use of ROLEPL-AI will be assessed using the Flow Short Scale (FSS) developed by Rheinberg, Vollmeyer, and Engeser. This validated instrument is designed to measure the extent to which individuals experience a state of optimal engagement—commonly referred to as “flow”—while performing a task.

Flow is a psychological state characterized by deep concentration, a sense of control, intrinsic motivation, and the feeling that time is passing quickly. It is particularly relevant in educational and simulation contexts, where engagement and immersion are key to effective learning.

The Flow Short Scale includes 13 items, administered immediately after the ROLEPL-AI session, to capture the learner's subjective experience during the simulation. Each item is rated using a 7-point Likert scale, ranging from 1 (Not at all) to 7 (Very much), with 5 (Partly) as the midpoint. This scale offers a nuanced understanding of the participant's cognitive, emotional, and motivational states during the training:

1. I feel just the right amount of challenge.
2. My thoughts/activities run fluidly and smoothly.
3. I don't notice time passing.
4. I have no difficulty concentrating.
5. My mind is completely clear.
6. I am totally absorbed in what I am doing.
7. The right thoughts/movements occur of their own accord.
8. I know what I have to do each step of the way.
9. I feel that I have everything under control.
10. I am completely lost in thought.
11. Something important to me is at stake here.
12. I won't make any mistakes here.
13. I am worried about failing.

This tool enables a comprehensive evaluation of the flow experience across cognitive clarity, emotional absorption, perceived control, and performance confidence—factors which are essential in determining the effectiveness and engagement level of the ROLEPL-AI training.

3.3.4 Hypothesis H4: User Experience

The overall user experience (UX) of the ROLEPL-AI application will be assessed using the User Experience Questionnaire (UEQ) developed by Laugwitz, Held, and Schrepp. This evaluation will take place immediately after the ROLEPL-AI session, alongside other post-session assessments.

The UEQ is a standardized tool designed to capture users' perceptions of the usability and emotional impact of interactive systems. It measures both pragmatic qualities (e.g., efficiency, clarity) and hedonic qualities (e.g., stimulation, novelty), offering a well-rounded view of the user's experience.

For this study, we will use the short version of the UEQ, which consists of a series of semantic differential scales. Each item presents a pair of opposite adjectives, and participants are asked to rate their experience by selecting a position on a 7-point scale. For example, users indicate where their experience falls between "obstructive" and "supportive."

The items used in this study are:

- Obstructive ○ ○ ○ ○ ○ ○ Supportive
- Complicated ○ ○ ○ ○ ○ ○ Easy
- Inefficient ○ ○ ○ ○ ○ ○ Efficient
- Confusing ○ ○ ○ ○ ○ ○ Clear
- Boring ○ ○ ○ ○ ○ ○ Exciting

In addition to the closed-question format of the UEQ, we will also include a set of open-ended questions to gather qualitative feedback. These questions aim to capture users' suggestions, challenges, and insights, with a particular focus on interface design and usability, difficulties encountered during the session, desired features or improvements, as well as overall satisfaction and future expectations.

The open-ended questions are as follows:

1. Have you encountered any technical bugs? If so, which ones?
2. What are your overall impressions of the ROLEPL-AI environment?
3. Does the educational material provide sufficient information for you to use the application effectively?
4. What specific information do you feel is missing from the educational material?
5. What improvements would you suggest to make the educational material more useful?
6. Were there any features you didn't understand? If so, which ones and why?
7. Did you feel any tools or actions were missing from the application?
8. Do you have any suggestions for improving the ROLEPL-AI metaverse?
9. Do you have any other comments about the tool?

This combination of quantitative and qualitative feedback will provide valuable insights into how the ROLEPL-AI application is perceived by users and how it can be improved to better support learning engagement and effectiveness.

3.3.5 Hypothesis H5: Open-mindedness about AI

Following insights gained during the first experimentation phase, it became evident that participants' perception of AI technology plays a significant role in shaping their overall experience with the ROLEPL-AI application. This observation led to the addition of a new hypothesis—H5—focused on evaluating whether and how the use of the ROLEPL-AI tool influences participants' views on AI in general.

To explore this variable, a short questionnaire was developed, comprising two closed-ended questions and one open-ended question. The goal is to assess the evolution of participants' perception of AI-based applications and how this perception may impact their user experience and openness to AI-enhanced learning environments.

Participants are asked to respond to the questionnaire both before and after using the ROLEPL-AI simulation, allowing for a direct comparison of their attitudes toward AI and any shift that may have occurred as a result of their experience.

