Research Square

Preprints are preliminary reports that have not undergone peer review. They should not be considered conclusive, used to inform clinical practice, or referenced by the media as validated information.

Multimodal Generative AI Chatbots for Enhancing Scientific Research Skills: A Self-Determination Theory-Based Mediation Model

Sami AlMazrouii

sami.m@aou.edu.om

Arab Open University of Oman

Research Article

Keywords: AI Chatbots, Video Summaries, Scientific Research Skills t, Information and Summary Quality, Cognitive Load, Self-Determination Theory, Infographics

Posted Date: October 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10939304/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

This study examined the effectiveness of generative AI-enhanced chatbots incorporating video, infographic, and audio summaries in developing scientific research skills among teacher qualification students enrolled in a Research Methods course in the Sultanate of Oman. Drawing on Self-Determination Theory (SDT), it also investigated whether autonomy, competence, and relatedness mediated the relationships between perceived information and summary quality, cognitive load reduction, and scientific research skills. A quantitative posttest-only quasi-experimental design was employed with 308 students assigned to video (n = 98), infographic (n = 104), and audio (n = 106) conditions. Data were analyzed using group comparisons and partial least squares structural equation modeling (PLS-SEM). Significant modality effects emerged across all study outcomes, with the video condition producing the highest scientific research skills scores, followed by infographic and audio conditions. Information and summary quality and cognitive load reduction significantly predicted autonomy, competence, and relatedness, which in turn significantly predicted scientific research skills; competence showed the strongest association with performance. All six specific indirect effects were significant, supporting complementary partial mediation through basic psychological needs. The full model explained 52.8% of the variance in scientific research skills. Experimental mediation analysis further showed that video summaries improved performance both directly and through the proposed SDT-based mechanism, whereas the infographic advantage over audio was primarily direct. These findings demonstrate the value of multimodal generative AI chatbots, particularly video-enhanced designs, for cognitively efficient and motivationally supportive research-skills instruction.

1. INTRODUCTION

In recent years, the rapid advancement of Artificial Intelligence (AI) has significantly transformed educational practices, particularly in higher education. AI-driven technologies, including intelligent tutoring systems and conversational agents, have enhanced student engagement, learning efficiency, and academic achievement. Among these technologies, AI-powered chatbots have emerged as effective tools due to their ability to provide personalized, real-time feedback and interactive learning experiences (Huang et al., 2025).

The increasing availability of video-based educational content has created both opportunities and challenges for learners. While videos offer rich and engaging learning experiences, navigating lengthy instructional materials can be cognitively demanding and time-consuming. Video summarization techniques have been introduced to address this issue by condensing extensive content into concise and meaningful representations, thereby improving learning efficiency and accessibility (Yadav et al., 2024). Moreover, integrating video summarization with AI chatbots enables learners to interact dynamically with content, facilitating targeted information retrieval and deeper understanding (Reddy et al., 2025).

In addition to video-based learning, visual tools such as infographics play a crucial role in enhancing comprehension and knowledge retention. Infographics provide structured visual representations of information, making complex data easier to understand and recall (Bhat & Alyahya, 2024). Research has shown that both video-based and infographic-based instructional formats can significantly improve learning outcomes, although their effectiveness may vary depending on learner preferences and contexts (Ebrahimabadi et al., 2018).

Recent studies have emphasized the effectiveness of combining AI chatbots with summarization strategies. For instance, AI-based chatbots integrated with graphical summarization have been found to enhance students' academic achievement, motivation, self-efficacy, and cognitive load management compared to traditional teaching methods. Similarly, generative AI-enabled chatbots that incorporate summarization and interactive quizzing have demonstrated significant improvements in learners' engagement and knowledge retention by simplifying complex information and promoting active learning (Huang et al., 2025).

From a theoretical perspective, Self-Determination Theory highlights the importance of autonomy, competence, and relatedness in fostering intrinsic motivation, academic performance and relatedness in motivating learner engagement (Ellikkal & Rajamohan, 2024, Deci & Ryan, 2000). Additionally, Cognitive Load Theory emphasizes the necessity of minimizing unnecessary mental effort to optimize learning processes (Sweller et al., 2011) as AI-based chatbots based on video summaries and interactive infographics -generated content can reduce extraneous cognitive load while amplifying Germane processing (Yuan et al., 2025). Integrating these theoretical frameworks with AI-driven educational tools provides a comprehensive understanding of how technological, cognitive, and motivational factors interact to influence learning outcomes.

Despite these advancements, literature remains fragmented, as most studies have examined chatbots, video summarization, and infographics independently rather than within an integrated framework. Furthermore, there is limited research exploring these technologies in postgraduate teacher education contexts, particularly in the Arab region. Critical psychological constructs such as academic self-efficacy and cognitive load reduction have also received insufficient attention in such integrated learning environments and learning outcomes in the context of Oman higher education. Most existing research is either conceptual or situated in Western or East Asian educational systems, lacking relevance to the unique socio-cultural and technological landscape of Oman.

Despite these advancements, several gaps remain in literature. First, limited studies have explored the combined impact of generative video summaries and chatbots on higher-order skills such as innovation and research competencies. Second, most research focuses on general learning outcomes rather than specialized domains like scientific research skills. Third, there is a lack of empirical studies examining the long-term effectiveness of such systems in higher education contexts.

The study examines the effectiveness of Generative AI-enhanced chatbots with multimodal summaries—video, infographics, and audio—in developing scientific research skills among teacher qualification

students in Oman. It also investigates whether these effects are mediated by autonomy, competence, and relatedness within the framework of Self-Determination Theory (SDT), while considering information quality and cognitive load reduction as key features of the AI-supported learning environment.

2.LITERATURE REVIEW

Generative AI-enhanced chatbots with multimodal summaries can support scientific research by improving information and summary quality and reducing cognitive load, thereby facilitating more efficient and focused learning. In this study, these features represent the independent variable and are expected to enhance scientific research skills, the dependent variable, through the mediation of basic psychological needs derived from Self-Determination Theory (SDT): autonomy, competence, and relatedness. Accordingly, the study examines how multimodal AI-supported learning environments contribute to research-skill development both directly and indirectly through the satisfaction of learners' psychological needs.

2.1 Generative AI-Enhanced Chatbots in Scientific Research Education

Generative artificial intelligence (GenAI) has substantially expanded the educational capabilities of chatbots, transforming them from rule-based conversational systems into adaptive learning environments capable of generating explanations, feedback, summaries, examples, and multimodal instructional resources. Large language model–based chatbots can respond dynamically to learners' queries, personalize explanations, and represent information through textual, visual, auditory, and video-based formats (Chiu, 2024). These affordances are particularly relevant to higher education, where GenAI is increasingly used for academic writing, information searching, idea generation, and other activities associated with academic inquiry (Ahmed & Sultan, 2026).

Within scientific research, AI-enhanced chatbots can support multiple stages of the research process, including literature exploration, information synthesis, academic writing, reference identification, data interpretation, and the organization of research ideas (Jee, 2023). Large language models may also support qualitative research processes through functions such as coding, thematic extraction, and data summarization (Barros et al., 2024). Accordingly, GenAI chatbots have the potential to reduce procedural demands and provide learners with timely scaffolding during complex research activities.

Nevertheless, technological efficiency should not be equated with the development of scientific research competence. Scientific research skills require learners to formulate research problems, identify and evaluate relevant evidence, synthesize literature, understand methodological decisions, interpret findings, and communicate evidence-based conclusions. These processes depend on critical reasoning and independent judgment rather than information retrieval alone. Excessive dependence on AI may therefore undermine critical evaluation, originality, and academic integrity when generated outputs are accepted without verification (Stokel-Walker & Van Noorden, 2023).

2.2 GenAI-Enhanced Multimodal Chatbots for Scientific Research Skills

Generative artificial intelligence (GenAI) has expanded educational chatbots from predefined conversational systems into adaptive learning environments capable of generating explanations, feedback, summaries, and multimodal instructional resources. Such systems can dynamically respond to learners' queries, personalize explanations, and present information in multiple representational forms, including text, images, video, and audio (Chiu, 2024). In higher education, these capabilities are increasingly relevant to complex academic activities such as information searching, academic writing, knowledge construction, and research-related tasks. However, concerns remain regarding inaccurate responses, overreliance on AI, and insufficient critical evaluation of AI-generated content.

These issues are particularly important for developing scientific research skills, which require learners to formulate research questions, locate and evaluate evidence, distinguish relevant information, synthesize findings, interpret evidence, and communicate conclusions academically. GenAI-enhanced chatbots may scaffold these processes by explaining unfamiliar concepts, organizing information, providing immediate feedback, and condensing complex academic materials. Nevertheless, their instructional value depends on supporting active cognitive engagement rather than replacing scholarly reasoning. Thus, effective AI-supported learning should reduce unnecessary cognitive burden while maintaining learners' active reasoning, problem solving, and intellectual autonomy (Jose et al., 2025).

A particularly important instructional affordance of GenAI is its ability to transform extensive academic information into concise multimodal summaries, including video, infographics, and audio. Although these modalities can convey the same research content, they differ substantially in how information is organized and cognitively processed. Multimedia learning principles suggest that instructional effectiveness depends not merely on the presence of multiple media, but on how effectively information is selected, organized, and integrated across limited visual and auditory processing channels (Mayer, 2021).

Video summaries integrate visual illustrations, narration, textual cues, demonstrations, and temporal sequencing, making them particularly suitable for explaining research procedures, methodological stages, and relationships among variables. Advances in AI-based video summarization enable systems to identify important segments, preserve semantic relationships, and generate adaptive or query-based summaries (De Silva et al., 2025; Kim et al., 2025; Liang et al., 2021). When combined with chatbot interaction, video summaries can transform passive viewing into interactive exploration through questioning, clarification, and personalized content navigation (Mahapatra et al., 2018; Reddy et al., 2025). However, poorly coordinated multimedia elements may increase rather than reduce cognitive load.

