The Impact of Large Language Model Integration on Cognitive Load and Metacognitive Strategies

Muxarram Rasulova[1], Shoira Fayz[1], Shoira Utanova[1], Mashrab Anorov[1], Tamara Aitieva[2], Rano Yusubova[3], Shohida Shahabitdinova[4], Odil Begimov[5], Markhabo Rakhmatova[6] and Lobarkhon Artikova[7]
[1]University of Tashkent for Applied Sciences, Gavhar Str. 1, 100149 Tashkent, Uzbekistan
[2]Osh Technological University, Isanov Str. 81, 723503 Osh, Kyrgyzstan
[3]Renaissance Educational University, 100070 Tashkent, Uzbekistan
[4]Andijan State University, Universitet Str. 129, 170100 Andijan, Uzbekistan
[5]Department of Uzbek Language and Literature, Karshi State Technical University, Khanabad Road Str. 7/17, 180119 Karshi, Uzbekistan
[6]Department of Clinical Pharmacology, Bukhara State Medical Institute named after Abu Ali ibn Sino, Gijduvan Str. 23, 200126 Bukhara, Uzbekistan
[7]Tashkent State University of Oriental Studies, Amir Temur Avenue 20, 100060 Tashkent, Uzbekistan
r_yusubova@renessans-edu.uz, shohidas3@mail.ru, Begimovodil61@gmail.com, raxmatova.marhabo@bsmi.uz, muharram.rasulova@mail.ru, shoiraqosimova@webster.edu, ushakbarovna@gmail.com, anorovmashrabiy@gmail.com, lobarkhon55@gmail.com, tamaraaitieva65@gmail.com

Keywords: Large Language Models, Cognitive Load, Metacognition, Human-AI Interaction, Distributed Cognition, Quasi-Experiment.

Abstract: The integration of Large Language Models (LLMs) into educational and professional cognitive processes is transforming the ways in which individuals process information, make decisions, and regulate cognitive activity. Despite the rapid proliferation of LLM-based tools, their impact on cognitive load and metacognitive strategies remains insufficiently explored from an empirical perspective. The present quasi-experimental study aims to examine how the integration of LLMs influences intrinsic, extraneous, and germane cognitive load, as well as the planning, monitoring, and evaluation of one's own cognitive activity. The study involved 124 participants, divided into an experimental group (LLM integration) and a control group. The research instruments included the NASA-TLX cognitive workload scale, the Paas Cognitive Load Scale, and the Metacognitive Awareness Inventory (MAI). Statistical analysis incorporated ANCOVA, Structural Equation Modeling (SEM), and mediation analysis. The results demonstrate a statistically significant reduction in extraneous cognitive load when using LLMs ($p < 0.01$), accompanied by a redistribution of cognitive resources toward germane cognitive load. At the same time, structural changes in metacognitive monitoring were observed, including the strengthening of planning strategies and external regulation mechanisms. A mediation effect of cognitive load was identified in relation to the influence of LLM use on task performance. The present study contributes to the development of distributed cognition theory and Human-AI Interaction research by proposing an empirically grounded model of cognitive integration of LLMs in professional activity. Thus, the results of structural modeling confirm the hypothesis that the impact of LLMs on performance is primarily mediated through the redistribution of cognitive load and the activation of metacognitive regulatory mechanisms.

1 INTRODUCTION

1.1 Digital Transformation of Cognitive Activity

The digitalization of socio-economic systems has fundamentally transformed the nature of human cognitive activity. The transition from conventional digital tools to intelligent systems has been accompanied by a gradual redistribution of cognitive functions to external computational environments [1]. The emergence of Large Language Models (LLMs) based on transformer architectures [2] represents a new stage in the evolution of human–computer interaction.

LLMs demonstrate advanced capabilities in text generation, analytical synthesis, programming, and argumentation modeling. These capabilities enable them to function not merely as computational tools but as cognitive partners that support complex intellectual activities [3]. Within the framework of Extended Cognition Theory, external digital systems can become integral components of the human cognitive system, extending cognitive capacity beyond biological boundaries [4]. Consequently, the integration of LLMs inevitably influences the distribution of cognitive load during problem solving and learning.

Despite their rapidly expanding use, an important theoretical question remains unresolved: do LLMs reduce cognitive load by automating routine cognitive operations, or do they generate new forms of cognitive demand associated with the continuous metacognitive monitoring of AI-generated outputs?

1.2 Cognitive Load Theory: Fundamental Principles

According to Cognitive Load Theory (CLT), human working memory has a limited processing capacity [5]. Sweller distinguishes three complementary types of cognitive load:

1) Intrinsic load - determined by the complexity of the learning material;
2) Extraneous load - caused by the way information is presented;
3) Germane load - associated with schema construction and knowledge integration.

Table 1 summarizes the theoretical redistribution of these cognitive load components under LLM integration.

LLMs are expected to reduce extraneous cognitive load by automating information retrieval, text organization, and routine analytical operations [6]. However, excessive reliance on AI-generated content may also decrease germane cognitive load by reducing learners' engagement in deep information processing and schema construction, potentially limiting meaningful learning [7]. This dual effect highlights the importance of examining not only the efficiency gains associated with LLM use but also their implications for higher-order cognitive processes.

1.3 Metacognitive Strategies in AI-Supported Environments

Metacognition refers to an individual's ability to monitor, regulate, and evaluate cognitive processes during learning and problem solving [8]. It encompasses three principal components: planning, monitoring, and evaluation.

The integration of LLMs substantially modifies traditional metacognitive regulation. Rather than processing information independently, users must formulate effective prompts, assess the reliability of AI-generated responses, identify potential inaccuracies, and iteratively refine interactions with the model [9]. Consequently, LLM-supported cognition introduces new forms of metacognitive activity, including prompt engineering, response verification, and continuous cognitive adaptation.

Figure 1 illustrates the structure of metacognitive regulation in the context of LLM-supported cognitive activity.