The pre-use questionnaire includes two closed-ended questions rated on a 5-point Likert scale (1 – Very negative to 5 – Very positive), as well as one open-ended question:

1. What is your general perception of AI?
2. Do you believe AI can be beneficial in your field of work or daily life?
3. What concerns, if any, do you have about AI?

Following the simulation, the post-use questionnaire includes a set of three follow-up questions designed to explore how participants' views may have evolved:

1. How has your perception of AI changed after using the application?
2. Do you now see AI as more beneficial in your field of work or daily life?
3. What, if anything, did you learn about AI that changed your perspective?

By combining quantitative ratings with qualitative insights, this addition to the research framework supports a more comprehensive understanding of how direct interaction with an AI-based educational tool can influence broader attitudes toward artificial intelligence—particularly in the context of education, training, and future professional use.

3.4 EXPERIMENTAL WORKFLOW

The day of the experimentation, each participant will receive the ROLEPL-AI access link and a numbered test account. Along with this, they will receive direct access to the pre-test and post-test survey links. The account number will serve as an anonymous identifier for both surveys to match responses.

Before beginning the session, participants will be asked to check that all links are working properly on their laptops. A brief introduction and explanation of the session will then be provided by the facilitator, including the objectives of the experiment and instructions for using the ROLEPL-AI tool. Participants will also be informed about the data collection process and asked to provide their informed consent for the use of their anonymised responses (see Consent Form in Annexe).

The session will begin with the pre-test survey, which includes:

- Basic demographic questions (e.g., age, gender, familiarity with immersive environments and AI)
- A self-assessment of soft skills self-efficacy, using a 3-item Likert scale questionnaire
- (In the second experiment only) A short survey about participants' perception of AI before using the tool

Once the pre-test is complete, participants will log into ROLEPL-AI using the assigned test accounts in order to complete a short tutorial and set up their avatar. This familiarisation phase allows participants to become comfortable with the environment and the core features they will use during the experimentation.

The tutorial (see Tutorial in Annexe) includes essential actions such as navigating within the virtual world, using the toolbar, interacting with avatars. Completing this phase in advance is critical to ensure that participants are technically ready and do not lose time during the actual session.

After the familiarisation phase, participants will complete two or more simulation modules designed for the experiment, each representing a simulated roleplay scenario within the virtual environment.

At the end of the training modules, participants will be asked to complete the post-test survey, which includes:

- The same soft skills self-efficacy scale, to measure perceived change
- (In the second experiment only) A short survey about participants' perception of AI after using the tool
- The Multimodal Presence Scale (MPS), to assess their sense of presence during the simulations
- The Flow Short Scale (FSS), to evaluate their level of engagement
- The User Experience Questionnaire (UEQ), to assess usability and satisfaction
- Additional open-ended questions to collect qualitative feedback on their experience

The post-test portion is expected to take approximately 30 minutes, and the full session—including introduction, training, and evaluation—will last around 90 minutes.

Table 1 Implementation Steps

Description	Duration	Documents & Material	Human Resources
Explanation of the process	5 min	Experimenter briefing, computer and a room for all participants	1 experimenter/teacher
Placement of participants at their workstation	5 min	Computer, headset & desk for each participant	1 experimenter/teacher
Handover of Pre-Test Questionnaires: - Consent form - Socio-demographic questionnaire - Questionnaire on perception of soft skills competencies - Questionnaire on open-mindedness about AI	10 min	Online pre-test questionnaire (link already open on all computers)	1 experimenter/teacher

Connection to the Work Environment and Familiarisation	20 min	Numbered test account assigned to each participant for anonymous responses (number to be recalled in pre- and post-test questionnaires)	2 experimenters/teachers
Completion of Roleplay 1	10 min	Computer, headset & desk for each participant	2 experimenters/teachers
Completion of Roleplay 2	10 min	Computer, headset & desk for each participant	2 experimenters/teachers
Handover of Post-Test Questionnaires: - Perception of soft skills competencies - Open-mindedness about AI - MPS (Multimodal Presence Scale) - FSS (Flow State Scale) - UEQ (User Experience Questionnaire)	30 min	Computer, headset & desk for each participant	1 experimenter/teacher
END	Approx. 90 min		

3.5 USE CASE

Following a series of collaborative meetings, the project partners agreed that the AI-based simulation training would focus on conflict management, a core soft skill relevant across multiple roles in the tourism and services sector. To provide a realistic and engaging environment for this training, the context of a job fair was selected as the setting in which learners would apply their conflict resolution skills. This decision is documented in detail in Deliverable 3.1.