Infographic summaries provide a more spatially organized representation of concepts, comparisons, sequences, methodological stages, and relationships. Their integration of visual and textual information

can facilitate comprehension, retention, information organization, and deeper reflection (Heimbürger et al., 2020; Lyra et al., 2016). Embedded within chatbot environments, infographics can provide visual scaffolding while conversational interaction enables learners to request further explanations. Their effectiveness, however, depends on visual simplicity and informational coherence, as overly dense representations may impose excessive processing demands.

Audio summaries provide flexible access to academic content and may support motivation, engagement, review, clarification, and independent learning (Díaz & Carmona-Galindo, 2026; Jafarian & Kramer, 2025; Sturgeon Delia, 2025). AI-generated podcasts have also demonstrated potential for strengthening research literacy and communicating scientific findings in accessible forms (Desmedt et al., 2025; Hamilton et al., 2026). Nevertheless, because spoken information is transient, its effectiveness may depend strongly on pacing, prior knowledge, content complexity, and playback control.

Collectively, these modality-specific characteristics indicate that video, infographic, and audio summaries should not be assumed to produce equivalent learning outcomes. Importantly, the reviewed literature provides limited direct experimental comparison of these three GenAI-generated summary modalities for developing scientific research skills, representing a substantive empirical gap addressed by the present study.

2.3 Multimodal Summary Quality and Cognitive Load Reduction

Cognitive Load Theory (CLT) provides a theoretical basis for conceptualizing information and summary quality and cognitive load reduction as key attributes of generative AI-enhanced chatbots with multimodal summaries. CLT assumes that novel information is processed through a working-memory system with limited capacity and duration; consequently, learning is constrained when instructional materials impose unnecessary demands for knowledge construction (Sweller et al., 2011; Sweller et al., 2019). Effective multimodal design should therefore minimize extraneous processing and direct cognitive resources toward information essential for schema acquisition and meaningful learning.

This principle is especially relevant to scientific research education, where learners must integrate interacting elements, including research problems, variables, theoretical frameworks, sampling, instrumentation, analysis, interpretation, and academic reporting. Generative AI-enhanced chatbots can reduce these demands by transforming complex content into concise video, infographic, or audio summaries that organize key concepts, clarify relationships, and reduce unnecessary searching. Evidence from AI-supported learning indicates that ChatGPT can facilitate information access and break complex material into more manageable components, thereby supporting cognitive load management (Patac & Patac, 2025). However, unexpected or poorly aligned AI responses may increase mental effort because learners must clarify, verify, or reconcile generated information.

Accordingly, cognitive efficiency depends not only on modality but also on summary quality. Fabbri et al. (2021) identify four central dimensions of summarization quality: coherence, consistency, fluency, and

relevance. In this context, high-quality multimodal summaries should be logically organized, factually aligned with source information, clearly communicated, and restricted to essential content. These criteria are particularly important because LLM-generated summaries may contain factual inconsistencies, while existing evaluation approaches can struggle to detect them reliably (Luo et al., 2024).

Thus, modality effects should not be interpreted independently of summary quality. Video, infographic, and audio conditions should maintain comparable informational accuracy, relevance, and completeness so that differences in scientific research skills reflect modality affordances rather than uncontrolled content variation. In this study, information and summary quality and cognitive load reduction are therefore treated as complementary attributes of the independent variable that explain how GenAI-enhanced multimodal summaries create cognitively efficient conditions for developing scientific research skills.

2.4 Basic Psychological Needs as a Mediating Mechanism in GenAI-Enhanced Learning

Self-Determination Theory (SDT) provides the theoretical basis for explaining how multimodal GenAI-enhanced learning may influence scientific research skills through the satisfaction of three basic psychological needs: autonomy, competence, and relatedness. Within Basic Psychological Need Theory, these needs are regarded as essential psychological resources for adaptive motivation, engagement, and functioning across educational and cultural contexts (Chen et al., 2015; Guay, 2022; Vansteenkiste et al., 2020).

Autonomy reflects learners' experience of volition and meaningful choice. In GenAI-supported learning, autonomy may be strengthened when students can determine the timing, pace, sequence, and form of interaction, formulate personalized questions, and revisit explanations as needed. Competence concerns feeling capable and effective in mastering challenging tasks. GenAI chatbots can potentially support competence through immediate feedback, individualized explanations, and scaffolded learning activities; however, inaccurate or poorly structured AI-generated content may undermine this experience (Chiu, 2024). Relatedness refers to feeling connected and meaningfully supported. Although chatbots can provide responsive interaction, AI-mediated relatedness should be interpreted cautiously because human teacher and peer interaction remains important in technology-enhanced learning (Li et al., 2025).

Accordingly, psychological need satisfaction is conceptualized as a mediating mechanism rather than a direct technological outcome. Li and Chiu (2025) found that need satisfaction mediated relationships between instructional support and multiple dimensions of engagement in GenAI-chatbot learning. Similarly, Ahmed and Sultan (2026) showed that autonomy, competence, and relatedness contribute to motivational pathways within GenAI-supported environments. Extending this logic to multimodal GenAI summaries, video, infographic, and audio formats may generate different perceptions of learner control, clarity, accessibility, and support, thereby differentially satisfy psychological needs and shape engagement with complex research tasks.

2.5 Scientific Research Skills as the Learning Outcome

Educational GenAI research has largely emphasized adoption, continuous-use motivation, language learning, and academic writing, while integrated scientific research competence remains comparatively underexplored (Annamalai et al., 2025). Scientific research skills represent complex, integrative learning outcomes involving problem and question formulation, literature searching and review, methodological decision-making, data analysis, interpretation, and evidence-based scientific writing (Vígh, 2024; Rjoub et al., 2025). Accordingly, GenAI-enhanced learning should be evaluated through learners' performance on authentic, integrative research tasks rather than technology perceptions alone (Maddens et al., 2023). Multimodal summaries may scaffold information processing and comprehension, but their educational value depends on supporting—rather than replacing—independent reasoning. Within an SDT framework, environments supporting autonomy, competence, and relatedness may provide motivational conditions conducive to stronger research-skill development (Maddens et al., 2023).

2.6 Research Gap in the Current Study:

Existing literature supports the educational potential of GenAI-enhanced chatbots, yet three interrelated gaps remain. First, prior studies emphasize engagement, adoption, language learning, and academic writing, with limited attention to scientific research skills as higher-order outcomes. Second, the comparative effects of GenAI-generated video, infographic, and audio summaries remain unclear when content and summary quality are controlled, particularly regarding information quality and cognitive load reduction. Third, although Self-Determination Theory highlights autonomy, competence, and relatedness in AI-supported learning, limited evidence explains whether these needs mediate the effects of multimodal chatbot design on research-skill development. Accordingly, this study compares summary modalities and tests whether psychological need satisfaction explains their effects on scientific research skills, integrating technological, cognitive, and motivational mechanisms within an explanatory model.

2.7 Research Hypotheses

Based on the proposed conceptual model, Generative AI-enhanced chatbots with multimodal summaries, particularly in terms of information and summary quality and cognitive load reduction, are expected to influence scientific research skills both directly and indirectly through the satisfaction of the three basic psychological needs—autonomy, competence, and relatedness—based on Self-Determination Theory (SDT). Accordingly, the following hypotheses are proposed.

Effects of Multimodal Summary Format

H1. There are significant differences among students exposed to generative AI-enhanced chatbots with video, infographic, and audio summaries in their scientific research skills.

H2. There are significant differences among students exposed to generative AI-enhanced chatbots with video, infographic, and audio summaries in their perceived information and summary quality.

H3. There are significant differences among students exposed to generative AI-enhanced chatbots with video, infographic, and audio summaries in their perceived cognitive load reduction.

H4. There are significant differences among students exposed to generative AI-enhanced chatbots with video, infographic, and audio summaries in their satisfaction of basic psychological needs.

H4a. There are significant differences among the three experimental groups in perceived autonomy.

H4b. There are significant differences among the three experimental groups in perceived competence.

H4c. There are significant differences among the three experimental groups in perceived relatedness.

Effects of Information and Summary Quality on Basic Psychological Needs

H5. Information and summary quality has a significant positive effect on students' satisfaction of basic psychological needs.

H5a. Information and summary quality has a significant positive effect on students' autonomy.

H5b. Information and summary quality has a significant positive effect on students' competence.

H5c. Information and summary quality has a significant positive effect on students' relatedness.

Effects of Cognitive Load Reduction on Basic Psychological Needs

H6. Cognitive load reduction has a significant positive effect on students' satisfaction of basic psychological needs.

H6a. Cognitive load reduction has a significant positive effect on students' autonomy.

H6b. Cognitive load reduction has a significant positive effect on students' competence.

H6c. Cognitive load reduction has a significant positive effect on students' relatedness.

Effects of Basic Psychological Needs on Scientific Research Skills

H7. Satisfaction with basic psychological needs has a significant positive effect on students' scientific research skills.

H7a. Autonomy has a significant positive effect on students' scientific research skills.

H7b. Competence has a significant positive effect on students' scientific research skills.

H7c. Relatedness has a significant positive effect on students' scientific research skills.

Direct Effects on Scientific Research Skills

H8. Information and summary quality has a significant positive effect on students' scientific research skills.

H9. Cognitive load reduction has a significant positive effect on students' scientific research skills.

Mediating Effects of Basic Psychological Needs

H10. Satisfaction of basic psychological needs mediates the relationship between information and summary quality and students' scientific research skills.

H10a. Autonomy mediates the relationship between information and summary quality and students' scientific research skills.

H10b. Competence mediates the relationship between information and summary quality and students' scientific research skills.

H10c. Relatedness mediates the relationship between information and summary quality and students' scientific research skills.

H11. Satisfaction of basic psychological needs mediates the relationship between cognitive load reduction and students' scientific research skills.

