Managing generative AI as organizational change: evidence from a multi-institutional study in education

Journal of Organizational Change Management

Tilia Stingl de Vasconcelos Guedes
Department of Digital Economy, FHWien der WKW, Vienna, Austria and NSI – Next Society Institute, Kazimieras Simonavicius University, Vilnius, Lithuania, and
Patrick Rupprecht and Isabel Rodenas
Department of Digital Economy, FHWien der WKW, Vienna, Austria

Received 27 January 2026
Revised 3 March 2026
Accepted 1 April 2026

Abstract

Purpose – This paper presents findings from a 1.5-year multi-institutional study on the integration of Generative Artificial Intelligence (GenAI) into educational institutions under the Vienna Chamber of Commerce and Industry (WKW). Grounded in Luhmann's social systems theory, the study conceptualizes GenAI integration as a challenge of organizational programming, requiring a shift from conditional (rule-based) to purposive (goal-oriented) approaches that accommodate technological uncertainty while fostering competency development. With the rapid rise of GenAI, educational institutions are facing a major organizational change that challenges established structures and routines and requires purposeful adaptation.

Design/methodology/approach – A mixed-methods design was applied across four institutions spanning secondary, vocational, tertiary, and adult education. The empirical study combined a large-scale learner survey (n = 1,577), educator focus groups (n = 23), and controlled experiments (n = 47) comparing student teams with and without access to GenAI.

Findings – Six critical success factors were identified. Beyond technical skills, outcomes depended on: (1) clear communication and role definition, (2) shared responsibility and team cohesion, (3) prior GenAI experience, (4) critical reflection on AI-generated content, (5) intrinsic motivation, and (6) capacity for innovative thinking. Strong teamwork yielded superior results regardless of AI use; unstructured AI adoption consistently weakened outcomes. Findings informed the establishment of a KI-Info-Hub, a centralized digital platform providing educators with access to training materials and evidence-based practices.

Originality/value – This study offers one of the first comprehensive, practice-based evaluations of GenAI in German-speaking vocational and higher education and provides actionable strategies for competency-oriented, human-centered GenAI integration.

Keywords Generative AI, Education, Human-AI interaction, Competencies, Organizational change

Paper type Research article

1. Introduction

Generative Artificial Intelligence (GenAI) has rapidly moved from a peripheral technological innovation to a pervasive element of everyday work and learning practices. Its widespread adoption is already reshaping professional fields such as marketing, journalism, communication, and management, with tangible effects on employment structures and entry-level job profiles (Hui *et al.*, 2023; Reddy *et al.*, 2025). Companies increasingly report that tasks traditionally assigned to junior professionals or freelancers are being partially automated or fundamentally transformed through GenAI applications (Scerri, 2024). Empirical research in higher education further indicates that tools such as ChatGPT are contributing to profound changes in learning processes, academic integrity, and the

Journal of Organizational Change Management
Emerald Publishing Limited
e-ISSN: 1758-7816
p-ISSN: 0953-4814
DOI 10.1108/JOCM-01-2026-0095

competencies required for responsible and critical use (Dikilitaş *et al.*, 2024). As a result, graduates face growing challenges when entering the labor market, as established competency profiles no longer align with emerging workplace demands.

Educational institutions therefore find themselves under acute pressure to respond. While GenAI is already deeply embedded in students' everyday practices – used to save time, generate or refine texts, and support learning and creativity – its integration into formal education remains largely unstructured. Empirical studies suggest that although students actively use tools such as ChatGPT and often perceive them as helpful, they may rely on AI-generated outputs without sufficient critical evaluation (Elkhodr *et al.*, 2023; Krupp *et al.*, 2023). High rates of unreflected acceptance and copy-paste strategies point to a growing gap between informal use and pedagogically guided, reflective engagement (Krupp *et al.*, 2023).

This situation creates a paradox: GenAI is omnipresent in learning practices, yet its educational use often lacks clear pedagogical intention, didactic framing, and explicit competency goals. The challenge is therefore not merely technical (Jain and Kiran, 2025). Public and professional discourse increasingly emphasizes that future competence requirements extend well beyond operational AI skills.

This tension requires a strategic reorientation of educational institutions, leading to the following research question:

> How can Generative AI be integrated into educational and professional development frameworks to enhance competency acquisition rather than merely accelerating output, and what core competencies are essential for effective human-AI collaboration?

This question is particularly urgent for institutions committed to practice-oriented education. In the absence of robust empirical evidence, educators risk either rejecting GenAI altogether or adopting it uncritically – both of which may undermine learners' long-term employability and professional judgment. This paper addresses this challenge by presenting insights from a 1.5-year multi-institutional study conducted across educational institutions under the umbrella of the Vienna Chamber of Commerce and Industry (WKW).

The study employs a mixed-methods research design combining a large-scale learner survey (n = 1,537), focus groups and workshops with approximately 90 educators in total, and controlled team-based experiments involving 47 learners with and without access to GenAI tools. The aim was to identify human, social, and organizational factors shaping meaningful human–AI interaction in educational settings. Rather than asking whether GenAI should be used in education, the study focuses on how it can be embedded to foster critical reflection, collaboration, and future-oriented competencies – supporting a shift from automation toward human–AI augmentation.

The research was conducted in cooperation with a diverse range of educational institutions, including a University of Applied Sciences (tertiary education), an Institute for Economic Promotion (adult education), a Tourism School (upper secondary vocational education), and a Business School (upper secondary vocational education). This diversity enabled the study to capture educational contexts ranging from secondary and vocational education to higher and adult education, strengthening the transferability of the findings.

Beyond empirical data collection, the project integrated continuous knowledge transfer and professional development through workshops and targeted training formats. As a structural outcome, a centralized digital platform was established to provide sustained access to research insights, training materials, and curated best practices, supporting long-term institutional learning.