The situation model developed for this use case (see Figure 2) guided the design of the simulation scenarios and defined the roles that students would assume during the training. Based on this model, students from the two partner institutions were assigned specific roles that align with their educational context and expected learning outcomes:

- Students at FHD take on the role of Fair Managers, responsible for overseeing the entire event and addressing high-level issues.
- Students at VUC act as Fair Assistants, interacting directly with visitors and managing more operational-level challenges.

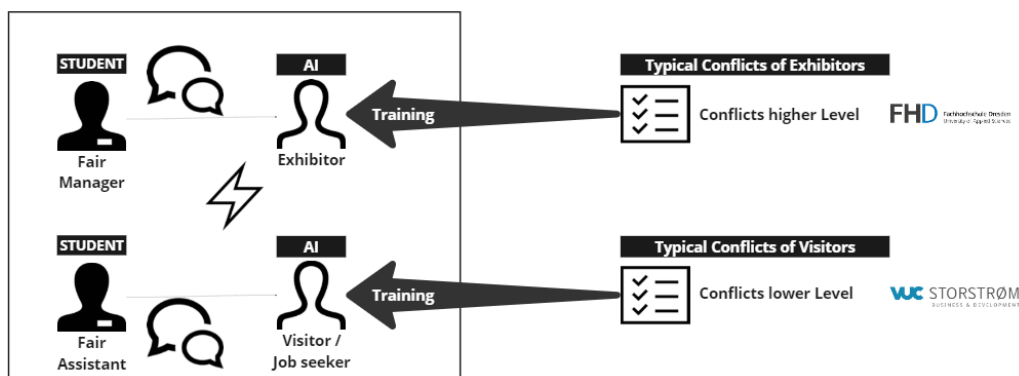


Figure 2 Situation model

The conflict scenarios featured in the simulation are derived from a typology developed by the partners during the creation of the pedagogical and technical documentation for AI training. Each simulation includes four types of conflict per role level (assistant and manager), as outlined below:

Table 2 Types of conflict per level

Visitor / Job seeker (low)	Exhibitors (high)
1. Communication and Information Flow	1. Booth Placement Issues
2. Customer Management / Visitor Flow / Traffic / Guidance System	2. Technical Difficulties
3. Health and Safety Concerns	3. Event Schedule
4. Disputes Over Services	4. Disputes Over Services

To ensure coherence and immersion, a dedicated AI character has been designed for each conflict type and level. Each bot consistently embodies the same fictional persona, allowing participants to engage with distinct characters and build contextual understanding. For example, a character like Noah Miller (see Figure 3) may play different roles depending on the level: addressing communication issues as a job seeker or raising booth-related concerns as an exhibitor.



Figure 3 Job fair setting with AI-based simulation character

The complete interaction scenario is presented in Figure 4. Each step in the scenario represents a phase in the simulation experience. Elements marked in

green indicate essential features necessary for the successful execution of the experimentation (e.g., avatar dialogue, conflict resolution feedback), while yellow squares indicate optional, “nice-to-have” features, the relevance of which will be evaluated during the first round of pilot testing.

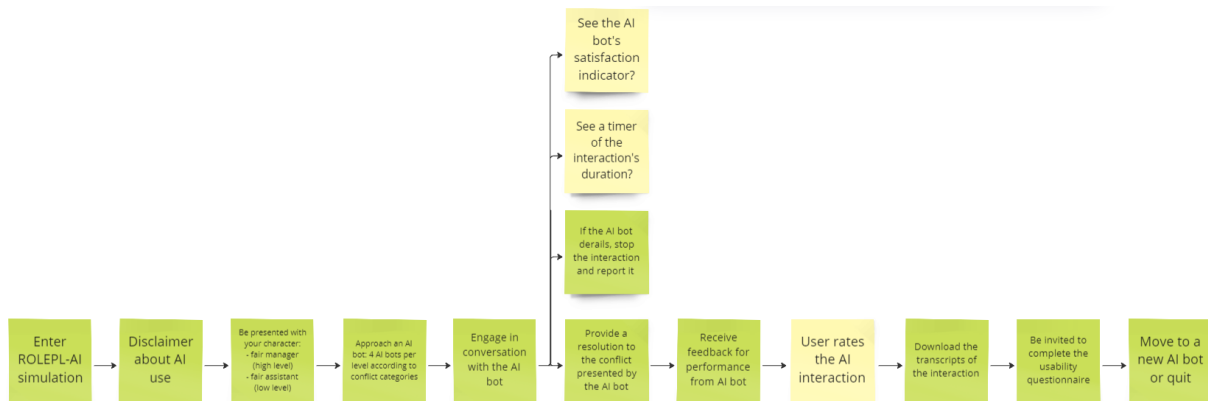


Figure 4 Scenario validated by partners

This use case forms the basis for both the technical deployment of ROLEPL-AI and the evaluation of its effectiveness in developing soft skills. It ensures a consistent, structured, and pedagogically grounded experience for all participants, while offering flexibility to refine or expand features based on feedback from the experimentation phases.