H11a. Autonomy mediates the relationship between cognitive load reduction and students' scientific research skills.

H11b. Competence mediates the relationship between cognitive load reduction and students' scientific research skills.

H11c. Relatedness mediates the relationship between cognitive load reduction and students' scientific research skills.

Overall Mediation Hypothesis

H12. Basic psychological needs satisfaction mediates the relationship between the perceived characteristics of generative AI-enhanced chatbot summaries, namely, information and summary quality and cognitive load reduction—and students' scientific research skills.

4. Methodology

4.1 Research Design

This study employed a quantitative quasi-experimental research design to investigate the effects of Generative AI-enhanced chatbots with multimodal summaries on students' scientific research skills. Three types of AI-generated multimodal summaries were examined: video summaries, infographic

summaries, and audio summaries. The study further adopted a mediation model grounded in Self-Determination Theory (SDT) to examine whether the satisfaction of the three basic psychological needs —autonomy, competence, and relatedness—mediated the relationship between students' perceptions of the GenAI-enhanced chatbot environment and their scientific research skills.

The conceptual model incorporated Generative AI-enhanced chatbots with multimodal summaries as the independent variable, represented by two dimensions: information and summary quality and cognitive load reduction. The three basic psychological needs of autonomy, competence, and relatedness were specified as mediating variables, whereas scientific research skills constituted the dependent variable.

The design enabled the study to examine both the comparative effectiveness of the three multimodal summary conditions and the direct and indirect relationships proposed in the conceptual model. Accordingly, the analysis combined between-group comparisons with structural equation modeling to assess the hypothesized mediation mechanisms.

4.2 Participants:

The study population consisted of students enrolled in the Educational Qualification Diploma Program in the Sultanate of Oman. The final sample comprised 308 students drawn from six academic sections. The participants were assigned to three experimental groups based on the type of multimodal summary integrated into the Generative AI-enhanced chatbot environment.

Specifically, Group 1 (n = 98) used a Generative AI-enhanced chatbot supported by video summaries, Group 2 (n = 104) used a chatbot supported by infographic summaries, and Group 3 (n = 106) used a chatbot supported by audio summaries. This distribution resulted in three experimental groups with relatively comparable sample sizes, allowing for the examination of differences in learning outcomes across the three multimodal chatbot conditions.

Participants were selected from the Educational Qualification Diploma Program because scientific research competencies constitute a key component of their academic preparation. As part of the program, students are enrolled in a Research Methods course, which requires systematic engagement with a range of research-related activities, including identifying research problems, formulating research questions, reviewing relevant literature, selecting appropriate research methodologies, interpreting research evidence, and preparing academic research outputs. Accordingly, this group provided an appropriate context for examining the development of scientific research skills within a Generative AI-enhanced learning environment.

The sample included students from different age groups, academic backgrounds, and demographic characteristics. Data was collected during the second semester of the 2025–2026 academic year, and a total of 300 valid responses were obtained and included in the final analysis. The study was geographically limited to the Arab Open University in the Sultanate of Oman.

Table 1
Participants

Demographic Variable	Category	Frequency	Percent (%)
Gender	Male	65	21.1
	Female	243	78.9
	Total	308	100.0
Age	20–29 years	277	85.3
	30–39 years	31	14.7
	Total	308	100.0
Qualifications	Bachelor's degree	308	100.0
Type of Institution	Government University	116	37.7
	Private University	192	62.3
	Total	308	100.0
Nationality	Omani	308	100.0
	Total	308	100.0
Experimental Group	Video summaries	98	31.8
	Infographic summaries	104	33.8
	Audio summaries	106	34.4
	Total	308	100.0

The final sample comprised 308 participants, predominantly female (n = 243, 78.9%) and aged 20–29 years (n = 237, 76.9%). All participants held a bachelor's degree and were Omani nationals, while 62.3% were affiliated with private universities and 37.7% with government universities. Participants were distributed across the video (n = 98, 31.8%), infographic (n = 104, 33.8%), and audio (n = 106, 34.4%) experimental groups, resulting in relatively comparable group sizes

4.3 Study Context

The study was conducted within the context of the Educational Qualification Program, where students undertake academic activities requiring the development and application of scientific research skills. The increasing integration of Generative AI into higher education provides opportunities to support these skills through interactive learning environments capable of generating, restructuring, and presenting academic information in multiple formats.

In the present study, a Generative AI-enhanced chatbot was used to provide learning support through three forms of multimodal summaries: video, infographics, and audio. Although the three experimental

conditions addressed comparable research-related learning content, the mode through which summaries were presented differed across groups.

The intervention was designed on the assumption that effective GenAI-supported learning does not depend solely on access to generated information. Rather, its educational effectiveness may depend on the perceived quality of the information and summaries, the extent to which the learning environment reduces unnecessary cognitive load, and its capacity to support students' psychological needs for autonomy, competence, and relatedness.

Within this framework, the study examined whether these characteristics contributed to the development of scientific research skills and whether basic psychological need satisfaction provided a psychological mechanism explaining these relationships.

4.4 Experimental Conditions and Intervention

The instructional intervention consisted of three Generative AI-enhanced chatbot environments that were comparable in their overall instructional objectives and scientific research content but differed in the modality of the summaries generated and provided to students.

In the video-summary condition, students interacted with a GenAI-enhanced chatbot that supported research-related learning activities and provided concise video-based summaries of relevant concepts and procedures.

In the infographic-summary condition, students received the same type of GenAI-supported instructional interaction, while key information, concepts, processes, and relationships were summarized through infographic-based representations.

In the audio-summary condition, the chatbot provided equivalent instructional support accompanied by audio summaries designed to communicate the essential research-related information in an auditory format.

Across all three conditions, the learning activities focused on developing scientific research skills. The instructional content and intended learning outcomes were maintained as consistently as possible across groups so that the principal experimental distinction concerned the multimodal summary format.

4.5 Data Collection Procedures

Data were collected in several stages. First, the six participating sections of the Educational Qualification Program, whose students were enrolled in the Research Methods course, were identified, resulting in a total sample of 300 students. The participants were then organized into three experimental conditions, with 100 students assigned to each multimodal chatbot environment.

Before the intervention, participants were informed about the academic purpose of the study, the procedures involved, and the voluntary nature of their participation. They were assured that their

responses would remain confidential and would be used exclusively for scientific research purposes.

During the intervention, each group interacted with its designated Generative AI-enhanced chatbot environment. Students in the first group used video-based summaries, students in the second group used infographic-based summaries, and students in the third group used audio-based summaries.

Following the intervention, data was collected to measure students' perceptions of information and summary quality, cognitive load reduction, and the satisfaction of the three basic psychological needs of autonomy, competence, and relatedness. Students' scientific research skills were also assessed using the designated research-skills assessment instrument.

The resulting dataset was subsequently screened for completeness, response quality, missing values, outliers, and suitability for the planned statistical analyses.

4.6 Measurement Instruments

The measurement framework was aligned with the study's conceptual model and included measures of information and summary quality, cognitive load reduction, basic psychological needs (autonomy, competence, and relatedness), and scientific research skills. Established scales were adapted where appropriate to fit the context of Generative AI-enhanced chatbots and multimodal summaries. Unless otherwise specified, questionnaire items were rated on a five-point Likert scale ranging from 1 (Strongly Disagree) to 5 (Strongly Agree).

4.6.1 Information and Summary Quality

Information and Summary Quality was measured using four items adapted from the Information Quality scale developed by Seddon and Kiew (1996), specifically Items 4, 3, 10, and 7. The adapted items assessed four dimensions: accuracy, clarity, relevance/needs fit, and timeliness. The wording was contextualized to evaluate students' perceptions of the information and multimodal summaries delivered by the Generative AI-enhanced chatbot in video, infographic, and audio formats, while retaining the conceptual meaning of the original items.

4.6.2 Cognitive Load Reduction

Cognitive Load Reduction (CLR) was assessed using four contextually adapted items informed primarily by the cognitive load measurement approach of Klepsch et al. (2017) and the mental-effort framework of Paas (1992). The items measured the extent to which the chatbot reduced the mental demands of research tasks, facilitated understanding of research content, reduced feelings of cognitive overload, and decreased the mental effort required to complete research-related activities. Higher scores indicated greater perceived cognitive load reduction.

4.6.3 Basic Psychological Needs

The three basic psychological needs derived from Self-Determination Theory (SDT)autonomy, competence, and relatedness—were measured using items adapted to the Generative AI-supported research-learning context, based on the conceptualization of Deci and Ryan (2000). Autonomy assessed students perceived choice and control over the pace and manner of using AI chatbots for research learning. Competence assessed perceived capability, confidence, and accomplishment in performing research-related tasks with AI support. Relatedness assessed perceived support, understanding, and social connection with the chatbot, instructors, and peers during the learning process. These constructs were examined as mediating motivational mechanisms linking the AI-supported learning environment to students' scientific research skills.

4.6.4 Scientific Research Skills

Scientific Research Skills were measured using a 16-item researcher-developed instrument aligned with the content and intended learning outcomes of the Research Methods course undertaken by students in the Educational Qualification Program. The items were formulated to assess the core research competencies emphasized in the course and targeted by the intervention. Accordingly, the instrument was designed as a curriculum-aligned measure of students' scientific research skills within the Generative AI-enhanced chatbot learning environment.

4.7 Validity and Reliability

The psychometric properties of the measurement instruments were examined before testing the structural relationships. The measurement model assessment included indicator reliability, internal consistency reliability, convergent validity, and discriminant validity.

Internal consistency was assessed using Cronbach's alpha and composite reliability (CR). Convergent validity was evaluated through indicator loadings and Average Variance Extracted (AVE). Discriminant validity was examined using appropriate criteria, including the Heterotrait–Monotrait ratio (HTMT).

Items demonstrating inadequate psychometric properties were examined carefully before any decision regarding their retention or removal was made. This procedure ensured that the constructs used in subsequent structural analyses demonstrated acceptable reliability and validity.