By integrating empirical research, educator engagement, and organizational learning mechanisms, this study contributes an evidence-based perspective on competency-oriented GenAI integration. Drawing on Niklas Luhmann's social systems theory, GenAI adoption is conceptualized as a challenge of organizational programming. While conditional programs ("if–then" rules) struggle to cope with technological uncertainty, purposive programs that define goals – such as fostering critical human–AI collaboration – allow adaptive and contextsensitive implementation (Luhmann, 2018). This programmatic openness is particularly relevant in contexts where GenAI adoption is already widespread but pedagogically unstructured. At the same time, it introduces responsibility: decisions about GenAI integration must be justified by shared institutional purposes rather than individual preference or technological determinism. The present study responds to this need by providing empirical grounding for purposive programming, identifying key dimensions of competency development and organizational learning that can inform adaptive, evidence-based goal-setting in GenAI integration.

2. Theoretical background and literature review

The integration of Generative Artificial Intelligence into education raises fundamental questions that go beyond tool use or instructional techniques. GenAI challenges established assumptions about learning, authorship, assessment, and responsibility, calling for a theoretical perspective that can account for uncertainty, organizational decision-making, and the social conditions of technology use (Jain and Kiran, 2025).

The theoretical background therefore combines insights from educational research on GenAI with Niklas Luhmann's Social Systems Theory. By conceptualizing educational institutions as organizational systems constituted by communication of decisions (Luhmann, 2018), this framework enables an analysis of how GenAI is interpreted and integrated, and how competency development in human–AI collaboration is shaped by organizational structures rather than by technology alone.

Recent research on collaboration technologies provides important insights into why GenAI does not automatically enhance collective performance. Zhang *et al*. (2025) show that technology-supported collaboration increases team creativity only when teams actively integrate distributed knowledge and engage in what the authors term "creative synthesis". Merely adding advanced tools to collaborative settings can even fragment team processes if roles, responsibilities, and shared sense-making remain underdeveloped. These findings resonate with earlier work on team cognition and coordination, emphasizing that technological affordances must be embedded in structured interaction patterns to support collective learning and innovation. Applied to GenAI-supported educational contexts, this suggests that team-based competencies – such as role clarity, shared responsibility, and communicative alignment – are not peripheral but constitutive conditions for meaningful human–AI collaboration.

2.1 Generative AI in educational contexts

Generative artificial intelligence (GenAI) that produces text, code, images, and other media, has rapidly become part of everyday educational practice. Policy and research communities increasingly emphasize that the key challenge is not merely easily available technology, but how education systems govern, integrate, and teach with/through these tools in a human-centered way (UNESCO, 2025).

In higher and vocational education, GenAI is commonly used for (1) drafting and revising academic texts, (2) summarizing readings, (3) brainstorming and outlining, (4) coding support, (5) generating practice questions and explanations, and (6) assisting with research-related tasks such as keyword generation and initial orientation. Research syntheses and position papers consistently describe a mix of learning-supportive applications (e.g. tutoring-like explanations, writing support, feedback generation) and assessment-related uses that blur boundaries between support and substitution (Chan and Colloton, 2024; Kasneci *et al*., 2023).

Survey evidence also suggests that student use is already widespread and often informal: for example, the UK HEPI/Kortext survey (Freeman, 2025) reports very high rates of GenAI use among students and substantial use in assessment contexts.

The literature converges on several recurring risks:

(1) *Authenticity and authorship:* GenAI complicates what it means for work to be "a student's own," especially when prompts, outputs, editing, and disclosure are not transparent. Guidance documents therefore stress clarity about permissible use, disclosure expectations, and alignment with learning outcomes (UNESCO, 2025).

(2) *Plagiarism:* GenAI can generate fluent text that is difficult to evaluate with traditional plagiarism tools; this shifts attention toward assessment design, integrity-by-design approaches, and explicit communication with students (Moorhouse *et al.*, 2023).

(3) *Overreliance:* When GenAI becomes a default shortcut, it may reduce opportunities to practice foundational skills (e.g. argumentation, writing, problem-solving), reinforcing "automation" rather than learning (Bobula, 2024).

(4) *Inaccuracy and hallucinations:* LLMs can produce plausible but false statements and even fabricated references, which is particularly relevant for research and academic writing tasks. This requires explicit instruction in verification practices and source evaluation (Bobula, 2024).

(5) *Governance:* Access to tools, paid features, and guidance differs across student groups and institutions; policy-oriented work highlights governance gaps, privacy concerns, and institutional preparedness as central issues (UNESCO, 2025).

Overall, the emerging consensus is that GenAI's risks and benefits depend strongly on task design, scaffolding, and norms of responsible use rather than on the tools alone (Kasneci *et al.*, 2023).

2.2 Educational institutions as social systems

We draw on Niklas Luhmann's Social Systems Theory to conceptualize educational institutions not as collections of individuals or technical arrangements, but as social systems constituted by communication. This perspective provides a theoretical lens for understanding how educational institutions respond to technological developments such as generative AI without being directly determined by them.

Educational institutions are organization systems. Their core operations revolve around decisions about curricula, assessment formats, or teaching methods. These are therefore not technical responses but communicative decisions that must be compatible with existing educational meanings, norms, and expectations (Luhmann, 2018).

This implies that the introduction of new technologies does not automatically lead to pedagogical change. Instead, technologies are selectively interpreted and integrated through organizational communication – often reinforcing existing practices before enabling transformation.

2.2.1 Irritation and internal adaptation processes. Irritations occur when developments in the environment cannot be ignored by the system but also cannot be directly integrated. GenAI can be understood as such an irritation for educational institutions: it challenges established assumptions about authorship, learning processes, assessment validity, and the role of expertise. Irritation triggers internal reflection, discourse, and adaptation processes within the system. Educational responses to GenAI – ranging from bans and restrictive policies to experimental integration and curriculum redesign – can thus be analyzed as system-specific attempts to restore coherence under new conditions.

From this perspective, GenAI is not an external disruptor that reshapes education from the outside, but a stimulus that provokes internally guided change. This theoretical framing allows the analysis to move beyond deterministic or purely instrumental views of technology and highlights the importance of organizational communication, meaning-making, and decision-making in shaping educational futures.

2.2.2 Organizational programming. Educational institutions observe technological developments translating them into internally meaningful distinctions (e.g. opportunity/risk, innovation/threat, support/substitution) and implementing them into programs. These programs guide decisions about acceptance, regulation, or rejection (Luhmann, 2008).

