ERASMUS-EDU-2022-PI-FORWARD-LOT1

[101086701] Discovery Space:
An AI-Enhanced Classroom for Deeper Learning in STEM

D.4.2 Impact Assessment Methodology and Instruments

WP4 Implementation Scenarios, Pilots and Impact Assessment

This document was prepared by **C. Conradty**

Document information page

Project Reference No	101086701 - Discovery Space
Deliverable	D4.2 Impact Assessment Methodology and Instruments
Work package	WP4: Implementation Scenarios, Pilots and Impact Assessment
Type	R - Document, report
Dissemination Level	PU - Public (fully open)
Editors	Catherine Conradty (UBT)
Contributors	Rosa Doran (NULCIO) & F.X. Bogner (UBT) E. Anagnostopoulou (NTUA), Olga Dziabenko (UDEUSTO)
Reviewer(s)	ATC; Evangelia Anagnostopoulou (NTUA)
Document version	Draft V8
Document status	☒ Final version
Date	28/6/2024

Disclaimer

Funded by the European Union under grant agreement No 101086701. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them.

Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
V1 Draft	4/2023		Cathérine Conradty
V2 Table of Content	3/2024		Cathérine Conradty
V3	24/4/2024		Cathérine Conradty
V4	07/6/2024		Cathérine Conradty
V5	10/6/2024		Cathérine Conradty
V6	11/6/2024	Online Tool Information QR codes	Cathérine Conradty
V7	17/6/2024	Review	Evangelia Anagnostopoulou
V8	17/6/2024	Integration of review	Cathérine Conradty
final	28/6/2024	Integration of results of discussion during the CM	Cathérine Conradty

Executive Summary

Deliverable 4.2, titled "Impact Assessment Methodology and Instruments," offers a thorough examination of various impact assessment options for the Discovery Space Project. It details these options and explores their potential applications. The document concludes by proposing an assessment framework designed to evaluate the effectiveness of the Discovery Space Project.

The primary objectives include evaluating teacher satisfaction with professional development programs, ensuring the quality of the Educational Learning Environment (ELE) tool, and assessing the development of deeper learning competencies among students.
The methodology combines both quantitative and qualitative approaches, including standardised tests, such as surveys, interviews, and focus groups. Key challenges such as resource limitations, cultural differences, and the complexity of educational environments are addressed to ensure the reliability and validity of the assessment.

This structured approach aims to provide empirical data that informs decision-making, policy development, and the continuous enhancement of educational practices. By systematically evaluating the impact of ICT-based educational tools, the Discovery Space project seeks to refine its methods and improve educational outcomes over time.

Table of Contents

List of Figures

List of Tables

Structure of the document

This report is organized into four chapters and a bibliography, beginning with the introduction (Chapter 1) and concluding with the appendix, providing questionnaires and guidelines (Chapter 4). Chapter 2 presents a theoretical framework of impact assessment that promotes project development and insights into the impact of the project, including the discussion of the target groups and challenging factors of assessment strategies. Chapter 3 outlines the Methodology of Data Collection and Analysis of the Discovery Space. It discusses the impact assessment and feedback tools, methods and tools for feedback collection, and concludes with the Tool Set for Discovery Space implementations. The evaluation is conducted on three levels: the first level involves immediate feedback from teachers about PD or ELE, the second level consists of a teachers' questionnaire, and the third level is a sophisticated impact study that measures the effect of the Discovery Space Scenarios on the competence development of participating students. The questionnaires used are described in detail and can also be found in the appendix.

1. Introduction

Over recent decades, scientific and technological advancements have reshaped modern society, necessitating rapid adaptation and continuous learning. Education, particularly through the lens of deeper learning, is crucial for preparing individuals to meet the challenges of the 21st century, remain competitive in the job market, and be engaged citizens.
D2.2 Deeper Learning Competence Framework discusses how Deeper Learning can be facilitated through innovative teaching strategies and technological infrastructure. Dede advocates for methods such as collaborative investigations, extended inquiries, and interdisciplinary projects supported by up-to-date technology– such as ELE of Discovery Space. Key technologies include collaboration tools, online and hybrid learning environments, maker tools, immersive media, and simulations.

This document outlines an assessment strategy for the Discovery Space project, based on the findings about Deeper Learning from D2.1. The report establishes the necessary feedback loops that are crucial for project evaluation, such as feedback on the user acceptance of our ELE. Additionally, it provides a detailed examination of how to evaluate deeper learning competencies and measure the impact of the Discovery Space scenarios on their development. As detailed in D2.2 “Deeper Learning Competence Framework”, paragraph 2, deeper learning involves developing a set of competencies that enable students to understand academic content deeply and apply their knowledge effectively. The William and Flora Hewlett Foundation defines these competencies as mastery of academic content, critical thinking, effective communication, collaboration, learning strategies, and academic mindsets. They can be categorised into three domains: cognitive, intrapersonal, and interpersonal.
Appropriately selected measurement instruments are discussed, and the questionnaire is presented in detail.

2. Impact Assessment Methodology and Instruments

Impact assessment in educational projects is a systematic process of evaluating the effects and outcomes of educational interventions and programs. It involves the collection and analysis of data to determine the extent to which educational activities achieve their intended goals and produce measurable changes in participants and institutions. This process is crucial in educational settings as it provides evidence-based insights that inform decision-making, policy development, and the continuous improvement of educational practices.
The introduction of innovative ICT-based tools into educational environments necessitates a rigorous assessment of their impact on teaching and learning processes. This report outlines the approach taken to adapt these tools to specific educational contexts and evaluate their effectiveness. Special attention is given to understanding their mode of action in this new area through meticulous assessment.
Impact assessment provides empirical data that help creators and educators make informed, evidence-based decisions. By understanding what works and what doesn't, stakeholders can allocate resources more effectively and implement strategies that are proven to yield positive outcomes (Leithwood et al., 2004). The feedback loops established by impact assessment allow for continuous improvement. Regularly assessing the impact of educational interventions enables the Discovery Space consortium to refine and enhance methods, leading to better educational outcomes over time (Guskey, 2000).
We consider several target groups in our assessment. Besides the students, who are the direct users of our AI-supported learning interventions, we are also interested in the impact on teachers and their perspectives.

2.1. Objectives of the Impact Assessment

To develop the Discovery Space assessment design, the project's goals, objectives, and target groups were first identified. The objectives of the Discovery Space Impact Assessment operate on three levels.

First, the satisfaction of teachers with the professional development programme and the ELE tool must be ensured. Relevant surveys feed directly into the ongoing development of scenarios, training sessions, and the ELE tool.

To ensure the quality of the ELE tool, it must be continuously evaluated and improved through process-oriented feedback. Therefore, teachers' concise yet meaningful feedback on the Discovery Space Scenario implementation is essential. This feedback loop adheres to best practices in developing new learning environments, especially in ICT programming. Feedback from teachers on the Educational Learning Environment (ELE), including their interests and AI-related needs in daily school life, informs the development of professional development programs and ELE improvements aligned with teachers' needs.

Thirdly, the effectiveness of ICT-based applications in diverse educational settings on DL Competencies needs to be evaluated. Therefore, the core assessment focuses on the benefits students gain from the ELE. Thus, measurement instruments were designed to align with the competencies the ELE aims to foster. The most comprehensive aspect of the assessment examines whether the ELE tool achieves the desired impact on students' competency development. A significant challenge lies in devising an evaluation strategy capable of identifying this beneficial effect. Unlike cognitive knowledge, competencies develop continuously and may only become evident after a certain latency period. Process-oriented assessment in education is particularly challenging, making it easy to overlook positive effects.

To address this, we plan a concise evaluation based on the model of competency development for deep learning. A pilot study will test and refine a comprehensive quantitative questionnaire to create a streamlined tool for the main study. This refined tool will then be implemented as part of a best practice framework.

2.2. Methods of Impact Assessment: Quantitative and Qualitative

There are two primary measurement strategies: qualitative surveys, which involve detailed interviews with a small group of individuals, and quantitative methods, which use standardised questionnaires to collect data from a larger population.