4 CONCLUSION

This research plan provides a structured, multi-dimensional framework for evaluating the ROLEPL-AI prototype, with a particular focus on its capacity to support the development of soft skills in vocational education. Grounded in established theories and adapted to the specificities of AI-based immersive learning environments, the methodology outlined in this report ensures consistency across the two experimentation phases while allowing for meaningful comparison and iterative refinement.

Five hypotheses have been formulated to guide the evaluation: from assessing perceived self-efficacy in soft skills and presence within the virtual environment, to flow, user experience, and evolving perceptions of AI. Each hypothesis is supported by dedicated tools and surveys, combining quantitative and qualitative data collection to provide a holistic understanding of user engagement and learning outcomes.

The use case—conflict management in the context of a job fair—offers a highly relevant and authentic setting for learners to apply and reflect on their soft skills. The scenario has been carefully designed to differentiate between roles (Fair Manager and Fair Assistant) and to introduce realistic conflict types, supported by distinct AI characters for each situation. This ensures a coherent and pedagogically meaningful learning experience for participants at both partner institutions.

By clearly defining the experimental workflow, participant preparation, and data collection methods, this research plan lays a solid foundation for the upcoming pilot phases. The integration of pre- and post-test surveys, coupled with structured roleplay modules, will allow for a detailed analysis of the impact of ROLEPL-AI on learners' perceptions, engagement, and skill development.

Finally, this research plan is not only instrumental in evaluating the current implementation of ROLEPL-AI but also serves as a reference for future iterations and adaptations of AI-based simulation tools in other educational contexts. It ensures that the evaluation process remains rigorous, replicable, and aligned with the broader pedagogical goals of the project.

Despite these strengths, a few limitations must be acknowledged, particularly the reliance on self-reported data, the difficulty of measuring soft skills directly, and the absence of longitudinal tracking or a control group. Technical and contextual constraints may also influence user experience across settings. These limitations will be taken into account when interpreting the results and refining the evaluation strategy in future phases of the project.

5 BIBLIOGRAPHY

Bandura, A. (2006). Guide for constructing self-efficacy scales. Self-efficacy beliefs of adolescents, 5(1), 307-337.

Bhattacharjee, A., & Premkumar, G. (2004). Understanding changes in belief and attitude toward information technology usage: A theoretical model and longitudinal test. *MIS quarterly*, 229-254.

Laugwitz, B., Schrepp, M. & Held, T. (2008). Construction and evaluation of a user experience questionnaire. In: Holzinger, A. (Ed.): USAB 2008, LNCS 5298, pp. 63-76.

Makransky, G., Lilleholt, L., & Aaby, A. (2017). Development and validation of the Multimodal Presence Scale for virtual reality environments: A confirmatory factor analysis and item response theory approach. *Computers in Human Behavior*, 72, 276-285.

Rheinberg, F., Vollmeyer, R., & Engeser, S. (2003). *Flow Short Scale* [Database record]. APA PsycTests.

<https://doi.org/10.1037/t47787-000>

6 ANNEXE

6.1 CONSENT FORM



CONSENT FORM

Dear Participant,

We invite you to participate in a research project exploring the impact of new technologies in the learning environment. This study examines the effects of simulation in virtual environments on learning soft skills. The entire experiment will last approximately 1 hour and 30 minutes and will be conducted in the following stages:

1. Participate in two soft skills learning simulations.
2. Complete a questionnaire.

These two stages will take place on November 12th, 2024.

How will the collected data be used?

The data collected will remain anonymous and confidential. All information will be grouped for statistical analysis, ensuring that individual participants cannot be identified.

Is participation in this study mandatory?

No. Participation in this study is voluntary. You are free to decide whether or not to participate, without needing to provide a reason or facing any negative consequences. Additionally, you may withdraw from the study at any time, with no need to justify your decision. If you choose to withdraw, all associated data will be destroyed.

Are there any risks, disadvantages, or benefits?

There are no known risks or inconveniences associated with this study. If at any point you feel uncomfortable, you are free to stop participating. This study does not offer monetary compensation.