4.8 Ethical Considerations

Participation in the study was voluntary. Participants were informed of the objectives and procedures of the research and their right to withdraw without academic consequences. Confidentiality and anonymity were maintained throughout the data collection and analysis processes.

No personally identifying information was included in the analytical dataset, and collected data was used exclusively for academic research purposes. The same instructional objectives and comparable learning content were provided across the three experimental conditions to ensure fairness, while the principal difference concerned the modality of the AI-generated summaries.

4.9 Data Analysis

Data analysis was conducted in two complementary stages to address both the comparative and explanatory objectives of the study.

First, descriptive statistics were calculated to summarize participant characteristics and the distributions of the study variables. Preliminary analyses were conducted to examine missing data, outliers, distributional characteristics, and other relevant statistical assumptions.

Differences among the video-summary, infographic-summary, and audio-summary groups were then examined using appropriate inferential statistical procedures. A one-way analysis of variance (ANOVA) was used to determine whether statistically significant differences existed among the three groups in scientific research skills and other relevant study constructs. Where significant overall differences were identified, appropriate post-hoc comparisons were conducted to determine which groups differed from one another.

Second, Partial Least Squares Structural Equation Modeling (PLS-SEM) was employed to test the proposed conceptual model. The analysis followed a two-stage procedure consisting of assessment of the measurement model followed by assessment of the structural model.

The structural model analysis examined the direct relationships between information and summary quality and cognitive load reduction and scientific research skills. It also examined their relationships with autonomy, competence, and relatedness, as well as the effects of these three psychological needs on scientific research skills.

Most importantly, mediation analysis was conducted to determine whether autonomy, competence, and relatedness mediated the effects of information and summary quality and cognitive load reduction on scientific research skills. The significance of the indirect effects was assessed using a bootstrapping procedure, providing estimates of path coefficients, standard errors, confidence intervals, and statistical significance.

The structural model was further evaluated using relevant criteria, including path coefficients (β), coefficient of determination (R^2), effect size (f^2), and predictive relevance (Q^2). Where appropriate, additional multigroup analyses were conducted to determine whether the structural relationships differed across the video, infographic, and audio chatbot conditions.

This analytical strategy enabled simultaneous examination of (a) differences among the three multimodal summary conditions, (b) the direct effects proposed in the conceptual model, and (c) the indirect effects operating through autonomy, competence, and relatedness.

5. Results

Generative AI-Enhanced Chatbots with Multimodal Summaries (Video, Infographics, and Audio) for Improving Scientific Research Skills: A Self-Determination Theory-Based Mediation Model

5.1 Data Screening and Descriptive Statistics

The final analytic sample comprised 308 participants (video n = 98, infographic n = 104, audio n = 106). Item-level missingness was low (1.51%) and was consistent with missing completely at random (Little's test, $\chi^2(2344) = 2290.94$, $p = .780$); missing values were therefore handled using the pre-specified mean-replacement rule for missingness below 5%. Screening for insufficient-effort responding and completeness removed cases that met the stated exclusion criteria. Univariate distributions were acceptable ($|skewness| \leq 0.84$; $|kurtosis| \leq 1.07$). Because the objective research-skills score was bounded and multivariate normality was not supported, PLS-SEM was retained as the primary estimator.

Table 2. Descriptive Statistics for Study Variables (N = 308)

Variable	M	SD	Min	Max
Information and summary quality	2.996	1.143	1.00	5.00
Cognitive load reduction	2.967	1.126	1.00	5.00
Autonomy	2.973	1.137	1.00	5.00
Competence	3.002	1.185	1.00	5.00
Relatedness	3.020	1.121	1.00	5.00
Scientific research skills	0.721	0.266	0.00	1.00

Note. Perceptual and motivational constructs were measured on five-point Likert scales. Scientific research skills is the proportion correct across 16 dichotomously scored items.

5.2 Measurement Model

The five perceptual and motivational constructs were modelled reflectively. Scientific research skills were represented by the proportion-correct score rather than as a reflective latent construct because the 16 achievement-test items sample a content domain; the test nevertheless showed adequate internal consistency (KR-20 = .873). For the five multi-item constructs, standardized loadings ranged from .870 to .912, composite reliability from .938 to .946, ρA from .915 to .925, and AVE from .792 to .815. Thus, indicator reliability, internal consistency, and convergent validity were satisfactory.

Table 3. Reliability and Convergent Validity

Construct	Items	Loading range	α	ρA	CR	AVE
Information & summary quality	4	.872–.912	.913	.915	.938	.792
Cognitive load reduction	4	.894–.900	.920	.921	.943	.806
Autonomy	4	.870–.903	.914	.918	.939	.795
Competence	4	.895–.910	.924	.925	.946	.815
Relatedness	4	.889–.906	.921	.923	.944	.807
Scientific research skills	1	—	—	—	—	—

Note. α = Cronbach's alpha; ρA = Dijkstra–Henseler's rho; CR = composite reliability; AVE = average variance extracted. Reliability of the research-skills test was KR-20 = .873.

Discriminant validity was also supported. HTMT values ranged from .239 to .537 and every bootstrap confidence-interval upper bound remained below .85 (highest upper bound = .624). The square root of AVE for each construct exceeded its correlations with the other constructs, and cross-loadings showed no problematic pattern. Full-collinearity VIFs ranged from 1.19 to 2.12 (< 3.3), indicating that common method variance was unlikely to threaten the findings. The estimated model also showed acceptable approximate fit (SRMR = .038).

5.3 Effects of Multimodal Summary Format (H1–H4c)

Summary modality significantly affected all six outcomes after Benjamini–Hochberg false-discovery-rate correction. Because variance homogeneity was violated for scientific research skills, Welch's ANOVA and Games–Howell comparisons were used for that outcome; conventional ANOVA with Tukey HSD was used for the five Likert composites. The largest modality effect occurred for scientific research skills (ω^2 = .164), whereas effects were moderate for information/summary quality and cognitive-load reduction and small for the three psychological needs.

Table 4. Effects of Summary Modality on Study Outcomes

Outcome	Video M (SD)	Infographic M (SD)	Audio M (SD)	F	FDR p	ω²
Research skills	0.86 (.19)	0.71 (.26)	0.59 (.27)	33.65	< .001	.164
Info./summary quality	3.39 (1.09)	3.07 (1.16)	2.59 (1.06)	13.40	< .001	.077
Cognitive-load reduction	3.43 (1.07)	2.82 (1.10)	2.67 (1.08)	13.76	< .001	.079
Autonomy	3.25 (1.14)	2.94 (1.12)	2.77 (1.12)	4.59	.011	.024
Competence	3.32 (1.14)	2.85 (1.16)	2.85 (1.21)	5.36	.006	.028
Relatedness	3.30 (1.07)	2.99 (1.17)	2.78 (1.07)	5.42	.006	.029

Note. Welch's F. H1, H2, H3, H4a, H4b, and H4c were supported. H4 was therefore supported at the aggregate conceptual level.

Pairwise comparisons sharpened this pattern. Scientific research skills followed a fully ordered video > infographic > audio pattern, with all three contrasts significant (video–infographic g = 0.64; video–audio g = 1.13; infographic–audio g = 0.46). For the perceptual and motivational outcomes, the video condition consistently exceeded audio, whereas infographic and audio were usually statistically indistinguishable except for information and summary quality. Thus, the video condition produced the most consistent gains across performance, appraisal, and need satisfaction.

5.4 Structural Model and Direct Effects (H5–H9)

Bootstrapping with 10,000 resamples supported all hypothesized paths in the primary structural model. Information and summary quality predicted autonomy, competence, and relatedness, with the strongest effect on competence (β = .464). Cognitive-load reduction also predicted all three needs, although with smaller coefficients (β = .140–.202). In turn, autonomy, competence, and relatedness positively predicted scientific research skills. The model explained 18.5% of autonomy, 26.3% of competence, 18.0% of relatedness, and 52.8% of scientific research skills. Inner VIFs were ≤ 1.58.

Table 5. Direct Structural Effects

Hyp.	Path	β	p	95% BCa CI	f^2	Decision
H5a	ISQ → Autonomy	.337	< .001	[.227, .434]	.13	Supported
H5b	ISQ → Competence	.464	< .001	[.369, .542]	.28	Supported
H5c	ISQ → Relatedness	.335	< .001	[.231, .427]	.13	Supported
H6a	CLR → Autonomy	.202	< .001	[.093, .304]	.05	Supported
H6b	CLR → Competence	.140	.004	[.042, .234]	.03	Supported
H6c	CLR → Relatedness	.197	< .001	[.087, .296]	.05	Supported
H7a	Autonomy → Skills	.269	< .001	[.173, .368]	.11	Supported
H7b	Competence → Skills	.316	< .001	[.215, .412]	.13	Supported
H7c	Relatedness → Skills	.182	< .001	[.092, .268]	.05	Supported
H8	ISQ → Skills	.107	.018	[.017, .198]	.02	Partial*
H9	CLR → Skills	.152	< .001	[.070, .230]	.04	Supported

Note. ISQ = information and summary quality; CLR = cognitive-load reduction. H5, H6, and H7 are supported through their respective sub-hypotheses. *H8 is significant in the reduced model but becomes non-significant when experimental condition is controlled (β = .050, p = .260), so it should not be interpreted as an independent direct pathway.

Predictive relevance was positive for all endogenous constructs. Q^2predict ranged from .105 to .222 for the three needs and was .288 for scientific research skills; the model outperformed the linear-regression benchmark on all indicators, supporting out-of-sample predictive relevance.

5.5 Mediation Effects (H10–H12)

All six specific indirect effects were significant because their bias-corrected bootstrap confidence intervals excluded zero. For information and summary quality, competence carried the largest indirect effect (β = .147), followed by autonomy (β = .091) and relatedness (β = .061). For cognitive-load reduction, the three indirect effects were smaller and similar in magnitude (β = .036–.055). Because the direct effects retained the same sign as the indirect effects, the pattern is consistent with complementary partial mediation.