2.2.1. Qualitative Methods

Focus groups are advantageous for international educational projects as they foster dynamic discussions that can reveal a wide range of perspectives and insights on educational interventions. The interactive nature of focus groups allows participants to build on each other's ideas, potentially uncovering deeper insights and generating new concepts. They are efficient for gathering diverse viewpoints simultaneously and can provide rich qualitative data that reflects the social and cultural contexts of different international settings. However, focus groups also have limitations.

They can be challenging to coordinate across different countries, especially considering time zone differences and logistical constraints. Language barriers and cultural nuances can affect group dynamics and the interpretation of data, potentially leading to misunderstandings or biased conclusions. Additionally, dominant personalities may overshadow quieter participants, skewing the data collected.

Focus groups, while valuable for creative workshops and identifying the needs, preferences, and priorities of target groups, are not utilized in Discovery Space's assessment design for measuring the effects of the Learning Companion. However, they could serve as a useful intervention in detailed workshops to evaluate concerns, hopes, and attitudes towards AI in the classroom.,

Interviews offer significant advantages as an evaluation tool in international education projects. They provide in-depth, detailed insights into participants' experiences, perspectives and cultural contexts, which are crucial for understanding the nuanced impact of educational interventions in different contexts. The flexibility of interviews allows for tailored questions and real-time customisation, enabling deeper exploration of complex issues and the collection of rich qualitative data. Building a relationship of trust through face-to-face interaction can lead to more honest and open answers, improving the data quality. However, interviews also come with challenges. They can be time and resource-consuming and require skilled interviewers and extensive logistical coordination, especially in an international context. Language barriers and cultural differences can also affect the clarity and interpretation of responses, which can lead to misunderstandings and bias.
The effect of the interview partner on the quality of the interview but also on the content of the statements must be critically considered and could sometimes be difficult to assess. The objectivity of the answers and whether the intervention or the interview had a decisive effect on the answers can sometimes be difficult to assess. Nevertheless, interviews could significantly enrich the Discovery Space assessment, particularly when conducted in relaxed settings such as summer schools.

2.2.2. Quantitative Assessment

Quantitative methods of assessment are surveys with standardised tests and statistical analysis of case studies. Surveys with standardised tests offer several advantages as assessment tools in international educational projects. They enable the collection of large amounts of data from diverse populations, allowing for comprehensive statistical analysis that can identify trends, measure outcomes, and compare results across different cultural and educational contexts. Standardised tests ensure consistency and reliability, making it easier to benchmark performance and assess the effectiveness of interventions objectively. However, these tools also have disadvantages. Designing surveys and tests that are culturally relevant and unbiased can be challenging, and language differences may affect participants' understanding and responses, potentially compromising the validity of the data. Additionally, while standardised tests can measure certain competencies effectively, they may not capture the full range of skills and contextual nuances important in international settings. Furthermore, quantitative questionnaires tend to be too long, which, in turn, has a negative impact on the quality of the answers. A balance must, therefore, be struck between the validity of the tool and the length of the questionnaire. Despite these challenges, when carefully designed and implemented, surveys with standardised tests provide valuable quantitative data that can enhance the understanding and evaluation of educational projects on a global scale.

2.3. Challenges in Impact Assessment and the Solutions

Despite its importance, impact assessment in education faces challenges. Ensuring high-quality, reliable data is often a challenge due to limited resources such as time (House, 1980). **Complexity of Educational Environments**: The dynamic and multifaceted nature of educational settings can make it difficult to isolate the effects of specific interventions (Cousins & Whitmore, 1998). Due to the complexity of the educational environment, its evaluation is influenced by many confounding and unmeasured factors, so that even a well-designed study resembles a poorly controlled field experiment. Therefore, these studies require large groups of participants willing to complete pre- and post-tests.
Another challenge is the diversity of **languages** used in European projects. For both quantitative questionnaires and qualitative interviews, the questions must be standardised and free of semantic shifts in meaning in all

languages. To this end, the tool sets must be carefully translated and checked before they can be used in in-depth studies.

Consequently, conducting comprehensive impact assessments can be **resource-intensive** and require a lot of time, money and expertise (Weiss, 1998). Unfortunately, time is a scarce commodity in schools, so despite the expertise and interest of participating schools, this factor can strongly influence the quality of the survey. This risk factor must be adequately considered in the design.

Therefore, we will use **digital data collection methods** as part of the Discovery Space Assessment Design. For questionnaires, we will utilise the tool "Umfrageonline.ch" ("survey online"), which is optimised for use on smartphones. The server is located in Switzerland, ensuring compliance with data protection regulations. With only one link, the survey is provided in all of the partners' languages; Participants can choose from the provided languages.

2.4. Evaluation Target Group

The following list outlines the target groups for our evaluation, the potential areas of interest and suitable assessment methods.

Target Group	**Assessment Method** and Area of Interest	**paragraph**
Teachers participating in PD	**Focus Groups** on Visionary Workshops: list of needs and wishes / **SCOT Analysis towards AI at School**	**3.1**
	Feedback: Satisfaction with the platform or Quality of training	**3.2**
	Interview: Attitudes towards AI in education	**3.3** not recommended
	Questionnaire: Competencies and Attitudes towards AI	**3.4**
	Questionnaire: measuring DL Competencies of Teachers	not applicable
Teachers implementing Scenarios in their classrooms	**Questionnaire:** Development of skills and competencies	not applicable
	Feedback: Quality of tool and Usability of the tool at school	**3.2**
	(**Questionnaire**: Attitudes towards AI)	**3.4** not recommended
Students participating in lessons with Discovery Space Scenarios	**Path Analysis of Metadata:** Problem-solving Competencies?	**3.5**
	Questionnaire: Development of skills and competencies	**3.6**
	Interview: Ideas and expectations regarding AI in school and professional life	not applicable
	(**Questionnaire**: Attitudes towards AI)	**3.4** not recommended

3. Methodology of Data Collection and Analysis

Based on the analyses in paragraph 2, the following recommendations for Discovery Space are proposed.

3.1. Focus Groups on PD with SCOT Analysis for AI in Schools

As part of the project assessment, focus groups can be conducted during Professional Development (PD) to compile a comprehensive list of teachers' needs and wishes regarding implementing AI in schools. These workshops provided a collaborative platform where teachers could openly discuss their experiences, expectations, and concerns about integrating AI into their teaching practices. SCOT (Strengths, Challenges, Opportunities, and Threats) analysis framework helps to systematically analyse and address these insights.

During the workshops, teachers will be guided through a SCOT analysis to evaluate AI in their educational environments. SCOT stands for Strengths, Challenges, Opportunities and Threats.
During the facilitation process, the moderator should ensure that teachers initially remain in the "problem space," thoroughly addressing their **Threats** and **Challenges**. Afterwards, the moderator guides them into the "solution space," where the focus shifts exclusively to **Strengths** and **Opportunities**. If participants mentally revert to the problem space during the Strength and Opportunity discussion, they should be gently reminded to leave behind the extensively discussed difficulties and concentrate on the possibilities, mindful of the limited time available.
Challenges could include concerns about data privacy, the need for significant professional development, and potential resistance to change among staff. Participants might explain in detail **Threats** such as the digital divide, dependency on technology, and the deficiency in their qualifications so that they feel unprepared for fundamentally different, modern forms of teaching. They will identify **Strengths**, which might be enhanced personalised learning, improved data-driven decision-making, and the potential for reducing administrative tasks. The teachers will also explore **Opportunities**, probably highlighting areas where AI could foster innovative teaching methods, support differentiated instruction, and provide real-time feedback.

This structured approach allowed us to gather valuable insights into the practical needs and aspirations of teachers, ensuring that the AI tools developed are aligned with their expectations and are feasible within the school context. The SCOT analysis facilitated a thorough understanding of the factors influencing AI adoption, thereby informing the development of targeted strategies to support teachers in effectively integrating AI into their classrooms.