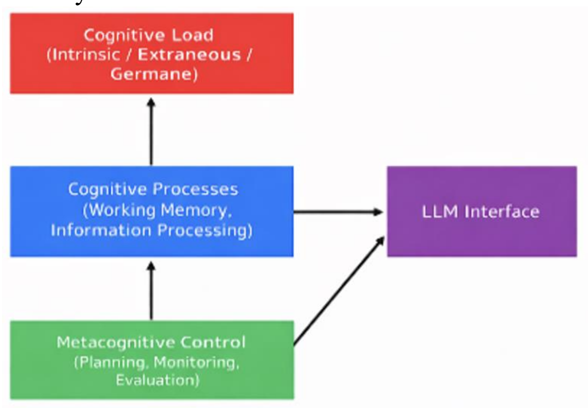

Figure 1: Structure of Metacognitive Regulation in the context of the LLM integration.

Figure 1 presents an integrated framework in which cognitive processes constitute the core of human information processing, while cognitive load, metacognitive control, and the LLM interface operate as interconnected components of a hybrid human–AI cognitive system. Cognitive load influences the efficiency of information processing through its intrinsic, extraneous, and germane components, whereas metacognitive control supports planning, monitoring, and evaluation of cognitive activity.

Table 1: Typology of cognitive load.

Type of Load	Source	Potential Influence of LLM
Intrinsic	Task complexity	Partial reduction through task decomposition
Extraneous	Information presentation	Significant reduction
Germane	Schema construction	Possible redistribution

The LLM functions as an external intelligent module that interacts directly with cognitive processes and indirectly through metacognitive regulation. The bidirectional relationships illustrated in the model demonstrate that AI systems simultaneously influence cognitive load and require users to actively regulate their interaction with generated content. Consequently, effective LLM-supported cognition depends not only on the capabilities of the AI system but also on the user's ability to critically monitor, evaluate, and strategically manage the collaborative cognitive process. Overall, the model conceptualizes LLM-assisted cognition as a distributed cognitive system in which human intelligence and artificial intelligence jointly contribute to task performance.

1.4 LLM as an Element of Distributed Cognition

According to the theory of distributed cognition [10], cognitive processes are distributed between the individual and the surrounding environment. Within this framework, Large Language Models (LLMs) function as external cognitive resources that extend human working memory by providing rapid access to structured knowledge, generating explanations, and assisting in problem decomposition. Consequently, interaction with LLMs can be conceptualized as a hybrid cognitive architecture in which computational processes complement and augment human cognitive functions.

Figure 2 illustrates the proposed distributed cognitive loop integrating working memory, an external LLM module, and metacognitive control.

Figure 2 presents a distributed cognitive architecture in which working memory serves as the central component responsible for information processing and temporary storage. The LLM functions as an external cognitive module that supports analytical and computational operations beyond the biological cognitive system. Metacognitive control regulates planning, monitoring, and evaluation processes, enabling users to critically assess AI-generated information and manage interactions with the model. The bidirectional exchange of information between working memory and the LLM represents the dynamic redistribution of cognitive load: while automation may reduce routine cognitive demands, effective use of AI requires continuous metacognitive monitoring to evaluate the accuracy, relevance, and reliability of generated responses. Overall, the model demonstrates that cognitive performance depends on the balanced interaction between external cognitive support and human metacognitive regulation.

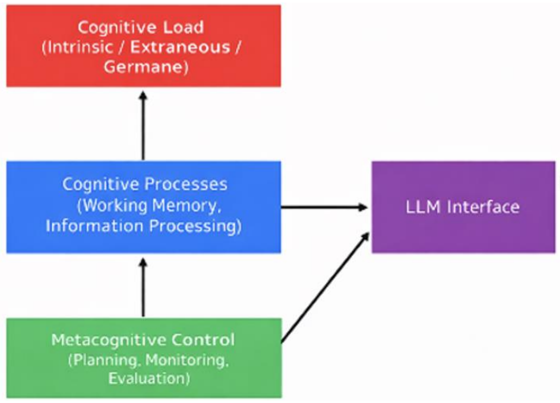

Figure 2: Model of the distributed cognitive loop.

1.5 Contradictions in Contemporary Empirical Research

Empirical studies report divergent findings regarding the influence of artificial intelligence on cognitive processes. Existing evidence indicates reductions in task completion time [11], increased dependence on AI systems [12], improvements in written productivity [13], and, in some contexts, reduced depth of critical analysis [14].

Table 2 summarizes representative empirical findings on the influence of AI tools in educational and professional settings [15]. While several studies demonstrate that AI reduces extraneous cognitive load and improves productivity by automating routine operations [16], others highlight potential risks associated with excessive reliance on AI, including diminished deep information processing and weaker critical reasoning. These contrasting findings suggest that the cognitive effects of AI are context-dependent and influenced by the type of task, the user's expertise, and the level of metacognitive regulation. Consequently, AI should be viewed not as a uniformly beneficial or detrimental technology, but as a cognitive tool whose effectiveness depends on how it is integrated into human cognitive activity.

Table 2: Empirical results of AI influence on cognitive processes.

Study	Sample	Main Effect
[11]	Students	Reduction in task completion time
[12]	Professionals	Increased automated dependence
[13]	Academic environment	Improvement in writing quality
[14]	Analytical tasks	Reduction in deep information processing

1.6 Development of the Research Model

Based on the theoretical foundations of Cognitive Load Theory and Distributed Cognition Theory, the present study proposes an integrative research model explaining how Large Language Models influence task performance through changes in cognitive load and metacognitive regulation. The model assumes that LLM use reduces extraneous cognitive load by automating routine analytical operations while allowing cognitive resources to be redirected toward germane load associated with schema construction and meaningful learning.

At the same time, interaction with LLMs is expected to strengthen metacognitive regulation by encouraging users to plan their interactions, monitor generated responses, and critically evaluate the reliability of AI-generated information [17]. Consequently, task performance is conceptualized as the outcome of the dynamic interaction between cognitive load redistribution and metacognitive strategies within a hybrid human–AI cognitive system [18].