Therefore, the impact of GenAI in education depends less on the technical capabilities of the tools themselves and more on how educational institutions frame them within their communicative structures – such as assessment rules, academic integrity norms, or competency frameworks.

Recent systems-theoretical work frames technological change as a reconfiguration of decision premises rather than linear adoption. Roth *et al*. (2025) emphasize the role of guiding distinctions as orientation structures that stabilize expectations in complex social environments. From this perspective, organizational change occurs when (guiding) distinctions are reformulated. GenAI integration can thus be understood as a modification of the distinctions that orient decisions about teaching, assessment, and professional responsibility.

Recent strategy research further reinforces the need to conceptualize GenAI adoption beyond operational or tool-centric perspectives. Doshi *et al*. (2025) show that GenAI reshapes how strategic decisions are evaluated, explored, and justified rather than simply improving decision accuracy. While GenAI can support exploratory reasoning and scenario generation, its impact depends strongly on how responsibility, judgment, and accountability are organized within decision processes. Without clear strategic framing, AI-supported decision-making risks decoupling outcomes from human responsibility. This insight aligns closely with Luhmann's distinction between conditional and purposive programming (see section "2.2.3 Conditional vs. Purposive Programs"). In contexts characterized by technological uncertainty, such as GenAI integration, purposive programs that define shared goals while leaving implementation open provide a more robust organizational response than rigid rule-based approaches.

2.2.3 Conditional vs. purposive programs. Within Niklas Luhmann's theory of organizations as social systems, programs function as central decision premises that orient organizational communication by specifying how decisions are to be made (Luhmann, 2018). A distinction is drawn between conditional programs (Konditionalprogramme) and purposive programs (Zweckprogramme), which embody different logics for handling uncertainty and time.

Conditional programs follow an input-oriented, rule-based logic: if a predefined condition occurs, a predetermined response follows ("if X, then Y"). Such programs presuppose a relatively stable environment in which relevant conditions can be anticipated and reliably classified (Luhmann, 2018). They are particularly effective in contexts that require legal certainty, procedural standardization, or risk minimization. In educational institutions, conditional programs typically manifest in assessment regulations, academic integrity policies, or sanctioning mechanisms – e.g. rules specifying penalties if prohibited tools are used in examinations.

Purposive programs, by contrast, are output-oriented and define desired future states while leaving open the means by which these goals may be achieved. Rather than prescribing actions, they articulate objectives that guide decision-making across situations. Luhmann emphasizes that purposive programming necessarily operates under conditions of uncertainty, since future developments cannot be fully known or controlled. As he states: "The main problem of all purposive programming is the unknownness and inaccessibility of all future" (Luhmann, 2018). While purposive programs can be revised in light of experience, they cannot eliminate the fundamental temporal gap between present decisions and future outcomes.

This distinction is particularly salient in the context of generative AI. As GenAI technologies evolve rapidly and their implications remain contested, conditional programming easily becomes outdated or misaligned with pedagogical aims. Rules such as penalizing AI use presuppose stable tool definitions and clear authorship boundaries – assumptions that GenAI fundamentally destabilizes.

Purposive programming therefore offers a more robust response. Extending this logic, Roth and Shedleski (2026) propose the multifunctional tetralemma as a program architecture that moves beyond binary oppositions without dissolving them. Applied to GenAI integration, this implies that minimal conditional routines (e.g. verification norms) can be embedded within broader goal orientations (e.g. critical AI literacy), thereby stabilizing expectations while preserving adaptive flexibility. Purpose-oriented programs (e.g. responsible human–AI collaboration) maintain coherence under uncertainty and relocate responsibility from individual compliance to institutional orientations.

2.3 Competency development in Human-AI collaboration

The distinction between conditional and purposive programming has direct implications for how competency development in AI-supported education is conceptualized. Early discourse on GenAI in education has largely focused on task-level capabilities such as prompt engineering, tool selection, or efficiency gains (Hsain and Housni, 2024). While these skills are not irrelevant, a predominantly task-based perspective risks reducing GenAI integration to technical optimization rather than educational transformation.

Recent research therefore calls for a shift toward competency-based frameworks that reflect the socio-technical nature of human–AI collaboration (Chan and Colloton, 2024; Kasneci *et al.*, 2023). From this perspective, technical AI literacy – understanding how tools function and how to interact with them – represents only one dimension of a broader competency set. Effective GenAI use also requires cognitive, social, and ethical competencies that enable individuals and teams to critically interpret, contextualize, and responsibly apply AI-generated outputs.

Beyond performance outcomes, empirical work highlights important motivational implications of human–GenAI collaboration. Wu *et al*. (2025) show that while generative AI can enhance short-term task performance, it may simultaneously undermine intrinsic motivation by reducing perceived autonomy and competence. When AI systems take over cognitively meaningful tasks rather than supporting reflective engagement, learners experience diminished ownership and motivational involvement. This is particularly relevant in educational contexts, where learning depends not only on output quality but on sustained engagement and competence development. These findings reinforce the importance of framing GenAI as an augmentative rather than substitutive resource.

Accordingly, frameworks for human–AI collaboration emphasize principles of complementarity, augmentation, and orchestration, highlighting factors such as shared goal clarity, role definition, mutual accountability, and open communication (Tarricone and Luca, 2002). GenAI does not replace human judgment but must be actively coordinated within collaborative settings. Such coordination is inherently collective: team communication, shared responsibility, and collective sense-making become key mediating factors in determining whether AI use fosters learning and innovation or merely accelerates output (Zhang *et al.*, 2025).

From a systems-theoretical perspective, this emphasis on collaboration aligns with the understanding of teams as organizational structures that stabilize decision-making under conditions of complexity. Teams function as structural decision premises that distribute responsibility and enable coordinated action. In Luhmann's terms, persons operate as communicative reference points through which roles, expectations, and responsibilities are attributed and negotiated (Luhmann, 2018). In AI-augmented contexts, such role attribution becomes increasingly critical, as decisions require negotiating the division of labor between human actors and technological systems.