3.2. Feedback of Discovery Space PD

Tools for Online Feedback about the ELE:
Using a star rating system, points, or feedback boxes to evaluate online offers has proven useful. This approach helps maintain customer satisfaction by allowing them to promptly report any frustrations, thereby fostering customer loyalty.
The star rating system allows users to provide a visual assessment, typically ranging from one to five stars, with more stars indicating higher satisfaction or quality. This type of rating is widely used on e-commerce platforms, review sites, and various other online portals to help gauge consumer satisfaction and guide potential users' decisions.
The feedback boxes have a very motivating effect on a certain group of people because they can provide direct feedback. Negative feedback, in particular, can be obtained in this way. This is of particular interest because users can discover shortcomings hidden from programming and programme development insiders. However,

the Discovery Space team can only offer this option if it has sufficient staff, as in particular, negative feedback requires immediate response.
Even though desirable, these features are not supported by the platform in the first version. The alternative would be a brief questionnaire for students provided by the agent to evaluate the overall experience with the Discovery Space platform, as well as the effectiveness and usefulness of the conversational agent.

Table 3: Feedback Questionnaire integrated into the Platform (designed by the ICCS team)

1. How would you rate the usefulness of the platform? ○ Extremely useful ○ Very useful ○ Somewhat useful ○ Not so useful ○ Not at all useful
2. The Discovery Space platform assisted me in understanding [pendulum concept] ** (fill in the subject matter of the scenario)* Entirely disagree ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Entirely agree
3. Overall, how satisfied are you with the platform? ○ Very satisfied ○ Somewhat satisfied ○ Neither satisfied or dissatisfied ○ Somewhat dissatisfied ○ Very dissatisfied
4. Have you used a similar application before? ○ yes → which ○ no
5. How would you rate the suggestions that the agent provided? ○ Very useful ○ Somewhat useful ○ Neither useful nor useless ○ Somewhat useless ○ Very useless
6. Discussion with the agent helped me to understand the [pendulum experiment] ** (fill in the subject matter of the scenario)* Entirely disagree ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Entirely agree
7. The agent's instructions were simple to understand and provided specific suggestions Entirely disagree○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Entirely agree
8. How engaged do you feel with the Discovery Space agent? Not engaged ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Highly engaged
9. How likely are you to use this platform in the future for your educational needs? Not at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ very likely
10. Did you notice improvements in understanding the [pendulum experiment] * after interacting with the platform? ** (fill in the subject matter of the scenario)* ○ yes ○ no

Feedback Questionnaire for Quality Assurance of Professional Development Training

In order to ensure the highest quality of our professional development training, we have designed a feedback questionnaire specifically for participating teachers. This questionnaire is essential for gathering valuable insights into the effectiveness of the training, identifying areas for improvement, and ensuring that the content meets the needs of educators. By collecting this data, we can refine our programs and provide an even better learning experience. The questionnaire is carefully structured to capture both qualitative and quantitative feedback, focusing on various aspects of the training, including content delivery, engagement, and practical application.

The questionnaire was designed in cooperation with the teams of UDEUSTO and UBT. It is offered as an online survey with a QR code: https://forms.gle/zq54QBWvR4er12d67

Tools for instant live Feedback during Classroom Courses:

This guide describes which feedback tools the workshop leader should integrate into the course. These include tools for the ongoing learning process, such as Mentimeter, a quiz to determine whether the students have understood the lesson's content. Other tools give us feedback on the participants' satisfaction, as well as self-reflection on learning content and skills development.

Live feedback systems are interactive tools that are particularly useful for gathering immediate feedback during training sessions while increasing participation and engagement (Wang & Tahir, 2020). Established offers are *Kahoot!, Poll Everywhere, Mentimeter.*

These tools will be used during the training sessions. They provide direct feedback and insight into the participants' ideas and learning environment. They also encourage participants to use them in their own teaching.

Feedback-Boards allow participants to post comments and ideas on a digital wall in real-time, creating an open and collaborative feedback environment (Thompson 2017). Established offers are *Padlet, Miro.*

These tools allow teachers to actively participate in the Discovery Space project and collect their feedback together on the notice board.

3.3. Interviews for Assessment of Attitudes toward AI at school

Interviews are a vital qualitative research method allowing in-depth exploration of participants' attitudes, experiences, and perceptions. In educational research, interviews provide rich, detailed data that can reveal the complexities of teachers' attitudes towards new technologies in schools, such as AI. According to Kvale (1996), interviews are effective for obtaining nuanced insights that other methods, such as surveys, may not capture. They facilitate a deep understanding of how teachers perceive AI, their challenges, and their needs for effective implementation (Creswell, 2013). Furthermore, interviews allow for real-time clarification and probing, which enhances the reliability and depth of the data collected (Patton, 2002).

An interview guide was prepared for suitable settings, such as multi-day training courses or conferences (Table 1). **The interview is suitable for both teachers and students.**

Table 1: Interview Guideline for Teachers

Interview Guide: Teachers' Attitudes toward AI at School (10 minutes)		
1. Introduction	1 min	*Greet and thank the participant for their time.* *Briefly explain the purpose of the interview*: We are interested in understanding your attitudes toward using AI in school settings.
2. General Attitude towards AI	2 min	(Can you describe your overall attitude towards AI in education?) What are your initial thoughts or feelings when you hear about AI being used in schools?
3. Perceived Benefits	2 min	What potential benefits do you see in using AI in your teaching practice? Can you provide an example of how AI could positively impact your classroom?
4. Concerns and Challenges	2 min	(What concerns or challenges do you have regarding using AI in education?) Are there specific issues you think might arise with integrating AI in your school?
5. Support and Resources Needed	2 min	What type of support or resources would you need to effectively integrate AI into your teaching? (How do you think the school administration can assist teachers in adopting AI tools?)
6. Conclusion	1 min	Do you have any additional thoughts or suggestions regarding using AI in schools? *Thank the participant for their insights and time.*

3.4. Teachers' Questionnaire: Competencies and Attitudes towards AI

An influential factor for students is the attitudes and competencies of their teachers. Therefore, the competency profile of teachers of participating student groups, as well as their attitudes towards AI in the classroom, would be of interest. To develop the competency profile, the competence questionnaire is suitable for both students and teachers (**paragraph 3.6**).

Instead of interviews, the AI Attitude Scale (AIAS-4 Scale) could be useful in this context. This tool has been successfully validated and was published by Grassini in 2023. However, the practical effectiveness of this scale in gathering reliable data about teachers remains to be seen.

- Likert scale: agree not at all 1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 – 10 completely agree
- Publication: Grassini S (2023) Development and validation of the AI Attitude Scale (AIAS-4): a brief measure of general attitude toward artificial intelligence. *Front. Psychol.* 14:1191628. doi: 10.3389/fpsyg.2023.1191628

Table 2: AI Attitude Scale (AIAS-4 Scale)

1. I believe that AI will improve my life
2. I believe that AI will improve my work
3. I think I will use AI technology in the future
4. I think AI technology is positive for humanity

3.5. Path Analysis of Metadata

We plan to apply path analysis of metadata in the Discovery Space ELE. Path analysis is a robust technique for uncovering user paths. We expect to reveal students' learning paths, offering profound insights into their interactions and progress through educational content (Figure 1). By examining the sequences of actions and decisions students make within a digital platform, patterns and critical decision points that influence learning outcomes might be identified. This method enables the mapping of student journeys, highlighting both common pathways and divergent routes that may signal areas of difficulty or interest. Jovanović et al. (2008) emphasise that path analysis elucidates the dynamic and nonlinear nature of learning processes, facilitating the customisation of educational experiences to better address individual student needs. Furthermore, these analyses can assess the effectiveness of various instructional designs and learning activities, providing data-driven insights that support the continuous enhancement of digital learning environments (Siemens, 2013). By leveraging metadata, path analysis not only deepens our understanding of student behaviour but also aids in developing more adaptive and personalised learning systems.

It is not yet clear whether the learning paths are more indicative of the users' problem-solving competencies or the design of the learning scenario. If the former, it would represent an innovative approach, potentially opening new avenues for researching the complex topic of problem-solving skills. If the latter, it would provide valuable feedback on the complexity and quality of the learning scenario, aiding in the development of guidelines for future learning scenarios.