Table 6. Specific and Total Indirect Effects on Scientific Research Skills

Hyp.	Indirect path / effect	β	p	95% BCa CI	Decision
H10a	ISQ → Autonomy → Skills	.091	< .001	[.053, .140]	Supported
H10b	ISQ → Competence → Skills	.147	< .001	[.098, .204]	Supported
H10c	ISQ → Relatedness → Skills	.061	< .001	[.030, .101]	Supported
—	ISQ total indirect	.298	< .001	[.234, .364]	—
H11a	CLR → Autonomy → Skills	.055	.001	[.024, .095]	Supported
H11b	CLR → Competence → Skills	.044	.004	[.015, .081]	Supported
H11c	CLR → Relatedness → Skills	.036	< .001	[.014, .067]	Supported
—	CLR total indirect	.135	< .001	[.076, .196]	—

Note. H10 and H11 are supported through all three specific mediators. Standardized estimates are based on 10,000 bootstrap resamples.

The aggregate mediation hypothesis (H12) was tested using a psychometrically justified second-order Basic Psychological Needs construct. Autonomy, competence, and relatedness loaded .748, .835, and .709 on the second-order construct; CR = .809 and AVE = .586. Information and summary quality (β = .499) and cognitive-load reduction (β = .231) predicted aggregate need satisfaction, which strongly predicted research skills (β = .594). The resulting indirect effects were .297, 95% CI [.236, .363], and .137, 95% CI [.083, .199], respectively. H12 was therefore supported as an aggregation-level confirmation of H10 and H11, not as a duplicate mediation test.

5.6 Experimental Mediation and Robustness Checks

When the experimental manipulation was entered using dummy coding, explained variance in research skills increased from 52.8% to 57.9%. Relative to audio, video showed a significant total effect (β = .474), comprising both a direct effect (β = .275) and a significant total indirect effect (β = .200). Infographic also outperformed audio overall (β = .217), but its indirect effect was not significant (β = .049, 95% CI [−.036, .137]); its advantage was therefore carried mainly by the direct pathway. This pattern indicates that video summaries improved performance both directly and through the proposed motivational mechanism, whereas the infographic advantage over audio did not measurably operate through the same mediation chain.

Robustness analyses supported the stability of the main conclusions. MICOM established partial measurement invariance sufficient for most structural comparisons; only one modality difference survived multiplicity adjustment, namely a stronger cognitive-load-reduction → research-skills path in video than audio. A confirmatory CB-SEM re-estimation yielded close coefficient convergence (maximum absolute discrepancy = .032) and acceptable fit (CFI = .994, TLI = .993, RMSEA = .023, SRMR = .028).

Excluding 17 influential cases increased R^2 for research skills from .528 to .595 without changing the sign or significance of any structural path.

5.7 Summary of Hypothesis Testing

Overall, the findings support the proposed model. H1–H4c confirm that summary modality affects performance, perceived summary affordances, and psychological need satisfaction. H5–H7 show that information/summary quality and cognitive-load reduction predict autonomy, competence, and relatedness, which in turn predict research skills. H9 is supported, whereas H8 should be interpreted cautiously because its small direct effect disappears after controlling for experimental condition. H10–H12 confirm that basic psychological needs—especially competence in the information-quality chain—carry a substantial portion of the relationship between perceived chatbot-summary characteristics and scientific research skills.

Table 7. Compact Hypothesis Decision Matrix

Hypotheses	Core test	Decision
H1–H4c	Modality effects on skills, ISQ, CLR, autonomy, competence, relatedness	Supported
H5 / H5a–c	ISQ → basic psychological needs	Supported
H6 / H6a–c	CLR → basic psychological needs	Supported
H7 / H7a–c	Basic psychological needs → research skills	Supported
H8	ISQ → research skills	Partially supported
H9	CLR → research skills	Supported
H10 / H10a–c	ISQ → needs → research skills	Supported
H11 / H11a–c	CLR → needs → research skills	Supported
H12	Aggregate basic-needs mediation	Supported

6. Discussion

6.1 Overview of Findings

This study investigated how Generative AI-enhanced chatbots with multimodal summaries contribute to the development of students' scientific research skills, with particular attention to two instructional characteristics—information and summary quality and cognitive load reduction—and to the mediating role of the three basic psychological needs specified by Self-Determination Theory (SDT): autonomy,

competence, and relatedness. The study further examined whether learning outcomes differed across three multimodal summary conditions: video, infographic, and audio summaries.

Overall, the findings support the proposed conceptual model by indicating that the educational value of GenAI-enhanced chatbots cannot be explained merely by students' exposure to artificial intelligence. Rather, their effectiveness appears to depend on the quality and usability of the information generated, the extent to which the learning environment manages unnecessary cognitive demands, and the degree to which the experience satisfies students' psychological needs. This interpretation is consistent with emerging research suggesting that the effectiveness of GenAI-supported learning is shaped not only by technological affordances but also by motivational and self-regulatory mechanisms. Chiu (2024), for example, conceptualized GenAI-supported activities through SDT and demonstrated how appropriately designed ChatGPT activities may support autonomy, competence, and relatedness while promoting self-regulated learning.

The present findings therefore extend the discussion beyond whether GenAI "works" in education toward the more theoretically informative question of how and under what instructional and psychological conditions GenAI contributes to higher-order learning outcomes.

6.2 Information and Summary Quality as a Foundation for Scientific Research Skill Development

The positive role of **information and summary quality** suggests that the instructional effectiveness of GenAI-enhanced chatbots depends substantially on the clarity, relevance, coherence, organization, and perceived usefulness of the information they provide. Scientific research is inherently intensive information. Students must distinguish relevant from irrelevant evidence, identify conceptual relationships, understand methodological procedures, connect theoretical constructs, and integrate multiple sources of information into coherent research decisions. Consequently, well-designed summaries may function as cognitive scaffolds that make complex research information more accessible and manageable.

This interpretation is consistent with recent evidence showing that combining AI-based chatbots with effective summarization strategies can enhance learning. Huang et al. (2026), for example, found that graphical summarization combined with an AI-based chatbot improved learning achievement and supported motivation, self-efficacy, critical thinking, and the management of cognitive load compared with less integrated instructional approaches. Such findings reinforce the argument that AI-generated content should not be evaluated simply according to its availability or speed of generation; rather, its instructional quality and representational structure determine whether it functions as meaningful learning support.

Within scientific research education, high-quality summaries may be particularly useful because they can highlight relationships among research problems, variables, theoretical frameworks, methodological choices, sampling procedures, instrumentation, data analysis, and interpretation. In this sense,

information and summary quality may facilitate the transition from fragmented methodological knowledge toward a more coherent understanding of the research process.

6.3 Cognitive Load Reduction and Scientific Research Skills

The positive contribution of cognitive load reduction is also theoretically meaningful. Scientific research tasks typically require students to coordinate numerous interacting elements simultaneously. Understanding a research problem, formulating questions, identifying variables, selecting an appropriate design, choosing sampling procedures, evaluating instruments, and determining appropriate statistical analyses can impose considerable demands on working memory.

From the perspective of Cognitive Load Theory, appropriately designed GenAI summaries may reduce extraneous processing by organizing information, highlighting essential concepts, providing concise explanations, and minimizing unnecessary searching across dispersed learning resources. In this way, cognitive resources can be redirected toward processes that are more directly relevant to understanding and applying scientific research concepts.

However, the present finding should not be interpreted as implying that less cognitive effort is always educationally superior. This distinction is particularly important in GenAI-supported learning. Stadler et al. (2024) found that students using an LLM experienced lower cognitive load than students using conventional web search, but they also produced lower-quality reasoning and argumentation in a scientific inquiry task. Thus, cognitive ease can become counterproductive when GenAI substitutes for rather than supports students' cognitive processing.

Accordingly, the contribution of cognitive load reduction in the present model is better interpreted as productive cognitive load management rather than indiscriminate cognitive offloading. Effective GenAI summaries should reduce unnecessary processing while preserving the intellectual effort required for evaluation, synthesis, methodological reasoning, and decision-making. This distinction is especially critical when the targeted outcome is scientific research skill rather than simple factual recall.

6.4 Basic Psychological Needs as a Mediating Mechanism

One of the central contributions of the study concerns the mediating role of autonomy, competence, and relatedness. According to SDT, learning environments are more likely to promote adaptive motivation and effective functioning when these three fundamental psychological needs are satisfied.

The findings are consistent with emerging empirical evidence from GenAI-supported learning contexts. Xia et al. (2023) examined psychological need satisfaction as a mediating mechanism in learning through an AI conversational chatbot, demonstrating the theoretical importance of need satisfaction in explaining how learner characteristics translate into self-regulated learning. More recently, Li and Chiu (2025) found that satisfaction of autonomy, competence, and relatedness significantly predicted

multiple dimensions of student engagement in a GenAI-chatbot learning environment and mediated the relationships between instructional support and engagement.

The present study extends this line of inquiry by positioning basic psychological needs as mechanisms connecting the instructional characteristics of GenAI-generated multimodal summaries with the development of scientific research skills.

Autonomy

The mediating role of autonomy suggests that students benefit when GenAI-supported learning enables them to exercise meaningful control over their learning process. Chatbots provide learners with opportunities to formulate questions, request clarification, revisit difficult concepts, control the pace of interaction, and access information in formats that accommodate their immediate learning needs.

Such affordances can strengthen students' sense of ownership over learning, which is especially important in research education, where learners must progressively develop the capacity to make independent methodological and conceptual decisions. Chiu (2024) similarly argues that GenAI learning activities can be designed to support student choice and self-directed regulation.