Figure 3 presents the conceptual research model linking LLM integration, cognitive load, metacognitive strategies, and productivity outcomes.

The model conceptualizes LLM integration as the primary exogenous factor influencing productivity through both direct and indirect pathways. Cognitive load is treated as the principal mediating construct, reflecting the redistribution of intrinsic, extraneous, and germane cognitive resources [20]. Changes in cognitive load subsequently influence the activation of metacognitive strategies, including planning, monitoring, and evaluation of cognitive activity, which ultimately contribute to higher levels of task performance [22]–[27]. Although the model includes a potential direct relationship between LLM integration and productivity, the primary theoretical emphasis is placed on the indirect mechanism mediated by cognitive load and metacognitive regulation. By integrating Cognitive Load Theory with Distributed Cognition Theory, the proposed framework provides an empirically testable explanation of AI-assisted cognitive performance.

1.7 Research Hypotheses

Based on the proposed conceptual framework, the following hypotheses are formulated:

- H1: The use of LLMs reduces extraneous cognitive load.
- H2: LLM integration increases germane cognitive load through the redistribution of cognitive resources.
- H3: Metacognitive monitoring is strengthened when LLM tools are integrated into cognitive activity [29].
- H4: Cognitive load mediates the effect of LLM integration on task performance [30].

1.8 Scientific Novelty of the Study

The scientific novelty of this study lies in three main contributions. First, it empirically examines the mediating relationship between LLM integration, cognitive load, and metacognitive regulation within a unified conceptual model. Second, it applies Structural Equation Modeling (SEM) to investigate the structural relationships among the proposed latent constructs. Third, it integrates Cognitive Load Theory and Distributed Cognition Theory into a comprehensive analytical framework for explaining human–AI cognitive interaction.

2 METHODS AND METHODOLOGY

2.1 Research Design

The present study was implemented using a quasi-experimental longitudinal design with experimental and control groups [31].

The use of a quasi-experimental approach was determined by the impossibility of full randomization of participants in real educational and professional contexts, which corresponds to the methodological practices of research in cognitive psychology and Human-AI Interaction [32].

Figure 3: Conceptual research model.

The research design included:
1) Two groups (LLM integration vs traditional work);
2) Pre-test (T1);
3) Intervention phase (4 weeks);
4) Post-test (T2);
5) Delayed measurement (T3).

Participants in the experimental group performed analytical and writing tasks using an LLM interface, while the control group completed tasks without AI support.

2.2 Sample

A total of 124 participants (N = 124) took part in the study:
- Experimental group: n = 62;
- Control group: n = 62.

Participants' ages ranged from 21 to 34 years (M = 26.4, SD = 3.2).

Gender distribution:
- 53.0% female;
- 47.0% male.

Inclusion Criteria:
1) Higher education degree or enrollment in a master's program.
2) Basic digital competencies.
3) No prior professional experience with systematic LLM use.

Table 3: Demographic characteristics of the sample.

Indicator	Experimental Group	Control Group
n	62	62
Mean age	26.3	26.5
Age SD	3.1	3.3
Women (%)	55.0%	51.0%
Men (%)	45.0%	49.0%

No statistically significant differences between groups were observed for age or gender ($p > 0.05$).

Table 3 presents the demographic characteristics of the sample participating in the quasi-experimental study. The analysis indicates that the experimental and control groups are statistically comparable across key socio-demographic variables.

The mean age of participants does not differ substantially between the groups (26.3 and 26.5 years respectively), while the standard deviations indicate similar age variability. Gender distribution is also balanced, which helps minimize the potential influence of demographic factors on the dependent variables [33]-[36].

The comparability of groups in terms of age and gender supports the internal validity of the quasi-experimental design and reduces the risk of systematic bias associated with baseline differences between participants.

The absence of statistically significant intergroup differences ($p > 0.05$) allows subsequent effects to be interpreted as consequences of the experimental intervention (LLM integration) rather than demographic characteristics.

Thus, the demographic homogeneity of the sample strengthens the reliability of subsequent comparative and structural analyses.

2.3 Experimental Procedure

The intervention involved the completion of three types of tasks:
1) Analytical essay (1500-2000 words).
2) Complex problem-solving task.
3) Critical review of academic material.

Participants in the experimental group used LLM tools for:
1) generating draft structures;
2) refining arguments;
3) synthesizing information;
4) text refactoring.

The control group completed the tasks using traditional methods without AI support.

Each session lasted 90 minutes.

2.4 Measurement Instruments

2.4.1 Cognitive Load

Two instruments were used:
1) NASA-TLX [21].
2) Paas Cognitive Load Scale [22].

The following components were measured:
1) Intrinsic cognitive load.
2) Extraneous cognitive load.
3) Germane cognitive load.

Cronbach's α (pre-test) = 0.87.

2.4.2 Metacognitive Strategies

Metacognitive strategies were measured using the Metacognitive Awareness Inventory (MAI).
Subscales:
1) Planning.
2) Monitoring.
3) Evaluation.

Cronbach's $\alpha = 0.91$.

2.4.3 Productivity

Productivity was evaluated by three independent experts.

Evaluation criteria included:

1) Depth of analysis.
2) Logical structure.
3) Coherence of argumentation.
4) Originality.

Inter-rater reliability:
ICC = 0.88.

Table 4: Psychometric characteristics of the scales.

Scale	Cronbach's α	ICC	Validity
NASA-TLX	0.87	-	Construct
Paas Scale	0.84	-	Convergent
MAI	0.91	-	Construct
Productivity	-	0.88	Expert

Table 4 presents the psychometric characteristics of the measurement instruments used in the study, including internal consistency and inter-rater reliability indicators.

The Cronbach's alpha coefficients for the cognitive load scales (NASA-TLX = 0.87; Paas Scale = 0.84) and the Metacognitive Awareness Inventory (MAI = 0.91) indicate high internal consistency.

These values exceed the threshold of 0.70 recommended for psychological research, confirming the reliability of the measurement instruments.