This structural dimension has an ethical counterpart. As Laflamme (2025) argues, ethical programs cannot be reduced to compliance rules but depend on guiding distinctions that orient responsible decision-making. In the GenAI context, questions of authorship and accountability must therefore be embedded in communicative routines, shifting responsibility from individual users to organizational program structures.

Consequently, effective GenAI integration requires not less but more human competence. As effort shifts from content production toward coordination, evaluation, and justification, competencies related to critical reflection, ethical reasoning, communication, and teamwork gain prominence. Team-based competencies, in particular, emerge as crucial enabling conditions for productive human–AI collaboration, mediating how AI outputs are interpreted, contested, and integrated into collective work processes (Wu *et al*., 2025; Zhang *et al*., 2025).

3. Methodology

Our study adopts a mixed-methods research design to examine how GenAI is integrated into educational institutions and how competencies in human–AI collaboration are developed. Given the exploratory and practice-oriented nature of the research problem, a single-method approach would be insufficient to capture both the structural conditions shaping the use of GenAI and the lived experiences of educators and learners. The methodology therefore combines a cross-sectional learner survey, focus groups with educators, and controlled group-work experiments comparing teams with and without access to GenAI tools.

Empirical data collection comprised three complementary stages. An institution-based learner survey was conducted between October 2024 and February 2025 across educational institutions of the Vienna Chamber of Commerce and Industry (WKW), covering secondary, vocational, higher, and adult education and involving heterogeneous learner populations. Focus groups with educators were conducted between October and November 2024, while experimental studies took place from December 2024 to February 2025. All participants provided informed consent prior to participation.

3.1 Learner survey

The learner survey was designed as a cross-sectional quantitative study to establish a baseline of GenAI use among learners. The objective was to capture current usage patterns, purposes, frequency, and learner perceptions of GenAI across different educational contexts. The survey followed descriptive and exploratory logic and provided an empirical foundation for the qualitative focus groups and the experimental team-based observations reported in subsequent sections.

Data was collected using an online questionnaire implemented in Microsoft Forms, consisting of 20 questions with an average completion time of approximately 10 min. The instrument was self-developed to reflect the institutional context and research focus. Items covered four thematic areas: (1) prevalence and frequency of GenAI use, (2) typical use cases (e.g. text production, ideation, information structuring), (3) perceived benefits and risks associated with GenAI use in learning contexts, and (4) self-assessed competencies relevant for human-AI collaboration, including the critical evaluation of AI-generated outputs. The questionnaire combined five-point Likert-scale items, multiple-choice questions, and selected open-ended questions.

The final sample comprised 1,577 learners across four institutions within the Vienna Chamber of Commerce and Industry (WKW), representing secondary, vocational, tertiary, and adult education. The overall mean age was 20 years (range 13–66). The sample included participants from both upper secondary and higher education contexts, ensuring coverage of heterogeneous learner populations. Participation was voluntary and based on informed consent. Descriptive analyses were conducted to identify usage patterns and self-assessed competencies across educational levels. No inferential hypothesis testing was conducted. All descriptive analyses were conducted in Microsoft Excel.

The survey instrument and data collection procedure were reviewed and approved by the participating institutions and their directors prior to field deployment. In addition, an internal pretest with university staff members was conducted to assess clarity, comprehensibility, and face validity of the questionnaire.

As a self-report instrument, the survey is subject to common limitations such as social desirability bias and potential overestimation of self-assessed competencies. Self-assessment measures are therefore interpreted as indicators of perceived rather than objectively measured competence. Survey findings are triangulated with focus group data and experimental observations at the interpretation stage to strengthen overall validity.

3.2 Focus groups with educators

The educator focus groups explored current uses of GenAI in teaching practice, educators' attitudes toward GenAI, perceived opportunities and risks, and institutional as well as pedagogical support needs for meaningful integration.

3.2.1 Participants and recruitment. Four focus groups were conducted (one per institution: University of Applied Sciences, Institute for Economic Promotion, Tourism School, and Business School), comprising a total of $N = 23$ educators (14 female, 9 male). Focus groups included educators from mixed subject backgrounds. Recruitment followed institution-specific procedures: participation at the University of Applied Sciences and the Institute for Economic Promotion was coordinated centrally via email and scheduling, whereas the Tourism School and the Business School relied on internal contact persons. At the University of Applied Sciences, the Tourism School, and the Business School, recruitment intentionally prioritized educators with a strong affinity for technology and AI, whereas the Institute for Economic Promotion also included less AI-affine educators to increase heterogeneity ("first come, first served", with waitlist substitution to diversify professional backgrounds).

3.2.2 Procedure and setting. Focus groups were conducted between October and November 2024.

Each session lasted 120 min and followed a standardized structure: introduction and shared definitions; a "status quo" segment (tools used, purposes, most frequently used tools, self-assessed competence, and frequency of use); a split into two sub-groups (A: potentials and future wishes; B: current use and challenges); and a plenary synthesis and wrap-up.

Sessions were primarily held onsite, with limited remote participation when needed (e.g. illness/long travel time). Two sessions used digital collaboration boards (Miro), while two were conducted with analogue materials (Flipchart/post-its) session shifted to flipchart due to technical issues).

3.2.3 Data collection. Discussions were moderated and comprehensively documented through workshop artefacts (digital or physical post-its, flipcharts, and board outputs) and structured field notes captured by an observing researcher. All artefact-based responses were transferred into Word and Excel files to enable aggregation and visualization.

Audio recordings were collected as a contingency measure to resolve potential ambiguities, but the analysis reported here draws on the documented artefacts and notes.

All participants provided informed consent prior to participation, and materials were anonymized for analysis and reporting.

3.2.4 Data analysis and rigor. A thematic content analysis was conducted to identify recurring patterns across institutions in (a) current GenAI use, (b) perceived opportunities and risks, and (c) needs for guidance, infrastructure, and professional development.

In addition, responses to the structured "status quo" questions (items 1–6) were aggregated into positive and negative categories to summarize the most frequently cited benefits and barriers.