Nevertheless, autonomy support should not be equated with unrestricted dependence on AI. Darvishi et al. (2024), in a large, randomized study involving AI assistance, found that learners may come to rely on AI support rather than internalize the underlying regulatory processes. The implication is that GenAI should expand students' capacity to make research decisions rather than make those decisions on their behalf.

Competence

The mediating role of competence indicates that students' perceptions of effectiveness and mastery are another important pathway through which GenAI-enhanced learning may contribute to scientific research skills. Research methodology can be challenging for students because errors in one stage—for example, incorrectly identifying a variable or choosing an inappropriate design—can affect subsequent methodological decisions.

High-quality summaries, structured explanations, immediate responses, and opportunities to request clarification may provide students with repeated experiences of understanding and successful task completion. These experiences can strengthen perceived competence and encourage continued engagement with increasingly complex research activities.

This interpretation accords with SDT-based GenAI research emphasizing the importance of competence satisfaction within AI-supported environments. Li and Chiu (2025) found that psychological need satisfaction was positively associated with behavioral, emotional, cognitive, and agentic engagement in a GenAI chatbot context. Consequently, competence may provide an important motivational bridge

between understanding GenAI-generated research content and being able to apply that understanding independently.

Relatedness

The mediation through relatedness further suggests that effective learning with GenAI is not exclusively an individual cognitive process. Although a chatbot is not a human social partner, conversational responsiveness, immediate interaction, personalized responses, and integration with teacher-guided learning may contribute to students' perceptions that the learning environment is responsive and supportive.

This interpretation is consistent with research showing that SDT-based satisfaction remains relevant in technology-mediated and GenAI-supported environments. Li and Chiu (2025) demonstrated that needs satisfaction mediated the relationship between instructional support and several forms of engagement in learning with a GenAI chatbot.

However, relatedness should not be interpreted as evidence that AI replaces human relationships. Rather, the findings suggest that GenAI can contribute to a broader need-supportive learning ecology when appropriately integrated with instructor guidance, peer interaction, and meaningful academic support.

6.5 From Psychological Need Satisfaction to Scientific Research Skills

The positive relationships between autonomy, competence, relatedness, and scientific research skills provide an important extension of SDT into the domain of research-methods education. Scientific research competence requires more than memorizing definitions of research designs, variables, sampling methods, or statistical procedures. It involves independent judgment, methodological reasoning, persistence, critical evaluation, and the ability to integrate multiple forms of evidence.

Autonomy may encourage students to take ownership of research decisions; competence may increase confidence in applying methodological knowledge; and relatedness may provide the supportive learning conditions needed to sustain engagement with complex or ambiguous research problems. Collectively, these psychological resources provide a theoretically plausible explanation for why students who experience the GenAI environment as supportive may demonstrate stronger research capabilities.

The mediation findings are therefore particularly significant because they suggest that GenAI-enhanced multimodal summaries do not influence research skills solely through direct information delivery. Part of their educational impact is transmitted through changes in how learners experience themselves within the learning process—as autonomous decision makers, capable researchers, and supported participants in an academic learning environment.

6.6 Differences Across Video, Infographic, and Audio Summary Conditions

The differences observed among the three multimodal summary conditions provide an additional instructional contribution. In particular, the stronger performance associated with the video-summary condition suggests that the representation of information can influence how effectively students transform AI-generated explanations into research-related knowledge and skills.

Video summaries may be particularly well suited to scientific research instruction because many research procedures are sequential and relational. Students must understand not only individual concepts but also how one methodological decision leads to another. Video can combine narration, visual representation, examples, and temporal sequencing, potentially allowing learners to follow methodological processes in a coherent progression.

The effectiveness of visually structured information is also supported by emerging research on AI-assisted summarization. Huang et al. (2026) found that graphical summarization combined with AI-based chatbot support generated benefits for achievement and several cognitive and motivational outcomes. This evidence supports the broader principle that representation matters, although it does not imply that one modality is universally optimal.

Infographic summaries may be particularly effective for displaying conceptual relationships, comparisons, classifications, and methodological frameworks in a compact visual format. Their relative performance can therefore be understood in terms of their ability to externalize relationships among complex research concepts.

Audio summaries, by contrast, provide flexible and accessible learning opportunities but rely primarily on transient auditory processing. For complex methodological content involving multiple interacting variables, classifications, and procedural relationships, the absence of a persistent visual representation may make integration more demanding for some learners.

Importantly, the superiority of one format in the present study should be interpreted as context- and outcome-specific rather than as evidence of a universal modality hierarchy. The direct experimental literature comparing GenAI-generated video, infographic, and audio summaries simultaneously remains limited. Consequently, the present study contributes new comparative evidence rather than merely replicating an established ranking of modalities.

6.7 Integrating Cognitive and Motivational Explanations

A major theoretical contribution of the present study is that it integrates cognitive and motivational mechanisms rather than treating them as competing explanations.

Information and summary quality addresses the instructional integrity of the GenAI content. Cognitive load reduction addresses the cognitive manageability of that content. Autonomy, competence, and

relatedness explain the motivational experience generated by interaction with the learning environment. Scientific research skills represent the eventual higher-order learning outcome.

Accordingly, the proposed mechanism can be summarized as follows:

High-quality, cognitively manageable GenAI-supported multimodal learning → stronger satisfaction of autonomy, competence, and relatedness → stronger scientific research skills.

This integrated explanation advances existing GenAI research because it moves beyond technology acceptance or general perceptions of usefulness and specifies a psychologically grounded mechanism linking instructional design to learning outcomes. It also responds to recent calls to examine GenAI through theoretically grounded models rather than treating AI use itself as a sufficient explanatory variable. Chiu's (2024) SDT-based framework and subsequent empirical work on psychological needs in GenAI learning provide strong support for this direction.

6.8 Theoretical Contributions

This study makes several theoretical contributions. First, it extends Self-Determination Theory into GenAI-supported scientific research education by demonstrating the relevance of autonomy, competence, and relatedness to the development of research-related capabilities.

Second, the study conceptualizes basic psychological needs as mediating mechanisms, thereby explaining not simply whether GenAI-supported multimodal learning affects research skills but how these effects may emerge. This contribution builds on earlier AI-chatbot research demonstrating the mediating role of need satisfaction in self-regulated learning and engagement.

Third, incorporating cognitive load reduction alongside SDT provides a more comprehensive account of GenAI-supported learning. The findings suggest that successful learning depends simultaneously on maintaining manageable cognitive demands and creating a psychologically supportive learning experience.

Fourth, the comparison of video, infographic, and audio summaries extends multimodal learning research into an emerging GenAI context. Rather than treating multimodality as a single instructional feature, the study demonstrates the importance of examining the specific representational format through which AI-generated content is delivered.

6.9 Practical Implications

The findings have important implications for instructors, instructional designers, and developers of GenAI-enhanced learning environments. GenAI integration should prioritize instructional quality rather than mere technological availability. AI-generated summaries should be accurate, relevant, concise, logically organized, and aligned with the intended research-learning outcomes.

Designers should also distinguish between beneficial cognitive scaffolding and excessive cognitive offloading. Students should receive sufficient support to manage complex methodological information without allowing the chatbot to replace critical evaluation and independent research reasoning. This is particularly important given evidence that lower cognitive effort with LLMs does not necessarily translate into deeper scientific reasoning.

Furthermore, GenAI activities should deliberately support autonomy by allowing meaningful learner choice, competence through appropriately challenging tasks and constructive feedback, and relatedness through responsive interaction combined with human instructional support.

For research-methods courses specifically, video summaries may be particularly valuable for demonstrating sequential research processes, whereas infographics may complement them by representing conceptual and methodological relationships. Audio summaries can provide flexible reinforcement and accessibility. Thus, the pedagogical implication is not necessarily to replace all modalities with a single format but to align the modality with the cognitive characteristics of the research skill being developed.

6.10 Limitations and Directions for Future Research

Several limitations should be considered when interpreting the findings. First, the study was conducted within a specific educational program and course context; therefore, replication across universities, disciplines, educational levels, and cultural settings is required before broad generalizations can be made.

Second, although the model identifies theoretically grounded mediation through autonomy, competence, and relatedness, future longitudinal research should examine how these psychological needs change over time during sustained interaction with GenAI systems. Recent GenAI research increasingly emphasizes the importance of examining motivational and regulatory processes longitudinally rather than relying exclusively on single-time assessments.

Third, future studies should distinguish more carefully between productive cognitive load reduction and cognitive offloading. Measures of behavioral engagement, metacognition, independent task performance, and delayed transfer could determine whether students internalize research competencies or perform effectively only while AI assistance remains available. This issue is particularly important given evidence of reliance effects in AI-supported learning.

Fourth, further experimental research should directly compare different combinations and sequences of multimodal representations. For example, future research could investigate whether video explanations followed by infographic consolidation produce stronger outcomes than either modality alone, or whether learner characteristics moderate the effectiveness of video, infographic, and audio summaries.

Finally, future studies could incorporate additional motivational and cognitive mechanisms, including self-regulated learning, metacognitive monitoring, AI literacy, prior research experience, and epistemic

beliefs, to provide a more comprehensive explanation of how GenAI contributes to complex scientific research competencies.

6.11 Conclusion of the Discussion

Taken together, the findings indicate that the educational value of Generative AI-enhanced chatbots with multimodal summaries lies not simply in their capacity to generate information, but in their capacity to provide high-quality, cognitively manageable, and psychologically supportive learning experiences. Information and summary quality and cognitive load reduction appear to establish instructional conditions through which students can engage effectively with research-related content, while autonomy, competence, and relatedness provide important motivational mechanisms translating these conditions into stronger scientific research skills.

The study therefore supports a shift from a technology-centered interpretation of GenAI toward a learner-centered and theoretically grounded model. From this perspective, GenAI is most educationally valuable when it scaffolds rather than replaces thinking, strengthens rather than undermines learner agency, and uses multimodal representations strategically to support the cognitive and motivational demands of scientific research learning.