For further information, questions, or comments, please contact us by email at test.teemew@manzalab.com. We will respond as soon as possible.

I have read and understood the information and consent form regarding the "ROLEPL-AI" project. I am aware of the conditions, risks, and benefits of my participation and have received answers to any questions I had about the project.

- ☐ I am over 18.
- ☐ I freely agree to participate in this study. This consent applies solely to this study on the date and time indicated above.

Full Name:

Date:

Signature:

Time:

ROLEPL-AI - Confidentiel

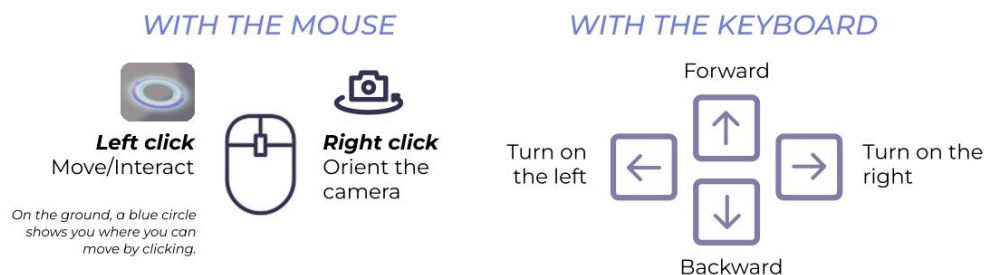
Figure 5 ROLEPL-AI's Consent Form Example

6.2 TUTORIAL

6.2.1 Paper Version – First Pilot

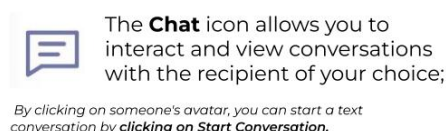
ROLEPL-ai FIRST STEPS

MOVING AROUND



USING THE MENU

On the left of the screen, you will find a lateral menu; use it for the following features:



INTERACTING *WITH OTHERS*

In the menu at the bottom of your screen, you will find the different available interactions:

THROUGH ANIMATION



THROUGH SPEECH



A microphone icon will appear in areas where you can speak with others. You can **activate it** or **mute it by clicking again**. To interact with another person, simply **click on their avatar**.

INTERACTING *WITH OBJECTS*



When you **hover** over a clickable object, a highlight effect is added. When you **click** on the object, the corresponding interactions are offered to you.

To exit an interaction, simply click on the icon (which varies depending on the context) to the right in the blue-bordered box.

Figure 6 ROLEPL-AI's First steps Paper Tutorial

6.2.2 In-app Version – Second Pilot

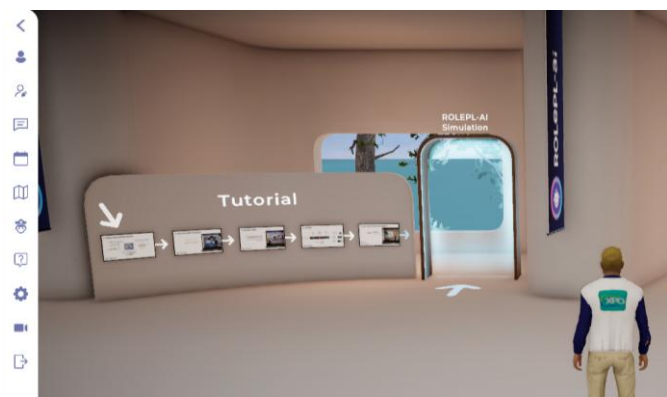
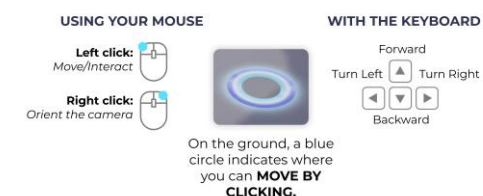


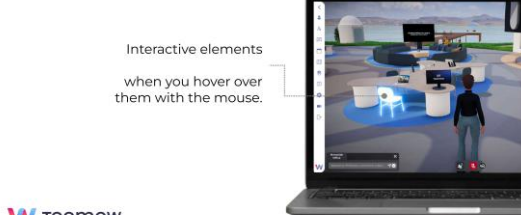
Figure 7 ROLEPL-AI's First steps In-app Tutorial screenshot (above) and the detailed content of the five screens displayed (below)

1- Move around the space



Teemew

2- Interact with Teemew



Teemew

3- Spatial audio



Teemew

4- Toolbar



Teemew

5- Let's start the AI-simulation!



ROLEPL-ai

Teemew