Coding and synthesis were led by one researcher and subsequently discussed in a peer debriefing session with a second project team member who had also observed the sessions.

3.3 Experiments: group work with and without AI

The experimental component of the study was conducted across six workshop runs between December 2024 and February 2025 and involved 47 participants from the University of Applied Sciences, the Tourism School, and the Business School. Participants represented a heterogeneous educational spectrum, ranging from secondary school students (approximately 16–18 years) to part-time Bachelor and full-time Master students (approximately 20–35 years). The experiments were designed as practice-based learning settings in which student teams worked on an authentic, domain-relevant problem under controlled yet realistic conditions.

Teams were tasked with analyzing the installation of photovoltaic systems on suitable rooftops in Vienna, addressing technical, economic, social, and legal challenges while developing stakeholder communication strategies and a final concept presentation. The experimental design contrasted two conditions: teams required to integrate Generative AI tools into their collaborative work and teams explicitly prohibited from using GenAI. Following a pragmatic approach, participants were organized into teams of four, with one team consisting of three students due to an absence. Team composition, task structure, and tool availability were aligned with typical educational settings. All sessions followed a standardized procedure consisting of a briefing with an identical problem scenario, a 90-min collaborative work phase, and a 10-min presentation. GenAI teams were provided with laptops equipped with ChatGPT Plus (GPT-4o and o1 models) without access to search engines, while non-GenAI teams used standard browsers (e.g. Google search) without access to GenAI tools. Both conditions had access to standard collaboration software (MS Teams, PowerPoint, and Word).

The study followed a qualitative, theory-informed comparative analysis approach. Multiple data sources were triangulated, including structured observation protocols, participant self-assessment questionnaires, field notes, full-process video recordings, and final team presentations. Data analysis focused on identifying patterns in team processes, decision-making, and competence-related behaviors rather than on measuring individual performance outcomes.

Observational data were analyzed using predefined analytical categories derived from established teamwork research. These categories included shared goal clarity, role definition and leadership, mutual accountability, communication quality, social competence, and strategic task coordination. Rather than applying inductive open coding, the analysis followed a deductive–interpretive logic: empirical observations were systematically mapped against these theoretically informed dimensions to compare how teams operationalized shared goals under GenAI-supported versus non-GenAI conditions.

Self-assessment questionnaires were used descriptively to capture participants' perceptions of collaboration quality and perceived competence development. These data served as a complementary perspective to observational findings, enabling cross-validation between external observation and participant self-reports. Video recordings (made with participants' prior informed consent) and presentation artifacts were used to corroborate observations and to reconstruct critical moments in team interaction, such as decision shifts, coordination breakdowns, or changes in task orientation.

Several measures were taken to enhance validity and reduce observer bias. All sessions were observed by trained facilitators using standardized observation forms to ensure consistency across experimental runs. Observers were briefed on the analytical criteria prior to data collection, and observations were discussed within the research team to reflect on potential interpretive biases. Triangulation across multiple data sources further strengthened internal validity by allowing converging and diverging interpretations to be identified. Given the exploratory and practice-oriented nature of the study, formal interrater reliability statistics were not calculated; instead, methodological rigor was ensured through structured observation, transparent analytical criteria, and systematic cross-checking of empirical materials.

4. Findings

Drawing on survey data from learners, controlled experiments, and focus groups with educators, the findings shed light on how GenAI is currently used, perceived, and negotiated across educational institutions contexts. Rather than offering an exhaustive account of all data sources, the section focuses on those results most relevant to understanding emerging competency requirements, human–AI interaction dynamics, and institutional conditions shaping meaningful AI integration in education.

4.1 Survey results: current use of GenAI by learners

The survey results indicate that generative AI has become an established element of learners' day-to-day educational practice. A substantial share of respondents reported frequent use, with many indicating usage several times per week or daily. Usage intensity was particularly pronounced in higher education contexts, where around three quarters reported using GenAI often (two to three times per week) or very often (daily). In upper secondary education, the corresponding share was around two thirds, indicating routine adoption across educational levels.

4.1.1 Main purposes and reasons for use. Learners' reported use cases are strongly oriented toward text-based tasks and workflow support. GenAI is primarily applied for drafting and revising written work, completing homework assignments, and supporting research-related activities such as summarizing and structuring information. In addition, learners use GenAI for brainstorming and generating initial ideas, particularly in early task phases. When asked about their reasons for using AI tools, time savings emerged as the dominant driver across contexts, followed by the goal of producing more content and improving the quality of output. A smaller share indicated that AI outputs were perceived as better than their own performance, suggesting that learners mainly frame GenAI as a productivity and quality-enhancement resource.

Across both reported use and recommendation, ChatGPT emerged as the dominant GenAI tool, with DeepL particularly salient for translation-related tasks and writing support (DeepL Translate and DeepL Write). Gemini and Claude were rarely used and rarely recommended in upper secondary contexts; in higher education, Gemini showed somewhat higher uptake but remained secondary relative to ChatGPT and DeepL.

Learners' responses further indicate that access to GenAI tools is primarily organized privately and individually rather than through institutionally provided offerings. Paid/premium use differed markedly by educational context. In higher education, around one third reported using paid versions, while around two thirds relied on free access. In upper secondary education, paid usage was substantially lower, at around 10%, indicating the potential for differentiated access to features and functionality depending on subscription status.

4.1.2 Self-assessed GenAI competence and experience. Learners' self-assessments point to substantial perceived competence, but with clear differences across contexts. In higher education, over 50% of respondents classified themselves as advanced users or experts. In upper secondary education, the share reporting advanced or expert competence was just under 50%, pointing to uneven development of perceived mastery across educational segments.

4.1.3 Free-text insights: perceived learning potential and critical concerns. Open-ended responses on whether GenAI can help learners acquire knowledge and skills faster provide a differentiated picture. Positive statements clearly dominated (495 mentions). Learners most frequently highlighted summarizing content and gaining an overview (160 mentions), time savings and efficiency (150 mentions), and simplified explanations (130 mentions). Smaller clusters referred to personalization and interactivity (35 mentions) and additional use cases such as brainstorming, translations, or flashcards (20 mentions).