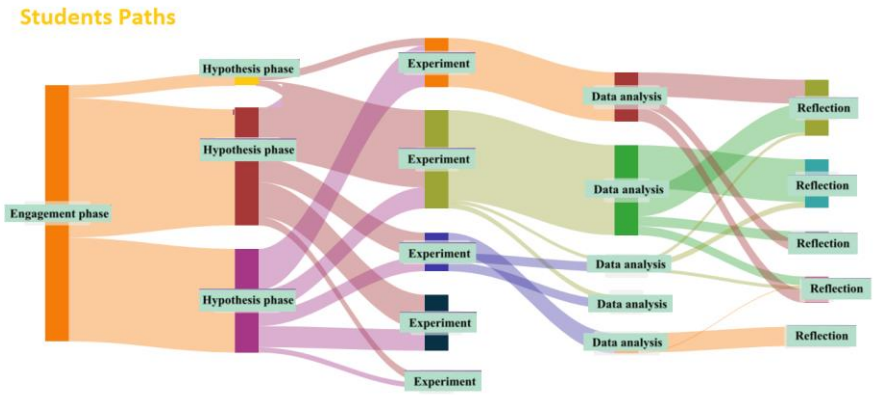

Figure 2 Concept Visualisation: a hypothetical Result of the Path Analysis based on Metadata of a class with the individual learning paths of the students

3.6. Students' Questionnaires

Questionnaires offer the possibility of surveying a large number of participants and analysing complex concepts such as attitudes. Based on the Discovery Space Deeper Learning Competencies Model described in D2.2 (Figure 2) the measurement instruments described below were selected. The **Assessment** Instruments are chosen based on their ability to measure effectively. Thus, proven instruments with published validity are selected. The design was intended to measure changes in educational practices, student learning outcomes and students' development of DL Competencies.

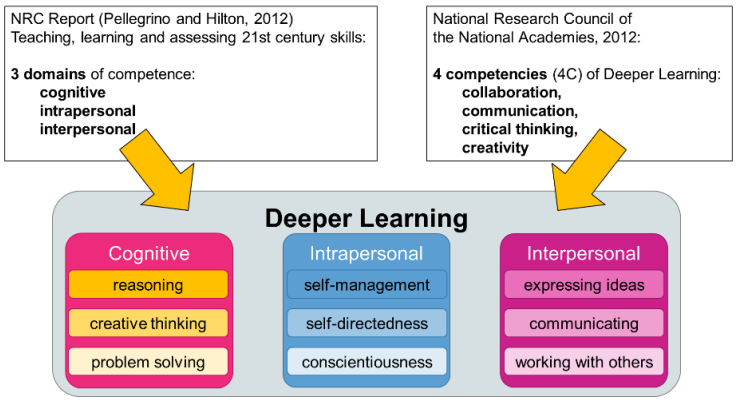

Figure 3 Illustration of the Deeper Learning Model

3.6.1. Competencies for Deeper Learning

A range of competencies need to be practised and developed to achieve deep learning. These competencies extend beyond mere knowledge acquisition, focusing on the application, analysis, synthesis, and evaluation of information. As outlined in the literature, these competencies are interrelated and often developed simultaneously through various learning experiences and environments. They equip students not just for academic success but also for effective and responsible participation in the modern world. The Deeper Learning competencies are summarised in Table 4.

Table 4: Deeper Learning Framework
based on the Three Domains of Competencies and the Four Core Competencies. (taken from D2.1)

	Cognitive Domain *intellectual abilities, understanding*	**Intrapersonal Domain** *self-management: emotional intelligence, self-regulation*	**Interpersonal Domain** *interacting effectively with others: communication, collaboration, empathy*
Ways of Thinking *critical thinking, creativity, learning to learn, metacognition*	deep understanding effective problem-solving	self-directed learning, innovative thinking	
Ways of Working *communication and collaboration*			*effective* teamwork and interaction
Tools for Working *information literacy, ICT literacy*	*effectively* utilise tools for learning and working		
Living in the World *Citizenship, career, personal and social responsibility*		self-awareness adaptability	understanding and participating *effectively* in societal and global contexts

Table 5: Deeper Learning Competencies (taken from D2.2)

Critical Thinking and Problem Solving	Analysing complex problems and devising innovative solutions is central to deeper learning.	[1,2]
Self-Regulated Learning	Self-directed learning trains students to set personal learning goals, organise time and evaluate their own progress. The student, instead of the teacher, takes responsibility for learning.	[3,4]

Collaboration and Communication	Working effectively in groups and communicating ideas needs training, self-reflection - and empathy (→ EQ)	[4,5]
Creativity and Innovation	Creativity is not (just) about arts or innovative output but also involves understanding complex concepts, developing skills, and fostering dispositions that are crucial for comprehensive learning.	[6,7]
Digital Literacy	Actively utilising the advantages and disadvantages of digital resources in lessons must be part of contemporary education of "digital natives".	[8,9,10]
Metacognitive Skills	Self-reflection and Mastery of one's own learning process. It contributes to self-efficacy, self-determination, and a fruitful error management culture.	[11,12]
Emotional Intelligence (EQ)	Recognising and managing one's own emotions, as well as empathising with others, is the basis for communication, collaboration, peaceful conflict management and error management culture.	[13]
Global and Cultural Awareness	Understanding and appreciating cultural differences and awareness of global issues is crucial for collaboration.	[14,15]
Ethical and Social Responsibility	Moral development is an important goal of education. Ethics are highly contingent on cultural circumstances. In a globalised world, contemporary liberal ethics standards and responsibility awareness are essential.	[16]

3.6.2. Design & Tools

*Deeper learning is not just about acquiring knowledge; it's about **unlocking the potential** within each learner to think critically, solve complex problems, and become an active participant in shaping their own future. (Discovery Space)*

The tools presented below are established questionnaires designed to assess competencies that contribute to deeper learning. Figure 3 illustrates the Deeper Learning Framework, which is based on the Three Domains of Competencies and the Four Core Competencies, and shows how these tools are connected to the competencies they measure.

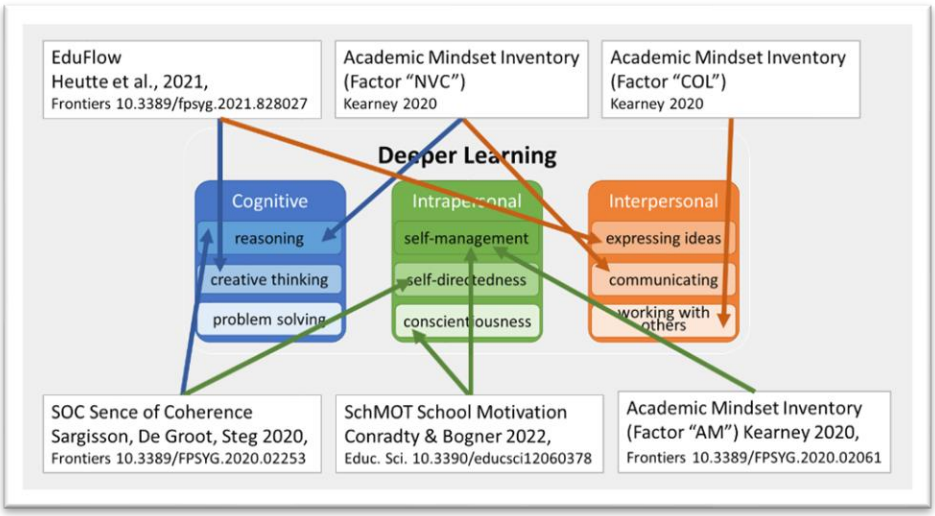

Figure 4 Deeper Learning Framework linked with the Tools measuring the Competencies.

Academic Mindset

The Academic Mindset Inventory (AMI) is designed to assess students' academic mindsets, encompassing their beliefs about learning and intelligence and students' attitudes towards academic challenges. The tool reliably measures various aspects of academic mindset, such as growth mindset, belongingness, and purpose. By using the AMI, educators can better understand and support students' academic beliefs and attitudes, tailoring interventions to promote a growth mindset, a sense of belonging, and a clear purpose in their educational journeys.