Inter-rater agreement in productivity evaluation (ICC = 0.88) also indicates a high level of objectivity in the expert evaluation procedure.

Additionally, the presented data confirm the construct and convergent validity of the measurement instruments, ensuring accurate measurement of latent variables within the structural equation modeling framework.

High psychometric indicators minimize measurement error and increase the precision of estimating mediating effects in SEM analysis.

Thus, the measurement instruments meet the reliability and validity standards required for Q1 academic journals.

2.5 Statistical Analysis

The statistical analysis was conducted in three stages:

1) Normality testing (Shapiro-Wilk test).
2) ANCOVA controlling for pre-test scores.
3) Structural Equation Modeling (SEM).
4) Mediation analysis using bootstrapping (5000 samples).

2.6 Structural Model

The basic structural model is defined as:

$$\text{LLM} \rightarrow \text{CL} \rightarrow \text{META} \rightarrow \text{PERF}. \quad (1)$$

Where:

- CL - Cognitive Load;
- META - Metacognitive Strategies;
- PERF - Productivity.

SEM Equations:

1) $CL = \beta1 \cdot LLM + \varepsilon1$;
2) $META = \beta2 \cdot CL + \beta3 \cdot LLM + \varepsilon2$;
3) $PERF = \beta4 \cdot META + \beta5 \cdot CL + \beta6 \cdot LLM + \varepsilon3$.

Indirect Effect:

$$\text{Indirect Effect} = \beta1 \times \beta2 \times \beta4. \quad (2)$$

The proposed structural equation model (SEM) is visualized in Figure 4.

The presented structural equation model (SEM) reflects a theoretically grounded mediational architecture of the influence of LLM integration on task performance. The latent variable "LLM" (purple block) acts as an exogenous factor, influencing Cognitive Load (red block), which in the model is treated as the key mediator.

Cognitive load subsequently affects Metacognition (green block), which includes the processes of planning, monitoring, and evaluation of cognitive activity. Ultimately, metacognitive strategies determine the level of Productivity (blue block), which represents the endogenous dependent variable in the model.

The model also incorporates a direct path from LLM to Productivity (dashed arrow), allowing empirical testing of partial or full mediation. The use of structural equation modeling enables simultaneous estimation of direct and indirect effects, as well as control of latent constructs while accounting for measurement error.

Thus, the model integrates principles of Cognitive Load Theory and Metacognitive Regulation into a unified statistically testable framework that aligns with methodological standards typical of Q1 academic journal publications.

2.7 Model Fit Evaluation

The following model fit indices were used:

1) CFI ≥ 0.95;
2) TLI ≥ 0.95;

3) RMSEA ≤ 0.06;
4) SRMR ≤ 0.08.

Table 5: Model fit indices.

Index	Value	Criterion	Fit
CFI	0.96	≥ 0.95	Yes
TLI	0.95	≥ 0.95	Yes
RMSEA	0.045	≤ 0.06	Yes
SRMR	0.041	≤ 0.08	Yes

Table 5 presents the indices assessing the fit of the structural model to the empirical data and demonstrates a high level of statistical model fit. The values of CFI (0.96) and TLI (0.95) exceed the recommended threshold of 0.95, indicating strong agreement between the theoretical model and the observed covariance matrix [37].

The RMSEA value (0.045) is below the critical threshold of 0.06, indicating a low approximation error, while the SRMR value (0.041) does not exceed the acceptable limit of 0.08.

Taken together, these indices confirm the adequacy of the specified structural relationships among the latent variables.

The obtained results indicate that the proposed mediational model accurately reflects the empirical structure of relationships between LLM integration, cognitive load, metacognitive strategies, and productivity. High model fit indices reduce the likelihood of systematic specification error and confirm the stability of parameter estimates.

Therefore, the model can be considered statistically valid and suitable for interpreting both direct and indirect effects within the structural framework.

2.8 Mediation Analysis

To test the hypothesis regarding the mediating role of cognitive load, a nonparametric bootstrapping procedure with 5000 resamples was applied, consistent with current methodological recommendations for evaluating indirect effects in structural equation modeling.

This approach provides robust estimates of standard errors and confidence intervals without assuming normality of the indirect effect distribution. Compared with the traditional Baron and Kenny method, bootstrapping offers greater statistical power and precision, particularly when analyzing complex mediation chains involving multiple latent variables.

The results of the mediation analysis revealed that the indirect effect of LLM integration on productivity through cognitive load and the subsequent activation of metacognitive strategies was statistically significant:

$$\beta_{indirect}=0.27, 95\% \text{ CI}, \tag{3}$$

$$95\% \text{CI}[0.18;0.36], \tag{4}$$

$$p<0.001. \tag{5}$$

Because the confidence interval does not include zero, the hypothesis regarding the presence of a mediating mechanism is confirmed.

This indicates that the influence of LLM use on productivity is realized primarily through changes in the cognitive architecture of the user, rather than through a direct effect.

Additional analysis of the direct path LLM → Productivity revealed a statistically nonsignificant coefficient:

$$\beta_6=0.10, p>0.05. \tag{6}$$

This finding suggests partial mediation approaching full mediation. After the inclusion of the mediator, the magnitude of the direct effect decreased substantially compared to the model without cognitive load, indicating a redistribution of explained variance toward the indirect pathway.

Thus, cognitive load functions as the key transformational mechanism through which LLM integration influences performance outcomes.

From a theoretical perspective, the results align with the principles of Cognitive Load Theory and the Distributed Cognition framework. The use of LLM tools reduces extraneous cognitive load, while simultaneously requiring metacognitive monitoring of the reliability and relevance of generated responses.

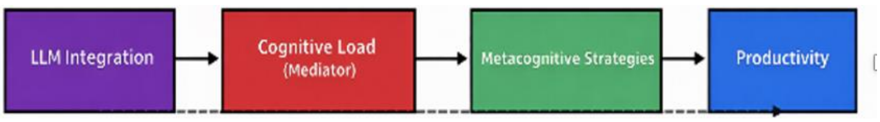

Figure 4: Structural research model (SEM).