Declarations

Author Contribution

The author is the sole contributor to this manuscript and was responsible for the conceptualization, methodology, data collection, data analysis, interpretation of the results, writing of the original draft, and reviewing and editing of the manuscript. The author has read and approved the final version of the manuscript.

Ethics approval:

Ethical approval for this study was obtained from the Arab Open University Research Ethics Committee prior to data collection. The study did not involve human experimentation, the use of human tissue, or the collection of biological samples. All procedures involving human participants were conducted in accordance with the ethical standards of the institutional research ethics committee, the principles of the 1964 Declaration of Helsinki and its subsequent amendments, as well as other applicable national ethical guidelines.

Ethical Considerations

The study adhered to established ethical standards throughout the data collection and analysis processes. Participation was entirely voluntary, and respondents were informed about the objectives and academic purpose of the study prior to participation. Participants were assured that all responses would remain confidential and anonymous, and that the collected data would be used solely for academic and research purposes.

No personally identifiable information was collected from participants, and all data were securely managed to ensure privacy and confidentiality throughout the research process.

Consent to participate:

Electronically informed consent was obtained from all individual participants prior to their participation in the study. All participants were diploma students in Educational Qualification and were aged 18 years or older. Participation was voluntary, and participants were informed of the study purpose and their right to withdraw at any stage without consequences. The confidentiality and anonymity of participants' responses were ensured throughout the research process.

Consent for publication: Not applicable.

Data availability: The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request.

Finding:

The authors declare that no financial support, funding, grants, or other assistance were received for the conduct of this research, the preparation of this manuscript, or its publication.

Acknowledgments

The authors gratefully acknowledge the Arab Open University, Sultanate of Oman, for its support and for providing the academic environment that facilitated the completion of this study..

References

1. Ahmed, A., & Sultan, A. (2026). Examining the impact of generative AI on student motivation and engagement: The mediating role of autonomy-support and autonomous motivation in education. *Frontiers in Psychology, 17*, 1852265. https://doi.org/10.3389/fpsyg.2026.1852265
2. Annamalai, N., Bervell, B., Mireku, D. O., & Andoh, R. P. K. (2025). Artificial intelligence in higher education: Modelling students' motivation for continuous use of ChatGPT based on a modified self-determination theory. *Computers and Education: Artificial Intelligence, 8*, 100346. https://doi.org/10.1016/j.caeai.2024.100346
3. Barros, C. F., Azevedo, B. B., Graciano Neto, V. V., Kassab, M., Kalinowski, M., do Nascimento, H. A. D., & Bandeira, M. C. G. S. P. (2024). *Large language model for qualitative research: A systematic mapping study*. arXiv. https://doi.org/10.48550/arXiv.2411.14473
4. Bhat, S. A., & Alyahya, S. (2024). Infographics in educational settings: A literature review. *IEEE Access, 12*, 1633–1649. https://doi.org/10.1109/ACCESS.2023.3348083
5. Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E. L., Van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, A., Ryan, R. M., Sheldon, K. M., Soenens, B., Van Petegem, S., &

Verstuyf, J. (2015). Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion, 39*, 216–236. https://doi.org/10.1007/s11031-014-9450-1

6. Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. *Educational Technology Research and Development, 72*, 2401–2416. https://doi.org/10.1007/s11423-024-10366-w
7. De Silva, U., Fernando, L., Bandara, K., & Nawaratne, R. (2025).*Video summarisation with incident and context information using generative AI*. https://arxiv.org/abs/2501.04764
8. Deci, E. L., & Ryan, R. M. (2000). The “what” and “why” of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
9. Desmedt, C. P., Budts, W., De Vos, M., & Moons, P. (2025). Artificial intelligence-generated podcasts open new doors to make science accessible: A mixed-method evaluation of quality and the potential for science communication. *European Journal of Cardiovascular Nursing, 24*(6), 885–895. https://doi.org/10.1093/eurjcn/zvaf074
10. Díaz, L., & Carmona-Galindo, V. D. (2026). Turning the page: Pre-class AI-generated podcasts improve student outcomes in ecology and environmental biology. *Education Sciences, 16*(1), 168. https://doi.org/10.3390/educsci16010168
11. Ebrahimabadi, M., Rezaei, K., Moini, A., Fournier, A., & Abedi, A. (2018). Infographics or video; which one is more effective in asthmatic patients’ health? A randomized clinical trial. *Journal of Asthma*. https://doi.org/10.1080/02770903.2018.1536143
12. Ellikkal, A., & Rajamohan, S. (2024). AI-enabled personalized learning: Empowering management students for improving engagement and academic performance. *Emerald Publishing Limited*.
13. Fabbri, A. R., Kryściński, W., McCann, B., Xiong, C., Socher, R., & Radev, D. (2021). SummEval: Re-evaluating summarization evaluation. *Transactions of the Association for Computational Linguistics, 9*, 391–409. https://doi.org/10.1162/tacl_a_00373
14. Guay, F. (2022). Applying self-determination theory to education: Regulations types, psychological needs, and autonomy supporting behaviors. *Canadian Journal of School Psychology, 37*(1), 75–92. https://doi.org/10.1177/08295735211055355
15. Hamilton, J., Barr, L., Costello, M., Thalib, L., & Jones, M. (2026). GenAI-assisted podcasts to support nursing students’ research literacy and evidence-based practice: A feasibility study. *Collegian, 33*(4), 285–293. https://doi.org/10.1016/j.colegn.2026.05.002
16. Heimbürger, A., Keto, H., & Isomöttönen, V. (2020). Learning via summarizing infographics assignment in software engineering management e-course. In *Proceedings of the IEEE Frontiers in Education Conference (FIE)*. https://doi.org/10.1109/FIE44824.2020.9274229

17. Huang, A. Y. Q., Lin, C. C., Su, S. Y., Yen, R. X., & Yang, S. J. H. (2025).Improving students' learning performance and summarization ability through a generative AI-enabled chatbot. *Educational Technology & Society, 28*(3), 82–111.
18. Jafarian, N. R., & Kramer, A.-W. (2025). AI-assisted audio-learning improves academic achievement through motivation and reading engagement. *Computers and Education: Artificial Intelligence, 8*, 100357. https://doi.org/10.1016/j.caeai.2024.100357
19. Jee, H. (2023). Emergence of artificial intelligence chatbots in scientific research. *Journal of Exercise Rehabilitation, 19*(3), 139–140. https://doi.org/10.12965/jer.2346234.117
20. Jose, B., Cherian, J., Verghis, A. M., Varghise, S. M., S, M., & Joseph, S. (2025). The cognitive paradox of AI in education: Between enhancement and erosion. *Frontiers in Psychology, 16*, 1550621. https://doi.org/10.3389/fpsyg.2025.1550621
21. Kim, K., Hahm, J., Kim, S., Sul, J., Kim, B., & Lee, J. (2025).*SummDiff: Generative modeling of video summarization with diffusion*.https://arxiv.org/abs/2510.08458
22. Klepsch, M., Schmitz, F., & Seufert, T. (2017). Development and validation of two instruments measuring intrinsic, extraneous, and germane cognitive load. *Frontiers in Psychology, 8*, 1997. https://doi.org/10.3389/fpsyg.2017.01997
23. Li, Y., & Chiu, T. K. F. (2025). The mediating effects of needs satisfaction on the relationship between teacher support and student engagement with generative artificial intelligence (GenAI) chatbots from a self-determination theory (SDT) perspective. *Education and Information Technologies, 30*, 20051–20070. https://doi.org/10.1007/s10639-025-13574-w
24. Li, Y., Zhou, X., & Chiu, T. K. F. (2025). Systematics review on artificial intelligence chatbots and ChatGPT for language learning and research from self-determination theory (SDT): What are the roles of teachers? *Interactive Learning Environments, 33*(3), 1850–1864. https://doi.org/10.1080/10494820.2024.2400090
25. Liang, G., Lv, Y., Li, S., Zhang, S., & Zhang, Y. (2021).Unsupervised video summarization with a convolutional attentive adversarial network.https://arxiv.org/abs/2105.11131
26. Luo, Z., Xie, Q., & Ananiadou, S. (2024). Factual consistency evaluation of summarization in the era of large language models. *Expert Systems with Applications, 254*, 124456. https://doi.org/10.1016/j.eswa.2024.124456
27. Lyra, K. T., Isotani, S., Reis, R. C. D., Marques, L. B., & Pedro, L. Z. (2016). Infographics or graphics+text: Which material is best for robust learning? In *IEEE International Conference on Advanced Learning Technologies*.
28. Maddens, L., Depaepe, F., Raes, A., & Elen, J. (2023). Fostering students' motivation towards learning research skills: The role of autonomy, competence and relatedness support. *Instructional Science, 51*, 165–199. https://doi.org/10.1007/s11251-022-09606-4
29. Mahapatra, D., Mariappan, R., Rajan, V., Yadav, K., Seby, A., & Roy, S. (2018). VideoKen: Automatic video summarization and course curation to support learning. https://doi.org/10.1145/3184558.3186988

30. Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.
31. Niu, B., & Mvondo, G. F. N. (2024). I am ChatGPT, the ultimate AI chatbot! Investigating the determinants of users' loyalty and ethical usage concerns of ChatGPT. *Journal of Retailing and Consumer Services, 76*, 103562. https://doi.org/10.1016/j.jretconser.2023.103562
32. Paas, F. G. W. C. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology, 84*(4), 429–434. https://doi.org/10.1037/0022-0663.84.4.429
33. Patac, L. P., & Patac, A. V., Jr. (2025). Using ChatGPT for academic support: Managing cognitive load and enhancing learning efficiency—A phenomenological approach. *Social Sciences & Humanities Open, 11*, 101301. https://doi.org/10.1016/j.ssaho.2025.101301
34. Reddy, P. A., Prasanna, G., Sulthana, S., Vaishnavi, P., Srilekha, S., & Arora, G. (2025). Smart video assistant: A video summarizer and interactive chatbot system. In *Proceedings of the 5th International Conference on Advancement in Electronics & Communication Engineering (AECE)*. https://ieeexplore.ieee.org
35. **Rjoub, A., Daraghma, M., Dumaidi, Y., Nour, M. A., Abusnaina, D., Fiqyat, T., Abdul-Hafez, H. A., Zaben, F., & Zaben, M. (2025).** Undergraduate medical students' perspectives on research education and their confidence in research skills: A cross-sectional study from Palestine. *BMC Medical Education, 25*, 1005. https://doi.org/10.1186/s12909-025-07586-w
36. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68
37. Seddon, P. B., & Kiew, M.-Y. (1996). A partial test and development of DeLone and McLean's model of IS success. *Australasian Journal of Information Systems, 4*(1), 90–109.DOI:10.3127/ajis.v4i1.379.https://ajis.aaisnet.org/index.php/ajis/article/view/379
38. **Stokel-Walker, C., & Van Noorden, R. (2023).** What ChatGPT and generative AI mean for science. *Nature, 614*(7947), 214–216. https://doi.org/10.1038/d41586-023-00340-6
39. Sturgeon Delia, C. (2025). A pilot practitioner's inquiry into students' reflections on AI-generated podcasts in higher education. *Education Sciences, 15*(12), 1680. https://doi.org/10.3390/educsci15121680
40. Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer. https://doi.org/10.1007/978-1-4419-8126-4
41. Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292. https://doi.org/10.1007/s10648-019-09465-5
42. Vansteenkiste, M., Ryan, R. M., & Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion, 44*, 1–31. https://doi.org/10.1007/s11031-019-09818-1

43. **Vígh, T. (2024).** Development of research skills through research-focused microteaching lesson study in preservice teacher education. *Teaching and Teacher Education, 145*, 104618. https://doi.org/10.1016/j.tate.2024.104618
44. Yadav, M., Kumari, N., Dogra, K., Kumar, Y., Arya, T., & Purwar, S. (2024). Advancements in educational technology and video summarization techniques: A comprehensive review. *IPEC Journal of Science & Technology*, 3(1), 6–12.
45. Yuan, L., Xu, J., & Zhan, Z. (2025). Effects of pedagogical agent-generated summaries on video-based learning: Evidence from eye-tracking and EEG. *Education Sciences*.

Figures

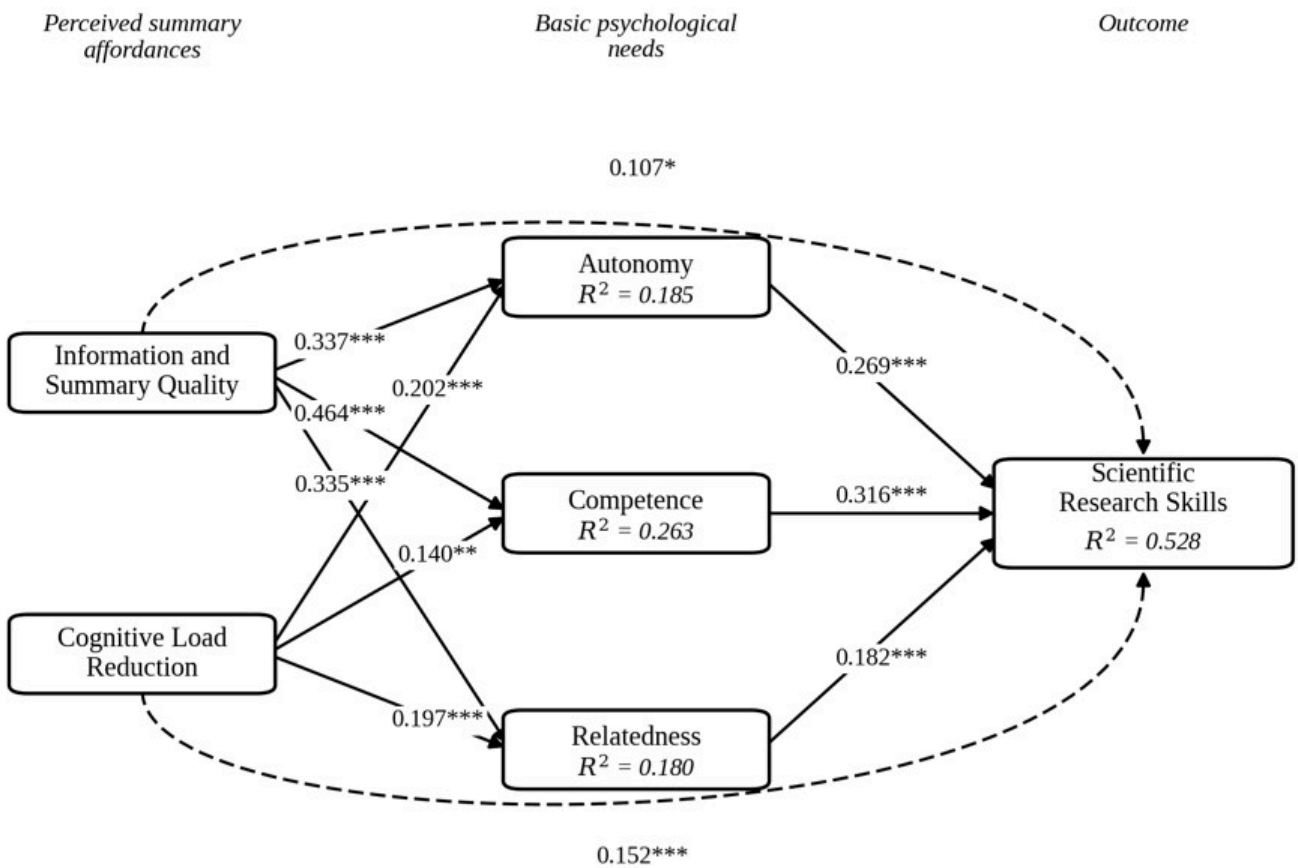

Figure 1

Estimated Structural Model with Standardised Path Coefficients

Note. N = 308. Solid arrows denote the hypothesised mediated pathways; dashed arrows denote direct effects. Values are standardised path coefficients from PLS-SEM with 10,000 bootstrap resamples and bias-corrected and accelerated confidence intervals. R^2 values are reported within each endogenous construct. * p < .05. ** p < .01. *** p < .001.

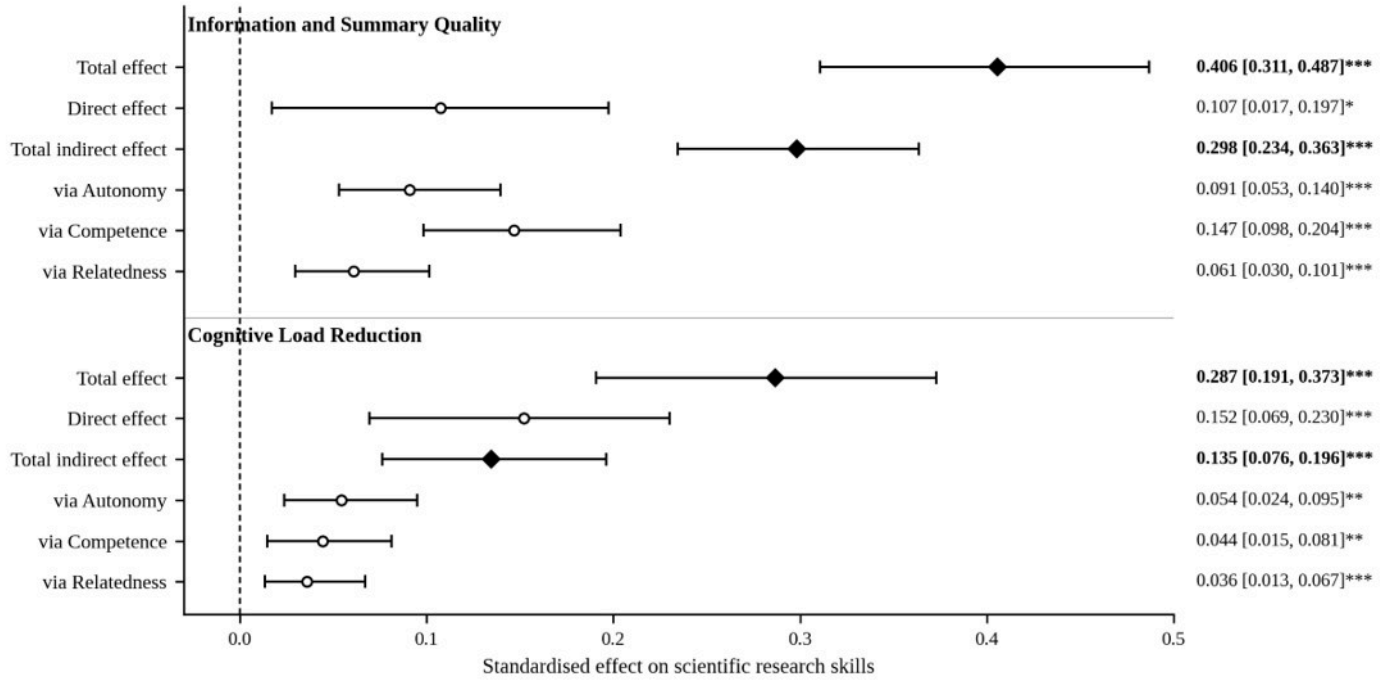

Figure 2

Decomposition of Direct, Specific Indirect, and Total Effects on Scientific Research Skills

Note. N = 308. Filled diamonds denote total and total indirect effects; open circles denote direct and specific indirect effects. Bars are 95% bias-corrected and accelerated bootstrap confidence intervals from 10,000 resamples. * p < .05. ** p < .01. *** p < .001.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- rusult.docx