In Discovery Space we implement the subscales "Academic Mindset", "Academic I do" scale and the "Social Emotional Learning" measuring aspects of communication and collaboration.

Both Academic Mindset subscales are asked with the question: "How much do the statements apply to you?". The Social Emotional Learning is asked with "How easy or difficult are the following challenges for you?".

- Likert scales; Subscales:
  Academic Mindset 1 = not at all true, 2 = a little true, 3 = mostly true, 4 = completely true
  Academic Mindset: "I do" Questions 1 = never, 2 = sometimes, 3 = often, 4 = always
  Social Emotional Learning (Collaboration & NVC) 1 = very difficult, 2 = difficult, 3 = easy, 4 = very easy
- Publication: Kearney, C. A., Sanmartín, R., & Gonzálvez, C. (2020). The school climate and academic mindset inventory (SCAMI): Confirmatory factor analysis and invariance across demographic groups. *Frontiers in Psychology*, 11, 2061. DOI: 10.3389/FPSYG.2020.02061

Table 6: Items of the Academic Mindset Subscales (Kearney et al 2020)

Academic Mindset
Challenging myself won't make me any smarter.
There are some things I am not capable of learning.
If I am not naturally smart in a subject, I will never do well in it.
I don't participate in discussions because I am afraid people might think I am foolish.
I don't ask questions in class because people might think my questions are not smart.
I stop doing work if I feel like I can't do it well.
I only volunteer to answer a question if I am sure my answer is right.
Academic Mindset: "I do" Questions (Rewording)
I prepare for class by doing the readings or other assigned work.
I turn in assignments on the due date.
I actively participate in class.
I have all of my class materials with me.
I do more than what is expected of me.
Social Emotional Learning (Collaboration & NVC): How easy or difficult are the following challenges for you?
Knowing what my strengths are.
Knowing ways, I calm myself down.
Knowing the emotions I feel.
Knowing when my feelings are making it hard for me to focus.
Knowing what people may be feeling by the look on their face.
Learning from people with different opinions than me.
Knowing when someone needs help.

Getting through something even when I feel frustrated.
Being patient even when I am really excited.
Finishing tasks even if they are hard for me.
Setting goals for myself.
Doing my schoolwork even when I do not feel like it.
Being prepared for tests.
Getting along with my classmates.
Respecting a classmate's opinions during a disagreement.
Thinking about what might happen before making a decision.

Sence of Coherence (SoC)

The Sense of Coherence (SOC) is vital for learning, encapsulating an individual's ability to perceive their environment as comprehensible, manageable, and meaningful. This psychological framework empowers students to navigate educational challenges effectively, fostering resilience and adaptability. SOC plays a significant role in deeper learning by enabling students to engage fully with complex material, persist through difficulties, and find personal relevance in their studies. This engagement enhances understanding and retention and promotes the development of critical thinking and problem-solving skills, essential for deeper learning. By supporting a positive and proactive mindset, SOC helps learners transform educational experiences into meaningful and lasting knowledge.

- Likert scale 1 = not at all true, 2 = a little true, 3 = mostly true, 4 = completely true
- Publication: Sardu, C., Mereu, A., Sotgiu, A., & Andrissi, L. (2012). Antonovsky's sense of coherence scale: Cultural validation of SOC questionnaire and socio-demographic patterns in an Italian population. *Clinical Practice and Epidemiology in Mental Health*, 8, 1-6. DOI: 10.2174/1745017901208010001

Table 7: Sence of Coherence (SOC) (Sardu et al, 2012))

I don't really care about what's going on around me.				
I've been surprised by the behaviour of people I thought I knew well.				
I've been disappointed by people I've relied on.				
I have felt unfairly treated before.				
I currently find myself in an unfamiliar situation and don't really know what to do.				
My daily tasks give me joy and satisfaction.				
My daily tasks bore me and are a torture.				
I have very mixed feelings and ideas.				
Some feelings I would rather not have to endure.				
Many people, even those with a strong character, sometimes feel like losers in certain situations. I often feel that way.				
It usually turns out that I underestimated or overestimated the significance of events beforehand.				
It usually turns out that I had previously assessed the significance of events completely correctly.				
I often believe that what I do in everyday life makes little sense.				
I usually can't control my feelings.				
Until now, my life has had	no clear goals 1	hardly any goals 2	rough goals 3	very clear goals and purposes 4

School Motivation Subscale Self-Efficacy (SchMOT; SE)

The SchMOT builds upon the established Science Motivation Questionnaire (SMOT) by adapting it into a general school motivation scale (SchMOT). The SchMOT advances educational psychology by providing a robust tool for measuring student motivation, which is crucial for developing effective educational strategies and improving student outcomes. It offers a reliable measure of students' general motivation towards schoolwork, which can be used to develop and evaluate teaching concepts designed to enhance motivation. Within the Discovery Space project, Self-Efficacy is of special interest.

- Likert scale Likert scale 1 = not at all true, 2 = a little true, 3 = mostly true, 4 = completely true
- Subscale Self-Efficacy
- Publication: Conradty, C., & Bogner, F. X. (2022). Measuring students' school motivation. *Education Sciences*, 12(6), 378. DOI: 10.3390/educsci12060378

Table 8: School Motivation Subscale Self-Efficacy (Conradty & Bogner 2022)

School Motivation (SchMOT) Subscale Self-Efficacy
I am confident that I will do well in the exams.
I believe that I can achieve very good exam results.
I believe that I can master the content and practical requirements in my studies.
I am confident that I will do well in internships, projects or presentations.

Creativity

Creativity is a process of recognizing problems or gaps in knowledge, searching for solutions, and contributing to deeper learning (Chow, 2010). Research shows that creativity and intrinsic motivation depend on each other (Csikszentmihalyi, 1996; Jesus De, Rus, Lens, & Imaginário, 2013; Runco, 2014). Amabile and Tighe (1993) described creativity as driven by intrinsic motivation. At the neuronal level, a link with motivation was found in dopamine-related activities, which are key to human creativity (Aberg et al., 2016). This offers a promising start for research into fostering motivation through engagement, as creativity can support intrinsic motivation.
For measuring creativity, we focused on motivation and attitudes linked to personal creativity. This includes cognitive (thinking) and non-cognitive (motivation) dimensions (Miller & Dumford, 2016). Flow, a special state of consciousness, positively impacts individuals by leading to complete absorption and enjoyment in activities. It enhances intrinsic motivation, promotes creativity, and improves overall task performance (Csikszentmihalyi, 2000).
We used the “Act” subscale of CPAC to measure active cognitive processes indicating creative problem-solving. To measure flow, we applied the EduFlow Scale, which focuses on flow experiences in educational settings (Heutte et al., 2021), including four subscales: “cognitive control,” “immersion and time transformation,” “loss of self-consciousness,” and “autotelic experience” (well-being provided by the activity).

We reduced the Likert scale from 5 levels (including 3 = “not sure”) to 4 levels to align it with the rest of the questionnaire. This reduces the burden on students and encourages intuitive responses.
Our experience showed that 12-year-old students struggled with the concepts of “cognitive control” and “immersion.” This issue could stem from academic language, so we simplified the wording. Table 10 compares the original questionnaire with the adapted version. We aim to make the questionnaire suitable for all age and education levels, ensuring no subscales need to be omitted.

- Likert scale; 1 = Strongly disagree; 2 = disagree; 3 = agree; 4 = strongly agree
- Publication:
  Conradty, C., & Bogner, F. X. (2018). From STEM to STEAM: How to Monitor Creativity. *Creativity Research Journal*, *30*(3), 233–240. https://doi.org/10.1080/10400419.2018.1488195
  Heutte et al. (2021). Optimal Experience in Adult Learning: Conception and Validation of the Flow in Education Scale (EduFlow-2) *Frontiers in Psychology, section Educational Psychology*, ***12*** (1-12). DOI: https://doi.org/10.3389/fpsyg.2021.828027

Table 9: Creativity Tool Set (CPAC Act)

Cognitive Processes Associated with Creativity (CPAC) Subscales Act	
CPAC-1	When I tackled a subject, I tried to find as many ideas as possible.
CPAC-2	To find a possible solution to a problem, I asked other people for help.
CPAC-3	To find a solution, I looked at the problem from different points of view.
CPAC-4	I have imagined a possible solution in my mind's eye to check its usefulness.
CPAC-5	I incorporated an old solution in a completely new way.
CPAC-6	Imagining potential solutions to a problem leads to new insights.