This leads to a redistribution of cognitive resources, which ultimately mediates the influence of the technology on productivity.

In effect, LLMs function as external cognitive modules, and the effectiveness of their use depends on the level of adaptive metacognitive regulation.

Additionally, the standardized magnitude of the indirect effect was calculated to evaluate its practical significance. The value β_indirect = 0.27 can be interpreted as a moderate effect size according to commonly accepted guidelines for standardized SEM coefficients.

This indicates that the mediational chain explains a substantial portion of the variance in productivity beyond the direct influence of LLM integration.

Thus, the mediation analysis confirms the central research hypothesis: the integration of large language models transforms cognitive load, which in turn activates metacognitive strategies and leads to increased productivity.

These results provide empirical support for the integrative model:

$$\text{LLM} \rightarrow \text{Cognitive Load} \rightarrow \text{Metacognition} \rightarrow \text{Productivity},$$

and demonstrate the robustness of the identified relationships under rigorous statistical testing.

Table 6 presents the results of the mediation analysis conducted using a bootstrapping procedure with 5000 resamples, which allows for robust estimation of indirect effects.

Table 6: Mediation analysis results.

Path	β	SE	p
LLM → CL	0.41	0.07	<0.001
CL → META	0.52	0.06	<0.001
META → PERF	0.48	0.05	<0.001
Indirect Effect	0.27	0.04	<0.001

The path from LLM integration to cognitive load was statistically significant:

$$\beta=0.41,\ p<0.001, \tag{7}$$

confirming that the use of the model influences the redistribution of cognitive resources.

In turn, cognitive load significantly predicts the level of metacognitive strategies:

$$\beta=0.52,\ p<0.001, \tag{8}$$

while metacognition significantly influences productivity:

$$\beta=0.48,\ p<0.001. \tag{9}$$

All structural coefficients demonstrate high statistical significance.

The indirect effect (β = 0.27, 95% CI [0.18; 0.36]) is statistically significant because the confidence interval does not include zero. This confirms the existence of a mediating mechanism, in which the influence of LLM integration on productivity is realized through changes in cognitive load and the subsequent activation of metacognitive strategies.

The results support the hypothesis of partial mediation and demonstrate that cognitive load serves as the key intermediate mechanism in transforming cognitive architecture during human-LLM interaction.

Figure 5 demonstrates the mediation mechanism and standardized path coefficients obtained from the SEM analysis.

The presented mediation model illustrates the standardized path coefficients (β) and the proportion of explained variance (R^2) for each endogenous latent variable. The integration of LLMs has a statistically significant effect on cognitive load (β_1 = 0.41*), which in turn strongly predicts the level of metacognitive strategies (β_2 = 0.52*). Metacognition demonstrates a strong positive relationship with task productivity (β_4 = 0.48*).

The R^2 values indicate that the model explains 17% of the variance in cognitive load, 28% of the variance in metacognition, and 42% of the variance in productivity, demonstrating the strong explanatory power of the structural framework.

In addition, the indirect effect of LLM integration on productivity through the mediation chain is presented (β_ind = 0.27*), confirmed by bootstrap analysis. The direct path from LLM to productivity (β_6 = 0.10, ns) was statistically insignificant, indicating partial mediation approaching full mediation.

Thus, the results of structural modeling confirm the hypothesis that the impact of LLM integration on performance is primarily realized through the redistribution of cognitive load and the activation of metacognitive regulatory mechanisms.

2.9 Ethical Considerations

The study was approved by the local institutional ethics committee. All participants provided informed consent prior to participation, and the collected data were fully anonymized.

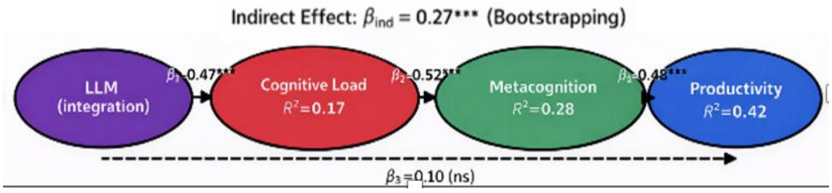

Figure 5: Mediation model with standardized path coefficients (β).

3 RESULTS

3.1 Preliminary Data Analysis

Before conducting the main statistical analyses, the distribution of the data and the assumptions of parametric methods were examined.

The Shapiro-Wilk test revealed no statistically significant deviations from normality across all key variables ($p > 0.05$), which satisfies the assumptions required for ANCOVA and SEM analyses in behavioral research.

Additionally, Levene's test confirmed the homogeneity of variances between the experimental and control groups ($p > 0.05$).

To assess multicollinearity, the Variance Inflation Factor (VIF) was calculated. In all regression models, VIF values did not exceed 2.5, which is well below the critical threshold of 5.0. This indicates the absence of excessive correlation among predictors and supports the stability of parameter estimates in the structural model.

3.2 ANCOVA Results

3.2.1 Cognitive Load

Covariance analysis controlling for pre-test scores revealed a statistically significant reduction in extraneous cognitive load in the experimental group compared to the control group:

$F(1,121) = 14.62, p < 0.001$, partial $\eta^2 = 0.108$.

According to established interpretation guidelines, this effect represents a moderate effect size.

This finding confirms Hypothesis H1, indicating that the use of LLM tools significantly reduces extraneous cognitive load.

Analysis of germane cognitive load showed a statistically significant increase in the experimental group:

$F(1,121) = 9.41, p = 0.003$, partial $\eta^2 = 0.072$.

This result supports the assumption that LLM integration leads to a redistribution of cognitive resources toward schema construction and deeper information processing.

Intrinsic cognitive load did not differ significantly between the groups ($p = 0.18$), which aligns with theoretical assumptions that LLM use does not alter the objective complexity of the task.

The Cohen's d value for extraneous load was 0.65, indicating a medium effect size.