At the same time, a smaller but meaningful set of critical responses (40 mentions) articulated risks, including the potential loss of independent thinking processes, error risks in AI-generated outputs, and concerns that reliance on AI could foster complacency or reduced effort. Taken together, the findings show high and normalized GenAI use oriented toward efficiency and text work, alongside early signals that reflective validation practices remain uneven and represent a central development need.

4.2 Focus group results: educator practices and perceptions

Educators across institutions reported that GenAI is already embedded in routine professional tasks, with ChatGPT being the most commonly used tool. In some cases, it is used as a pragmatic substitute for conventional web searches. Its use has also extended to translation and language support, and experimentation with specialized tools (e.g. media generation and research support) is particularly prevalent in adult/vocational education. However, use was frequently described as being driven by individuals rather than coordinated by institutions, which contributes to heterogeneous standards and uncertainty about "good practice".

4.2.1 Opportunities framed as pedagogical augmentation. Educators consistently highlighted the potential of GenAI to save time, for example by enabling faster topic entry, drafting and preparation, and accelerating the development of teaching materials for a relatively wide arrange of tasks. Many described GenAI as a resource for creativity and inspiration (e.g. AI acting as a "sparring partner" for lesson design). Beyond productivity, participants pointed to pedagogical benefits such as methodological variety and scaffolding, including the structured use of GenAI by students under supervision.

Another opportunity was presented strategically: since learners are already using GenAI, educators emphasized the importance of AI literacy as a core educational outcome, including the critical evaluation of outputs, verification practices and responsible integration.

4.2.2 Quality of work, epistemic uncertainty, and learner dependency. Concerns focused on the quality and uncertainty of outputs, including unreflective use, generic or misleading results, lack of authenticity/plagiarism, and bias. Educators also expressed concern about a reliance driven by convenience that may hinder long-term learning, for example by reducing effortful practice and writing/argumentation routines, unless reflective use is actively encouraged. Several discussions also raised broader ethical and societal considerations, such as energy consumption and the implications for expertise and work.

4.2.3 Supervision capacity and inequitable access. Implementation challenges were often described as capacity-related, such as supervising AI-supported learning in heterogeneous classrooms while educators were still developing their own skills. The need to fact-check and adapt outputs was considered time-consuming and sometimes offset the anticipated efficiency gains. Access constraints relating to devices, infrastructure and the limited functionality of free tool versions were also highlighted as posing a risk of inequity and inconsistent adoption.

4.2.4 Assessment validity and didactic redesign. Educators from various institutions expressed uncertainty about fair and meaningful assessment in the context of AI-assisted work and repeatedly advocated for instructional redesign. GenAI was presented as emphasizing the need to move away from polished end products towards process-oriented assessment, reflective documentation and explicit instruction in verification and justification.

4.2.5 Governance, infrastructure, and practice-oriented capacity building. The expressed needs were highly consistent: clear legal and data protection guidance and actionable rules, equitable provision of hardware and licenses, as well as practice-oriented professional development, including tool literacy, prompting practices and assessment strategies. Educators also requested structured exchange formats, such as repositories of best practices and communities of practice, as well as administrative relief through AI to free up time for pedagogical work and collaboration.

Institution-specific emphases reflected context: Discussions at the University of Applied Sciences were more polarized and frequently returned to concerns about writing and competence, as well as the need for structured resources. The Business School emphasized the need for clarity in legal and assessment matters within an already digitized environment. The Tourism School foregrounded infrastructure disparities and uncertainty in grading. Finally, Institute for Economic Promotion highlighted paywalls, ethical and societal considerations, and the desire for standardization in projects and final assignments.

Overall, the focus groups indicate that GenAI has become an integral part of educators' daily practice, primarily as a productivity and material development tool. It is also seen as an opportunity to foster students' AI literacy through guided, reflective use. However, adoption is constrained by concerns about output quality, epistemic uncertainty, supervision and verification workloads, and unequal access to tools and infrastructure. Assessment validity emerged as the central pressure point, driving calls for process-oriented didactic redesign, clear governance frameworks, equitable provision of licenses and devices, and practice-oriented professional development supported by structured peer exchange and shared resources.

4.3 Experimental results: group work with and without GenAI

Overall, across the experimental runs, differences in outcomes appeared to be strongly associated with team processes than with access to GenAI tools alone. While GenAI access shaped certain aspects of work, variation in performance was primarily observed along dimensions of collaboration, coordination, and role distribution. Teams that tended to produce more coherent and complete outcomes – regardless of tool condition – often showed early alignment on goals, some degree of task partitioning, emerging role clarity, and balanced participation during both the work phase and the final presentation. Conversely, less developed outcomes were more frequently observed in teams characterized by diffuse roles, uneven participation, or extended discussion phases without convergence on concrete deliverables. These tendencies broadly align with established teamwork dimensions such as shared goal clarity, leadership and role definition, mutual accountability, social competence, open communication, and purposeful team composition (Tarricone and Luca, 2002), which also informed the observational framework.

4.3.1 Tendencies related to GenAI access. 4.3.1.1 Speed and output. Under conditions of structured collaboration and basic AI literacy, GenAI-supported teams progressed more rapidly through early drafting phases and, in several cases, completed full presentations within the allotted time. This speed advantage diminished in teams with weaker coordination or limited familiarity with GenAI, where outcomes were comparable to those of well-organized non-GenAI teams.

4.3.1.2 Creativity. When team functioning was otherwise similar, GenAI access was occasionally associated with more unconventional solution elements. For instance, one team proposed photovoltaic panel shadowing as a building cooling mechanism, extending beyond standard economic or technical approaches. However, such effects were not consistently observed and should be interpreted as indicative rather than systematic.

4.3.1.3 Critical validation. Marked differences emerged in how teams engaged with AI-generated outputs. Some GenAI teams relied on generated content with limited validation, resulting in superficial reasoning or internal inconsistencies. Teams that applied explicit cross-checking strategies produced more coherent results, suggesting that GenAI shifts cognitive effort from content generation to evaluation, with varying success across teams.