Table 10: Creativity Tool Set (EduFlow)

EduFlow Subscale	**EduFlow (original, Pilot)**	**EduFlow (adapted, easy language)**
Cognitive control	I trust my ability to meet the high demands of the situation.	I believe I can handle everything that's happening.
Cognitive control	I feel completely in control of my actions.	I feel like I have full control over what I do.
Cognitive control	At each step, I know exactly what I have to do.	I know exactly what to do at each step.
Immersion and Time transformation	I am wholly absorbed in what I am doing.	I am totally into what I'm doing.
Immersion and Time transformation	I am deeply focused on what I am doing.	I am very focused on what I'm doing.
Immersion and Time transformation	I am losing track of time.	I am losing track of time.
Loss of self-consciousness	I don't care about what others may think of me.	I don't care what others think of me.
Loss of self-consciousness	I am not concerned about the judgement of others.	I'm not worried about what others say about me.
Loss of self-consciousness	I am not worried about what others might think of me.	I am not worried about what others think of me.
Autotelic experience	I have the feeling I am living a very exciting experience.	I feel like I am having a really exciting experience.
Autotelic experience	This activity brings me a sense of well-being.	My daily activities make me feel good.
Autotelic experience	When I talk about this activity, I feel such a deep emotion that I want to share it.	When I talk about this activity, I feel so excited that I want to share it with others.

3.6.3. Pilot and Standardisation

The evaluation design and tools will undergo rigorous testing in a pilot study to ensure optimal functionality and validity. Our goal is to refine internationally tested tools to perfectly suit our specific needs, enabling a comprehensive assessment of participant competencies. This approach aims to map out a broad spectrum of cognitive, interpersonal, intrapersonal, and motivational factors, thereby enhancing our understanding of sustainable and deeper learning processes.

To guarantee consistent usage of assessment instruments, we have standardised the distribution of questionnaires among participants. This standardisation reduces the need for extensive training of evaluation partners and ensures cost-effective utilisation of the tools. The partners carefully translated the questionnaires into all relevant languages and checked for semantic accuracy. Where available, previously validated translations of the questionnaires were used. The WP Lead Assessment (UBT) provides digital questionnaires to ensure uniform data collection.

Further, integrating these tools into the portal of our AI-supported learning environment links the intervention directly with evaluation. This integration allows for established quantitative tools to be complemented by metadata analysis of individual student behaviours on the learning platform. Such an innovative approach promises fresh insights into students' learning paths, potentially revolutionising our understanding of educational impact.

Since the pilot tests the general suitability of the questionnaire, a group of students will be asked to complete it once. The participants will not take part in a Discovery Space Scenario. This survey also serves as a control group, providing us with baseline competency values that are unaffected by the intervention.

3.6.4. Main Study: Data Collection and Analysis

A before-and-after design will be implemented to measure changes in the competencies due to the impact of the ELE. Baseline competency levels will be assessed before participating in a Discovery Space Scenario and reassessed after the implementation.

This design requires more effort, as the questionnaire should be completed approximately two weeks before the scenario. Additionally, an immediate follow-up survey after the lesson can be too demanding for the students; therefore, conducting the survey the next day may yield better acceptance. To accurately match the pre- and post-tests for individual-specific assessment rather than average changes, the questionnaires must be individually assigned while ensuring student privacy and anonymity are maintained.

The use of digital tools can simplify the collection and exchange of feedback. For large-scale evaluation studies, tools like Google Forms or SurveyMonkey can provide surveys and feedback forms, allowing teachers to quickly gather and analyse responses on various aspects of instruction with advanced analytics tools (Brooke 2016; Anderson 2018). Furthermore, online questionnaires make the data collection process more accessible and less time-consuming (Vanderlinde and van Braak 2010).

Two strategies for data collection will be pursued. Firstly, established online surveys will be offered to students via **QR code on their smartphones**. This has the weakness that the intervention is decoupled from the evaluation, and entire courses may forgo the evaluation. In addition, due to the high data protection regulations, difficulties may arise in assigning the questionnaires to each other in the pre-post-test design or linking them to the learning path in the ELE.

As a further option, we will test the **integration of the evaluation into the ELE**. Integrating data collection into the learning platform facilitates the process, eliminating the need for the time-consuming and costly distribution of paper questionnaires. Instead, students are presented with questions from our evaluation tool set as they

work on the platform. This streamlines the survey process, reduces the students' effort, and improves the quality of the survey data. Furthermore, the questionnaires can be assigned to a participant in a pre-post-test design, as well as to the metadata of the learning behaviour on the platform.

3.7. Outcomes and Insights

The evaluation will provide significant insights into the effectiveness of the deeper learning paradigm in fostering essential STEM competencies. These findings will inform the development of guidelines for implementing and enhancing deeper learning strategies.

Feedback and SCOT Analysis collected by trainer during professional development sessions will be gathered from the trainers by the Task Leader of T4.6 Impact Assessment, Reflection, and Recommendations (UBT). When appropriate, the results will be presented in meetings, which may occur during monthly online meetings or, at the latest, during consortium meetings, to discuss the results collectively and make decisions accordingly.

User feedback will be directly communicated to the developer of the Portal and Agent. The results from the questionnaire, which can be offered based on scenarios, will be forwarded by the Task Leader to the provider of both the scenario and the IT developers (UBT).

The complex data collections on competency developments will be thoroughly analyzed. For coding, matching, calibrating, factor analyzing, and relating the data, the data collections must be complete. Subsequently, the data will be analyzed using AMOS and SPSS software. It is desirable that Pilot Phase A (M17-23) can already provide complete datasets in a pre-post design, so that adjustments based on the results can be made for the pilot phase B. A final evaluation of the questionnaires will only begin after the completion of Pilot Phase B (M28-34). The results are expected at the end of the project.

4. References

References Assessment

- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. New Directions for Evaluation, 1998(80), 5-23.
- Fullan, M. (2001). The new meaning of educational change. Routledge.
- Guskey, T. R. (2000). Evaluating professional development. Corwin Press.
- House, E. R. (1980). Evaluating with validity. Sage Publications.
- Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). How leadership influences student learning. Wallace Foundation.
- Patton, M. Q. (2008). Utilization-focused evaluation. Sage Publications.
- Scriven, M. (1991). Evaluation thesaurus. Sage Publications.
- Weiss, C. H. (1998). Evaluation: Methods for studying programs and policies. Prentice Hall.

References Interview

- Creswell, J. W. (2013). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Sage Publications.
- Kvale, S. (1996). *Interviews: An Introduction to Qualitative Research Interviewing*. Sage Publications.
- Patton, M. Q. (2002). *Qualitative Research and Evaluation Methods*. Sage Publications.