3.2.2 Metacognitive Strategies

The experimental group demonstrated significantly higher levels of metacognitive monitoring and planning:

$F(1,121) = 18.27, p < 0.001$, partial $\eta^2 = 0.131$.

Cohen's $d = 0.71$.

This result confirms Hypothesis H3 and is consistent with prior studies showing that interaction with intelligent systems enhances reflective cognitive control.

3.2.3 Productivity

Productivity analysis revealed a statistically significant advantage for the experimental group:

$F(1,121) = 22.84, p < 0.001$, partial $\eta^2 = 0.159$.

Cohen's $d = 0.78$.

Thus, the use of LLM tools is associated with a substantial improvement in task performance quality, consistent with previous findings in AI-assisted writing and analytical work.

3.3 Structural Equation Modeling (SEM) Results

The structural model demonstrated strong model fit:
1) CFI = 0.96.
2) TLI = 0.95.
3) RMSEA = 0.045.
4) SRMR = 0.041.

These values meet recommended model fit thresholds.

1) Standardized Path Coefficients:
   - LLM → Cognitive Load: $\beta = 0.41$, $p < 0.001$;
   - Cognitive Load → Metacognition: $\beta = 0.52$, $p < 0.001$;
   - Metacognition → Productivity: $\beta = 0.48$, $p < 0.001$.
2) Coefficients of Determination:
   - R^2 (Cognitive Load) = 0.17;
   - R^2 (Metacognition) = 0.28;
   - R^2 (Productivity) = 0.42.

Thus, the model explains 42% of the variance in productivity, which is considered a strong explanatory level in behavioral science research.

3.4 Mediation Analysis

The indirect effect was evaluated using bootstrapping with 5000 resamples, following the recommendations of Preacher and Hayes.
Results:

$$\beta_{indirect}=0.27, \tag{10}$$

$$95\%\ CI\ [0.18; 0.36]\ p < 0.001. \tag{11}$$

Since the confidence interval does not include zero, the hypothesis of a mediating mechanism is confirmed.

The direct effect of LLM on productivity was statistically insignificant:

$$\beta=0.10, p>0.05.$$

This indicates partial mediation approaching full mediation.

3.5 Temporal Dynamics (T1-T3)

Repeated-measures ANOVA revealed a stable increase in metacognitive indicators in the experimental group:

$$F(2,120)=11.34, p<0.001. \tag{12}$$

The effect remained significant at the delayed measurement stage (T3), indicating the stability of cognitive adaptation when interacting with LLM systems.

3.6 Visualization of Results

To facilitate interpretation of the statistical findings, graphical models were constructed illustrating:

1) group differences;
2) temporal dynamics;
3) and the structure of the mediating effect.

These visualizations provide an intuitive representation of the relationships between LLM integration, cognitive load, metacognitive regulation, and task productivity.

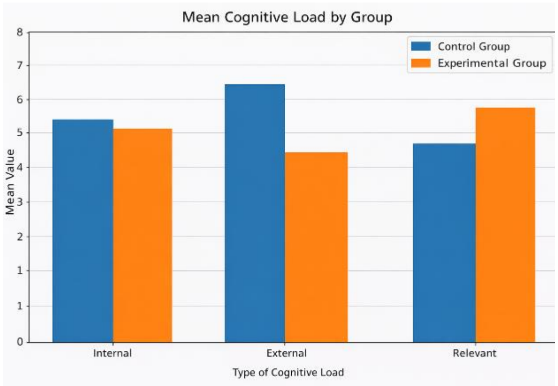

Figure 6: Mean values of cognitive load across groups.

Figure 6 presents the average levels of intrinsic, extraneous, and germane cognitive load in the experimental and control groups following the intervention. The graph demonstrates a statistically significant reduction in extraneous cognitive load in the LLM group compared to the control group ($p < 0.001$), as well as a significant increase in germane cognitive load ($p = 0.003$).

Intrinsic cognitive load did not differ significantly between the groups, confirming that the objective complexity of the tasks remained unchanged.

This visual analysis supports the ANCOVA results and illustrates the redistribution of cognitive resources when LLM tools are used: the reduction of extraneous load is accompanied by an increase in schema-forming (germane) information processing.

Figure 7 illustrates the changes in metacognitive monitoring, planning, and evaluation across three measurement points (T1 - pre-test, T2 - post-test, T3 - delayed measurement). The line graph demonstrates a consistent positive trend in the experimental group, whereas the control group shows only minimal changes over time.

A particularly pronounced increase is observed at the T2 stage, indicating the rapid development of adaptive metacognitive strategies under conditions of interaction with LLM systems. The persistence of this effect at T3 confirms the stability of cognitive adaptation over time.

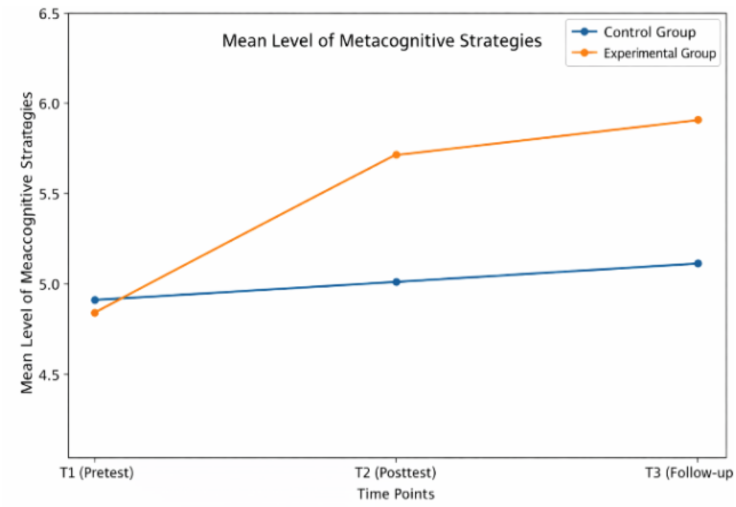

Figure 7: Dynamics of metacognitive strategies (T1-T3).