4.3.1.4 Motivation and ownership. Reflections revealed ambivalent motivational effects. While some students experienced reduced motivation in non-GenAI teams, others in GenAI-supported teams reported diminished authorship and ownership when AI contributions dominated. This points to a tension between efficiency gains and perceived responsibility in AI-supported collaboration.

4.3.2 Cross-context tendencies. Several tendencies appeared to recur across institutional and educational contexts. First, well-coordinated non-GenAI teams frequently achieved outcomes comparable to those of GenAI-enabled teams. Second, GenAI-enabled teams seemed to benefit most when baseline collaboration structures and shared task understanding were already in place. Third, non-GenAI teams sometimes demonstrated stronger scope control and time discipline, particularly when roles and objectives were clearly articulated. Finally, familiarity within homogeneous cohorts did not necessarily translate into efficiency; in some instances, it coincided with digressions rather than focused collaboration.

Taken together, the results suggest that team processes tend to mediate the effects of GenAI access. GenAI appeared to amplify existing patterns of collaboration rather than compensate for coordination deficits. When teams established shared goals, role clarity, and mutual accountability, GenAI could accelerate drafting and enrich ideas or visual elements. Where such conditions were weak, its potential advantages were less evident.

More broadly, the findings point toward a shift in the focus of challenge from production to selection. Teams that articulated intermediate goals, "done" criteria, time constraints, or validation routines appeared better positioned to integrate GenAI productively. In their absence, GenAI access alone did not systematically improve outcomes. These tendencies support a program-oriented perspective in which purposive goal setting provides the conditions under which GenAI can meaningfully extend available means, while minimal conditional routines (e.g. plausibility checks or source verification) help stabilize quality without over-determining the work process.

5. Discussion

This study provides an empirically grounded perspective on the integration of GenAI in educational contexts, showing that the primary challenges of AI adoption are not technological but human, social, and organizational. Across all empirical components, a consistent pattern emerges: while GenAI is already widely used in educational practice, its integration remains weakly institutionalized and pedagogically underdetermined.

Learners report frequent and confident use of GenAI, particularly for research, text production, and project-based work. However, this use is largely informal, dependent on freely available tools, and minimally guided by institutions. As a result, new inequalities arise in access, quality of use, and competency development, alongside unresolved questions concerning assessment, ethics, and learning outcomes. Educators, by contrast, adopt GenAI more cautiously and selectively, producing a structural asymmetry that institutions must actively address.

Experimental findings highlight the limits of technological substitution. While GenAI accelerates drafting and elaboration processes, it does not compensate for deficits in collaboration, coordination, or shared understanding. Instead, AI amplifies existing team dynamics: strong teams perform well regardless of AI access, whereas weak teams do not improve through technological support. These results challenge techno-solutionist narratives and suggest that GenAI shifts the main bottleneck in complex tasks from individual cognition to social organization and collective sense-making. Consequently, competencies related to communication, role clarity, leadership, and coordination become more, not less, relevant.

Educators focus groups further contextualize these findings. Participants recognize AI's potential for efficiency gains and personalization but express uncertainty regarding didactic integration, legal frameworks, and assessment practices. The absence of shared institutional guidelines, recommended tools, and exemplars emerges as a key barrier to systematic and reflective AI use. There is broad consensus that GenAI should be understood as augmentation technology rather than a substitute for pedagogical judgment.

Overall, the findings indicate that meaningful GenAI integration requires a shift from individual, tool-centered adoption toward institutionally anchored, competency-oriented strategies. The development of the KI-Info-Hub platform responds to this need by providing access to research insights, curated resources, and didactic examples. Theoretically, this supports conceptualizing AI integration as a challenge of organizational programming rather than individual usage decisions.

Several limitations must be acknowledged: (1) The learner survey relies on self-reported data, particularly regarding AI usage frequency and perceived competency levels. High self-assessments of AI expertise among learners may therefore reflect confidence rather than actual proficiency; (2) The experimental component was limited in scale and task scope, constraining generalizability: (3) Finally, the focus groups with educators primarily captured reflective and discursive perspectives rather than observed teaching practices. Although these insights are valuable for understanding perceived barriers and needs, future research could benefit from, e.g. classroom-based observations. Addressing these limitations in future research would strengthen the empirical foundation for sustainable, competency-oriented approaches to GenAI integration in education.

6. Conclusion

This study addresses the question *of how Generative AI can be integrated into educational and professional development frameworks to enhance competency acquisition rather than merely accelerating output, and which core competencies are essential for effective human–AI collaboration*. The findings demonstrate that successful GenAI integration is not achieved through increased efficiency or unrestricted tool use, but through deliberate organizational and pedagogical structuring that aligns AI use with learning objectives and professional competencies.

Empirical evidence shows that GenAI is already embedded in learners' everyday practices; however, without institutional guidance it tends to amplify unreflective work patterns rather than support sustainable learning. The study identifies six critical success factors that determine whether GenAI contributes to competency development: clear communication and role definition, shared responsibility and team cohesion, prior GenAI experience, critical reflection on AI-generated content, intrinsic motivation, and the capacity for innovative thinking. Strong teamwork consistently led to superior outcomes irrespective of AI access, while unstructured AI adoption weakened performance. These results indicate that GenAI does not substitute human competencies such as collaboration, judgment, or leadership; instead, it intensifies the need for them.

From an organizational change perspective, the integration of GenAI requires purposive programming: stable educational goals must be defined, while the means of achieving them remain flexible and adaptive under technological uncertainty. Effective GenAI integration therefore depends on institutional frameworks that provide orientation, foster reflective practice, and support competency development across roles. In response to these insights, the project informed the establishment of the KI-Info-Hub, a centralized platform offering educators access to training materials, research findings, and evidence-based practices.

Overall, this study advocates a shift from efficiency-driven automation narratives toward a human-centered augmentation paradigm. When embedded in coherent organizational strategies and guided by shared purposes, GenAI can enhance educational quality, innovation, and professional capability, preparing learners for meaningful and responsible participation in AI-supported work and learning environments.