References Students and Teachers Questionnaires: Design & Tools

- Grassini S (2023) Development and validation of the AI Attitude Scale (AIAS-4): a brief measure of general attitude toward artificial intelligence. *Front. Psychol.* 14:1191628. doi: 10.3389/fpsyg.2023.1191628
- Kearney, C. A., Sanmartín, R., & Gonzálvez, C. (2020). The school climate and academic mindset inventory (SCAMI): Confirmatory factor analysis and invariance across demographic groups. *Frontiers in Psychology*, 11, 2061. DOI: 10.3389/FPSYG.2020.02061 Conradty, C., & Bogner, F. X. (2022). Measuring students' school motivation. *Education Sciences*, 12(6), 378. DOI: 10.3390/educsci12060378
- Conradty, C., & Bogner, F. X. (2022). Measuring students' school motivation. *Education Sciences*, 12(6), 378. DOI: 10.3390/educsci12060378
- Sardu, C., Mereu, A., Sotgiu, A., & Andrissi, L. (2012). Antonovsky's sense of coherence scale: Cultural validation of SOC questionnaire and socio-demographic patterns in an Italian population. *Clinical Practice and Epidemiology in Mental Health*, 8, 1-6. DOI: 10.2174/1745017901208010001
- Conradty, C., & Bogner, F. X. (2018). From STEM to STEAM: How to Monitor Creativity. *Creativity Research Journal*, *30*(3), 233–240. https://doi.org/10.1080/10400419.2018.1488195
- Heutte, J., Fenouillet, F., Martin-Krumm, C., Gute, G., Raes, A. Gute, D., Bachelet, R. & Csikszentmihalyi, M. (2021). Optimal Experience in Adult Learning: Conception and Validation of the Flow in Education Scale (EduFlow-2) *Frontiers in Psychology, section Educational Psychology*, ***12*** (1-12). DOI: https://doi.org/10.3389/fpsyg.2021.828027

References Deeper Learning Competencies

1. Conradty, Cathérine; Bogner, Franz Xaver (2020): STEAM teaching professional development works: effects on students' creativity and motivation. In *Smart Learn. Environ.* 7 (1). DOI: 10.1186/s40561-020-00132-9

2. Zimmerman, Barry J. (2002): Becoming a Self-Regulated Learner: An Overview. In *Theory Into Practice* 41 (2), pp. 64–70. DOI: 10.1207/s15430421tip4102_2
3. Johnson, D. W.; Johnson, R. T. (2009): An educational psychology success story: Social interdependence theory and cooperative learning. In *Educational Researcher* 38 (5), pp. 365–379
4. Stahl, Gerry; Hesse, Friedrich (2009): Paradigms of shared knowledge. In *Computer Supported Learning* 4 (4), pp. 365–369. DOI: 10.1007/s11412-009-9075-7
5. Runco, Mark A. (2004): Creativity. In *Annual review of psychology* 55, pp. 657–687. DOI: 10.1146/annurev.psych.55.090902.141502
6. Conradty, Catherine; A. Sotiriou, Sofoklis; X. Bogner, Franz (2023): Full STEAM Ahead with Creativity. In : Pedagogy, Learning, and Creativity [Working Title]: IntechOpen
7. Meyers, Eric M.; Erickson, Ingrid; Small, Ruth V. (2013): Digital literacy and informal learning environments: an introduction. In *Learning, Media and Technology* 38 (4), pp. 355–367. DOI: 10.1080/17439884.2013.783597
8. Gilster, Paul (1997): Digital Literacy. New York: Wiley Computer Pub
9. Bawden, David (2008): Origins and concepts of digital literacy. In *Digital literacies: Concepts, policies and practices*, pp. 17–32
10. Flavell, John H. (1979): Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. In *American Psychologist* 34 (10), pp. 906–911. DOI: 10.1037/0003-066X.34.10.906
11. Veenman, Marcel V.J.; Verheij, Joke (2003): Technical students' metacognitive skills: relating general vs. specific metacognitive skills to study success. In *Learning and Individual Differences* 13 (3), pp. 259–272. DOI: 10.1016/S1041-6080(02)00094-8
12. Mai, Hoang Thi (2018): Developing Emotional Intelligence for Secondary School Students through Teaching Literature. In *IJLL* 6 (1). DOI: 10.15640/ijll.v6n1a4
13. Martin, Leisa A.; Smolen, Lynn Atkinson; Oswald, Ruth A.; Milam, Jennifer L. (2012): Preparing Students for Global Citizenship in the Twenty-First Century: Integrating Social Justice through Global Literature. In *The Social Studies* 103 (4), pp. 158–164. DOI: 10.1080/00377996.2011.601358
14. Merryfield, Merry M. (2000): Why aren't teachers being prepared to teach for diversity, equity, and global interconnectedness? A study of lived experiences in the making of multicultural and global educators. In *Teaching and Teacher Education* 16 (4), pp. 429–443. DOI: 10.1016/S0742-051X(00)00004-4
15. Resnik, David B.; Elliott, Kevin C. (2016): The Ethical Challenges of Socially Responsible Science. In *Accountability in research* 23 (1), pp. 31–46. DOI: 10.1080/08989621.2014.1002608
16. Pelegrino; Hilton (2012): Education for life and work. Developing transferable knowledge and skills in the 21st century. Washington, DC: The National Academies Press

5. Appendix

List of content of appendix

1. Interview Guideline for Teachers
2. AI Attitude Scale (AIAS-4 Scale)
3. Students' Feedback Questionnaire integrated into the Platform
4. Students' Questionnaire

5. Summary of Evaluation Methods and their benefits and challenges
6. Summary of Impact Assessment Challenges and why we should face them

Interview Guideline for Teachers

Interview Guide: Teachers' Attitudes Toward AI at School (10 minutes)

1. Introduction **1min**	***Greet and thank the participant for their time.*** *Note here of a concise, friendly way of introducing yourself to the interview partner:*
	Briefly explain the purpose of the interview: We are interested in understanding your attitudes toward using AI in school settings.
2. General Attitude towards AI **2min**	Have you got any experience with AI? *(Yes or no is expected. Immediately ask the next question)* What are your initial thoughts or feelings when you hear about AI being used in schools?
Notes	
3. Perceived Benefits	What potential benefits do you see in using AI in your teaching practice?
2min	
Notes	
	Can you provide an example of how AI could positively impact your classroom?
Notes	
4. Concerns and Challenges	*Use the question that suits you or your interview partner better.* (What concerns or challenges do you have regarding using AI in education?)
2min	Are there specific issues you think might arise with integrating AI in your school?
Notes	
5. Support and Resources Needed **2min**	*Use the question that suits you or your interview partner better.* What type of support or resources would you need to effectively integrate AI into your teaching? (How do you think the school administration can assist teachers in adopting AI tools?)
Notes	
6. Conclusion **1min**	Do you have any additional thoughts or suggestions regarding using AI in schools?
	Thank the participant for their insights and time. Note your own words.

AI Attitude Scale (AIAS-4 Scale)

This questionnaire is provided with a QR Code in all partners' languages.
The pilot's link is: https://www.umfrageonline.ch/c/DiscoverySpaceAI
Ask the UBT Team for links or QR codes.

SCAN ME

Dear Participant,
Welcome to the Discovery Space Project Survey!
Your participation helps us understand the needs and expectations of teachers, allowing us to develop better professional development programs. Your honest and independent insights are vital for improving education.
Rest assured, your responses will be kept strictly confidential. We deeply appreciate your time and effort, which make a significant contribution to the educational community.
Thank you for your participation!
Best regards,
The Discovery Space Team

My Age Group:
○ 20 – 30 ○ 31-40 ○ 41-50 ○ 51-60 ○ > 60

My Gender:
○ female ○ male ○ None / prefer not to tell

My Subject(s):
○ MINT ○ Languages ○ Arts ○ Sports ○ others

Below, you'll find a series of statements about attitudes. Don't overthink them—just go with your gut reaction and indicate how much you agree with each statement.