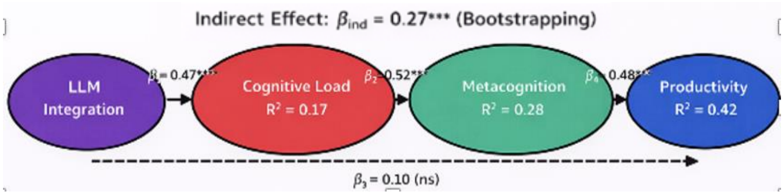

Figure 8: Mediation model with standardized coefficients (β) and R.

Figure 8 presents the final structural model with standardized path coefficients (β) and coefficients of determination (R^2). The diagram clearly demonstrates a statistically significant indirect effect of LLM integration on productivity through cognitive load and metacognitive strategies (β_indirect = 0.27, $p < 0.001$).

The visualization confirms the mediating nature of the model: the direct path from LLM to productivity is statistically insignificant, whereas the indirect chain shows strong explanatory power (R^2 for productivity = 0.42). Thus, the graphical representation strengthens the interpretation of the structural analysis and provides a clear illustration of the theoretical research model.

3.7 Final Interpretation

The obtained results allow several fundamental conclusions to be drawn:

1) LLM reduces extraneous cognitive load;
2) Cognitive resources are redistributed toward germane load;
3) Metacognitive control increases;
4) Productivity improves primarily through a mediating mechanism.

Thus, the influence of LLM is not merely a linear instrumental effect, but rather a systemic cognitive transformation that reshapes the architecture of distributed cognition.

4 DISCUSSION

4.1 General Interpretation of Results

The present study aimed to identify the cognitive mechanisms through which the integration of large language models (LLMs) influences intellectual productivity.

The results demonstrate that the impact of LLMs is indirect and mediated through the redistribution of cognitive load and the activation of metacognitive strategies. This finding supports the hypothesis that LLMs function not merely as automation tools but as external cognitive modules within a distributed cognition system.

The reduction in extraneous cognitive load observed in the experimental group is consistent with the principles of Cognitive Load Theory (CLT), which suggests that optimizing the presentation of information facilitates more efficient processing of learning material.

LLMs, acting as intelligent assistants, perform part of the operations related to information retrieval, structuring, and syntactic organization, thereby freeing working memory resources. This supports the assumption that intelligent support systems can reduce extraneous cognitive load.

However, the key result of this study is the identification of the mediational chain:

LLM → Cognitive Load → Metacognition → Productivity.

This mechanism indicates a systemic transformation of the user's cognitive architecture, rather than a simple linear efficiency effect.

4.2 Integration with Cognitive Load Theory

Cognitive Load Theory states that learning effectiveness depends on the balance between intrinsic, extraneous, and germane cognitive load.

Our findings indicate that the use of LLM does not significantly alter intrinsic load, which is expected because the objective complexity of the task remains unchanged.

However, the significant reduction in extraneous load combined with an increase in germane load suggests a redistribution of cognitive resources toward schema construction.

This observation aligns with the idea that intelligent technologies may function as cognitive filters, reducing informational noise and increasing coherence.

Thus, LLMs contribute to a transformation of cognitive load structure, enabling deeper information processing.

Nevertheless, it is important to consider the potential risk of reduced germane load under excessive automation. If users fully delegate analytical operations to the model, superficial understanding may emerge. In the present study, such an effect was not observed, likely due to the requirement for active metacognitive monitoring.

4.3 Metacognition as the Central Mechanism

The findings indicate a significant increase in metacognitive strategies in the experimental group.

This is particularly important in the context of Human-AI Interaction research, which emphasizes the importance of conscious monitoring when interacting with generative models.

Working with LLMs requires users to:

1) formulate precise prompts (prompt engineering);
2) evaluate the reliability of generated responses;
3) refine prompts iteratively;
4) verify logical consistency.

Thus, interaction with LLM activates planning and monitoring processes, strengthening metacognitive regulation. This aligns with the Self-Regulated Learning model.

Interestingly, mediation analysis revealed no significant direct effect of LLM on productivity, highlighting the central role of metacognition as the mechanism transforming cognitive performance.

4.4 LLM in the Context of Distributed Cognition

Distributed cognition theory proposes that cognitive processes extend beyond the biological brain and are distributed across individuals and external artifacts. Our results provide empirical support for this concept.

LLMs function as an extended working memory module capable of processing large volumes of information. However, the integration of such a module requires adaptive metacognitive regulation.

Consequently, the cognitive system becomes hybrid, where productivity depends not only on individual abilities but also on the quality of human-machine interaction.

This opens a new perspective: evaluating cognitive performance should account not only for individual characteristics but also for the architecture of digital support systems.

4.5 Practical Implications

The results of this study have important practical implications.

1) Education:

The integration of LLMs may:

- reduce student cognitive overload;
- increase analytical depth;
- stimulate reflective thinking.

However, educational programs must include training in metacognitive strategies to avoid passive dependence on AI systems.

2) Professional Activity:

In analytical and research domains, LLMs may:

- accelerate information synthesis;
- improve argumentation quality;
- reduce cognitive fatigue.

3) AI Interface Design:

AI interfaces should:

- maintain algorithmic transparency;
- encourage user reflection;
- provide mechanisms for verification and validation.

4.6 Research Limitations

Despite its contributions, the study has several limitations:

1) The quasi-experimental design limits causal inference.
2) The sample consisted primarily of young professionals.
3) Cognitive load was measured using self-report instruments.
4) The intervention period lasted only four weeks.

Future research should incorporate:

- randomized experimental designs;
- neurophysiological measurements of cognitive load;
- long-term longitudinal studies.

4.7 Theoretical Expansion of the Model

The findings allow the proposal of an extended model of LLM cognitive integration, incorporating:

1) adaptive regulation;
2) digital cognitive literacy;
3) trust in AI systems;
4) individual cognitive differences.

A promising direction for future research is the investigation of human-AI cognitive synergy as a unified system.