Ethics statement

This study was conducted within regular higher education settings and did not require formal ethics approval according to institutional guidelines. Participation was voluntary and based on prior informed consent. All data were collected and analyzed in anonymized form. No interventions beyond standard educational practice were carried out, and no foreseeable risks for participants were involved.

AI usage disclosure

The authors used ChatGPT (OpenAI) solely for language and grammar refinement of the manuscript. No content was generated by the AI tool. The authors remain fully responsible for the accuracy and integrity of the work.

References

Bobula, M. (2024), "Generative artificial intelligence (AI) in higher education: a comprehensive review of challenges, opportunities, and implications", *Journal of Learning Development in Higher Education*, No. 30, doi: 10.47408/jldhe.vi30.1137.

Chan, C.K.Y. and Colloton, T. (2024), *Generative AI in Higher Education: The ChatGPT Effect*, New York, Taylor & Francis. available at: https://library.oapen.org/bitstream/id/1f55f220-814d-4c52-a1c7-5e63546413f5/9781040024270.pdf

Dikilitaş, K., Furenes Klippen, M.I. and Keles, S. (2024), "A systematic rapid review of empirical research on students' use of ChatGPT in higher education", *Nordic Journal of Systematic Reviews in Education*, Vol. 2, doi: 10.23865/njsre.v2.6227.

Doshi, A.R., Bell, J.J., Mirzayev, E. and Vanneste, B.S. (2025), "Generative artificial intelligence and evaluating strategic decisions", *Strategic Management Journal*, Vol. 46 No. 3, pp. 583-610, doi: 10.1002/smj.3677.

Elkhodr, M., Gide, E., Wu, R. and Darwish, O. (2023), "ICT students' perceptions towards ChatGPT: an experimental reflective lab analysis", *STEM Education*, Vol. 3 No. 2, pp. 70-88, doi: 10.3934/steme.2023006.

Freeman, J. (2025), "Student generative AI survey 2025 – HEPI", available at: https://www.hepi.ac.uk/reports/student-generative-ai-survey-2025/ (accessed 3 January 2026).

Hsain, A. and Housni, H.E. (2024), "Large language model-powered chatbots for internationalizing student support in higher education", [Online], available at: https://arxiv.org/pdf/2403.14702

Hui, X., Reshef, O. and Zhou, L. (2023), "The short-term effects of generative artificial intelligence on employment: evidence from an online labor market".

Jain, N. and Kiran, M. (2025), "Rethinking education in the age of generative AI: cognitive offloading, assessment reform, and institutional adaptation".

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J. and Kasneci, G. (2023), "ChatGPT for good? On opportunities and challenges of large language models for education", *Learning and Individual Differences*, Vol. 103, 102274, doi: 10.1016/j.lindif.2023.102274.

Krupp, L., Steinert, S., Kiefer-Emmanouilidis, M., Avila, K.E., Lukowicz, P., Kuhn, J., Küchemann, S. and Karolus, J. (2023), "Unreflected acceptance – investigating the negative consequences of ChatGPT-assisted problem solving in physics education", [Online], available at: https://arxiv.org/pdf/2309.03087

Laflamme, D. (2025), "Be warned: Ethical programmes are open to changes, and the advent of artificial moral agents could change what counts as ethics [Conference presentation]", *Luhmann Conference 2025*, Cambridge, UK.

Luhmann, N. (2008), *Ökologische Kommunikation: Kann die moderne Gesellschaft sich auf ökologische Gefährdungen einstellen?*, 5th ed., Verlag für Sozialwissenschaften, Wiesbaden.

Luhmann, N. (2018), *Organization and Decision*, Cambridge University Press, Cambridge.

Moorhouse, B.L., Yeo, M.A. and Wan, Y. (2023), "Generative AI tools and assessment: guidelines of the world's top-ranking universities", *Computers and Education Open*, Vol. 5, 100151, doi: 10.1016/j.caeo.2023.100151.

Reddy, K.S., Gotur, R. and Bhat, V. (2025), "Generative AI adoption in enterprise: a comprehensive case study analysis of implementation strategies and outcomes across diverse sectors", *2025 6th International Conference on Recent Advances in Information Technology (RAIT)*, Dhanbad, India, 06.03.2025–08.03.2025, IEEE, pp. 1-6.

Roth, S. and Shedleski, B. (2026), "Multifunctional tetralemma: a systems-theoretical programme and its practical implications", *Journal of Organizational Change Management*, Vol. 39 No. 1, pp. 195-204, doi: 10.1108/jocm-10-2025-0928.

Roth, S., Watson, S., Dahms, H.F. and Atanesyan, A. (2025), "Guiding distinctions of social theory. Analogue guidelines or digital transformers? An introduction", *Current Sociology*, Vol. 73 No. 4, pp. 477-492, doi: 10.1177/00113921251341658.

Scerri, D. (2024), "Skilling for the future: enhancing vocational learning and workplace productivity with creative AI tools", *MCAST Journal of Applied Research and Practice*, Vol. 8 No. 1, pp. 150-178, doi: 10.5604/01.3001.0054.5099.

Tarricone, P. and Luca, J. (2002), "Successful teamwork: a case study", Research Outputs Pre 2011 [Online], available at: https://ro.ecu.edu.au/ecuworks/4008

UNESCO (2025), "Guidance for generative AI in education and research", 14 April, available at: httpswww.unesco.org/en/articles/guidance-generative-ai-education-and-research (accessed 3 January 2026).

Wu, S., Liu, Y., Ruan, M., Chen, S. and Xie, X.-Y. (2025), "Human-generative AI collaboration enhances task performance but undermines human's intrinsic motivation", *Scientific Reports*, Vol. 15 No. 1, 15105, doi: 10.1038/s41598-025-98385-2.

Zhang, X., Fang, Y., Zhou, J. and Lim, K.H. (2025), "How collaboration technology use affects IT project team creativity: integrating team knowledge and creative synthesis perspectives", *MIS Quarterly*, Vol. 49 No. 2, pp. 611-642, doi: 10.25300/misq/2024/16651.

Corresponding author
Tilia Stingl de Vasconcelos Guedes can be contacted at: tilia1@yahoo.com