1. I believe that AI will improve my life									
1 I agree not at all	2	3	4	5	6	7	8	9	10 I completely agree
2. I believe that AI will improve my work									
1 I agree not at all	2	3	4	5	6	7	8	9	10 I completely agree
3. I think I will use AI technology in the future									
1 I agree not at all	2	3	4	5	6	7	8	9	10 I completely agree
4. I think AI technology is positive for humanity									
1 I agree not at all	2	3	4	5	6	7	8	9	10 I completely agree

Students' Feedback Questionnaire integrated into the Platform

1. How would you rate the usefulness of the platform?
   - ○ Extremely useful
   - ○ Very useful
   - ○ Somewhat useful
   - ○ Not so useful
   - ○ Not at all useful

2. The Discovery Space platform assisted me in understanding [pendulum concept] ** (fill in the subject matter of the scenario)*

   Entirely disagree ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Entirely agree

3. Overall, how satisfied are you with the platform?
   - ○ Very satisfied
   - ○ Somewhat satisfied
   - ○ Neither satisfied or dissatisfied
   - ○ Somewhat dissatisfied
   - ○ Very dissatisfied

4. Have you used a similar application before?

   ○ yes → which ○ no

5. How would you rate the suggestions that the agent provided?
   - ○ Very useful
   - ○ Somewhat useful
   - ○ Neither useful nor useless
   - ○ Somewhat useless
   - ○ Very useless

6. Discussion with the agent helped me to understand the [pendulum experiment] ** (fill in the subject matter of the scenario)*

   Entirely disagree ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Entirely agree

7. The agent's instructions were simple to understand and provided specific suggestions

   Entirely disagree○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Entirely agree

8. How engaged do you feel with the Discovery Space agent?

   Not engaged ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ Highly engaged

9. How likely are you to use this platform in the future for your educational needs?

   Not at all ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ very likely

10. Did you notice improvements in understanding the [pendulum experiment] * after interacting with the platform? ** (fill in the subject matter of the scenario)*

    ○ yes ○ no

Students' Questionnaire

This questionnaire is provided with a QR Code in all partners' languages.
The pilot's link is: https://www.umfrageonline.ch/c/DiscoveySpacePilot
Ask the UBT Team for links or QR codes.

Item List of the Pilot Questionnaire

1		my year of birth is
2		my biological gender is:
3	SEL-19	Knowing when someone needs help.
4	SEL-24	Doing my schoolwork even when I do not feel like it.
5	SEL-17	Knowing what people may be feeling by the look on their face.
6	SEL-13	Knowing what my strengths are.
7	SEL-14	Knowing ways I calm myself down.
8	SEL-15	Knowing the emotions I feel.
9	SEL-25	Being prepared for tests.
10	SEL-27	Respecting a classmate's opinions during a disagreement.
11	SEL-16	Knowing when my feelings are making it hard for me to focus.
12	SEL-18	Learning from people with different opinions than me.
13	SEL-20	Getting through something even when I feel frustrated.
14	SEL-21	Being patient even when I am really excited.
15	SEL-22	Finishing tasks even if they are hard for me.
16	SEL-23	Setting goals for myself.
17	SEL-26	Getting along with my classmates.
18	SEL-28	Thinking about what might happen before making a decision.
19	SOC-1.	I don't really care about what's going on around me.
20	SMOT 17	I am confident that I will do well in the exams.
21	SOC-7a.	My daily tasks give me joy and satisfaction.
22	AM-1	Challenging myself won't make me any smarter
23	SOC-5. ®	I have felt unfairly treated before.
24	SMOT 18	I believe that I can achieve very good exam results.
25	SOC-10.	Many people, even those with a strong character, sometimes feel like losers in certain situations. I often feel that way.
26	AM-2	There are some things I am not capable of learning
27	SOC-3.	I have been disappointed by people I have relied on.
28	AM-3	If I am not naturally smart in a subject, I will never do well in it
29	SOC-8. ®	I have very mixed feelings and ideas.
30	SMOT 19	I believe that I can master the content and practical requirements in my studies.
31	SOC-12. ®	I often think that what I do in everyday life has little meaning.
32	AM-4	I don't participate in discussions because I am afraid people might think I am foolish
33	SOC-9. ®	Some feelings I would rather not have to endure.
34	AM-5	I don't ask questions in class because people might think my questions are not smart
35	SOC-2.	I have been surprised by the behaviour of people I thought I knew well.

36	AM-6	I stop doing work if I feel like I can't do it well*
37	SOC-6. ®	I currently feel myself in an unfamiliar situation and don't quite know what exactly to do.
38	SOC-13. ®	I usually cannot control my feelings.
39	AM-7	I only volunteer to answer a question if I am sure my answer is right
40	SOC-7b.	My daily tasks bore me and are a torment to me.
41	SMOT 20	I am confident that I will do well in internships, projects or presentations.
42	SOC-11. (a)	Most of the time it turns out that I had underestimated or overestimated the importance of the events before.
43	CPAC-1	When I tackled a subject, I tried to find as many ideas as possible.
44	DO-8	I prepare for class by doing the readings or other assigned work.
45	CPAC-2	To find a possible solution to a problem I asked other people for help.
46	DO-9	I turn in assignments on the due date.
47	CPAC-3	To find a solution, I looked at the problem from different points of view.
48	DO-10	I actively participate in class.
49	CPAC-4	I have imagined a possible solution in my mind's eye to check its usefulness.
50	DO-11	I have all of my class materials with me.
51	CPAC-5	I incorporated an old solution in a completely new way.
52	DO-12	I do more than what is expected of me.
53	CPAC-6	Imagining potential solutions to a problem leads to new insights.
54	EduFlow1	I believe I can handle everything that's happening.
55	EduFlow2	I am totally into what I'm doing.
56	EduFlow3	I don't care what others think of me.
57	EduFlow4	I feel like I am having a really exciting experience.
58	EduFlow5	I feel like I have full control over what I do.
59	EduFlow6	I am very focused on what I'm doing.
60	EduFlow7	I'm not worried about what others say about me.
61	EduFlow8	My daily activities make me feel good.
62	EduFlow9	I know exactly what to do at each step.
63	EduFlow10	I am losing track of time.
64	EduFlow11	I am not worried about what others think of me.
65	EduFlow12	When I talk about this activity, I feel so excited that I want to share it with others.
66		How likely is it that you would take part in such a questionnaire again if it contributes to improving schools?

Summary of Evaluation Methods and their benefits and challenges:

Category	Tools	Benefits	Challenges	Solutions
Qualitative Methods	**Focus Groups**	• Foster dynamic discussions • Reveal deeper insights • Efficient for gathering diverse viewpoints simultaneously	• Coordination across countries • Time zone differences and logistical constraints • Language barriers and cultural nuances • Dominant personalities overshadowing quieter participants	• Detailed planning and scheduling • Use of translation services and cultural sensitivity training • Facilitation training to manage group dynamics
	Interviews	• In-depth, detailed insights • Tailored questions for deeper exploration • Build trust through face-to-face interactions	• Time and resource-consuming • Require skilled interviewers and logistical coordination • Language barriers and cultural differences • Assessing objectivity and impact of the interview	• Hire or train skilled interviewers • Use interpreters or multilingual interviewers • Develop standardised interview protocols
Quantitative Methods	**Surveys with Standardized Tests**	• Collect large amounts of data • Allow comprehensive statistical analysis • Ensure consistency and reliability for benchmarking performance	• Designing culturally relevant and unbiased surveys • Language differences affecting understanding and responses • Limited measurement of full range of skills and contextual nuances	• Pilot testing and iterative design process • Use of multilingual surveys and cultural adaptation • Comprehensive question design and validation
	Digital Data Collection Methods	• Efficient data collection • Compliance with data protection regulations • Multilingual accessibility	• Ensuring high participation rates • Data privacy concerns	• Use mobile-optimized surveys • Ensure compliance with data protection regulations (e.g., hosting servers in Switzerland)

Summary of Impact Assessment Challenges and why we should face them

Tools	Challenges	Benefits	Solutions
General Impact Assessment Challenge	• Ensuring high-quality, reliable data • Complexity of educational environments • Resource-intensive assessments	• Comprehensive understanding of intervention impacts	• Use digital data collection methods • Plan for extensive pilot studies • Develop best practice frameworks for streamlined assessments
Complexity of Educational Environments	• Difficulty in isolating effects of specific interventions • Influence of many confounding and unmeasured factors	• Thorough evaluation of intervention effects	• Large sample sizes • Comprehensive pre- and post-tests • Use of control groups where possible
Diversity of Languages	• Standardizing questions and avoiding semantic shifts	• Accurate and inclusive assessments	• Careful translation and back-translation processes • Involvement of native speakers in survey design and testing
Resource-Intensive Assessments	• High demands on time, money, and expertise • Scarcity of time in schools affecting survey quality	• High-quality data and insights	• Efficient project management • Leveraging technology for data collection and analysis • Engaging stakeholders early to ensure buy-in and resource allocation