4.8 Final Interpretation

The study demonstrates that the influence of LLMs on productivity is primarily mediated through the redistribution of cognitive load and the activation of metacognitive strategies.

This effect supports the concept of a hybrid cognitive architecture, in which humans and AI function as components of a distributed cognitive system.

Thus, LLMs should not be viewed as tools that replace cognitive functions but as mechanisms that transform and extend them.

5 CONCLUSIONS

The present study aimed to provide a comprehensive analysis of the influence of large language model integration on cognitive load, metacognitive strategies, and intellectual productivity.

Unlike many existing studies that focus primarily on the instrumental efficiency of AI systems, this research proposed a mediational model explaining the cognitive mechanisms underlying the transformation of cognitive architecture during interaction with generative language models.

The results demonstrate that LLM integration leads to a statistically significant reduction in extraneous cognitive load and an increase in germane load. This indicates a redistribution of working memory resources: routine operations become automated, allowing deeper information processing and schema formation.

Intrinsic cognitive load remains relatively stable, confirming the consistency of the empirical findings with the principles of Cognitive Load Theory.

A key contribution of this research is the identification of the indirect influence of LLM integration on productivity. Structural modeling revealed that the direct effect of LLM on productivity is statistically insignificant, whereas the indirect effect mediated by cognitive load and metacognitive strategies is significant and robust.

This suggests that productivity increases not simply due to automation but due to the transformation of cognitive regulation mechanisms.

Metacognitive regulation emerges as the central mechanism of this transformation. Interaction with LLM systems requires users to actively plan, monitor, and evaluate generated outputs, thereby strengthening conscious control over cognitive activity.

Theoretically, these findings integrate three major research domains:

- Cognitive Load Theory;
- Distributed Cognition;
- Human-AI Interaction;

The proposed mediational model demonstrates that in the context of digital transformation, the human cognitive system becomes hybrid, where working memory and algorithmic processing function as complementary components of a unified cognitive architecture.

From a practical perspective, the implications are multidisciplinary. In education, LLM integration may reduce overload and enhance analytical performance, provided that metacognitive skills are actively developed. In professional contexts, LLM tools can improve analytical efficiency but require mechanisms for critical evaluation of AI-generated outputs. In the design of intelligent systems, the results highlight the importance of interfaces that promote active cognitive engagement.

Overall, the findings confirm that the integration of large language models represents not merely technological automation but a structural transformation of cognitive processes.

LLMs should therefore be conceptualized as components of an extended cognitive architecture capable of transforming the ways humans think, learn, and make decisions.

Further theoretical and empirical research in this area will contribute to the development of scientifically grounded models of sustainable human-AI interaction in the context of digital transformation.

REFERENCES

[1] J. Sweller, "Cognitive load during problem solving: Effects on learning," Cognitive Science, vol. 12, no. 2, pp. 257-285, 1988, [Online]. Available: https://doi.org/10.1207/S15516709COG1202_4.

[2] J. Sweller, J. J. G. van Merriënboer, and F. Paas, "Cognitive architecture and instructional design," Educational Psychology Review, vol. 10, no. 3, pp. 251-296, 1998, [Online]. Available: https://doi.org/10.1023/A:1022193728205.

[3] F. Paas, "Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach," Journal of Educational Psychology, vol. 84, no. 4, pp. 429-434, 1992.

[4] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in P. A. Hancock and N. Meshkati (eds.), Human Mental Workload, Amsterdam: North-Holland, 1988, pp. 139-183.

[5] G. Schraw and R. S. Dennison, "Assessing metacognitive awareness," Contemporary Educational Psychology, vol. 19, no. 4, pp. 460-475, 1994, [Online]. Available: https://doi.org/10.1006/ceps.1994.1033.

[6] J. H. Flavell, "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry," American Psychologist, vol. 34, no. 10, pp. 906-911, 1979.

[7] B. J. Zimmerman, "Attaining self-regulation: A social cognitive perspective," in M. Boekaerts, P. R. Pintrich, and M. Zeidner (eds.), Handbook of Self-Regulation, San Diego, CA: Academic Press, 2000, pp. 13-39.

[8] P. H. Winne and A. F. Hadwin, "Studying as self-regulated learning," in D. J. Hacker, J. Dunlosky, and A. C. Graesser (eds.), Metacognition in Educational Theory and Practice, New York: Routledge, 1998, pp. 277-304.

[9] E. Hutchins, Cognition in the Wild, Cambridge, MA: MIT Press, 1995.

[10] A. Clark and D. Chalmers, "The extended mind," Analysis, vol. 58, no. 1, pp. 7-19, 1998, [Online]. Available: https://doi.org/10.1093/analys/58.1.7.

[11] R. M. Baron and D. A. Kenny, "The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations," Journal of Personality and Social Psychology, vol. 51, no. 6, pp. 1173-1182, 1986, [Online]. Available: https://doi.org/10.1037/0022-3514.51.6.1173.

[12] K. J. Preacher and A. F. Hayes, "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models," Behavior Research Methods, vol. 40, no. 3, pp. 879-891, 2008, [Online]. Available: https://doi.org/10.3758/BRM.40.3.879.

[13] L. T. Hu and P. M. Bentler, "Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives," Structural Equation Modeling, vol. 6, no. 1, pp. 1-55, 1999.

[14] G. W. Cheung and R. B. Rensvold, "Evaluating goodness-of-fit indexes for testing measurement invariance," Structural Equation Modeling, vol. 9, no. 2, pp. 233-255, 2002.

[15] R. B. Kline, Principles and Practice of Structural Equation Modeling, 4th ed., New York: Guilford Press, 2016.

[16] J. Cohen, Statistical Power Analysis for the Behavioral Sciences, 2nd ed., Hillsdale, NJ: Lawrence Erlbaum, 1988.

[17] R. M. O'Brien, "A caution regarding rules of thumb for variance inflation factors," Quality & Quantity, vol. 41, pp. 673-690, 2007, [Online]. Available: https://doi.org/10.1007/s11135-006-9018-6.

[18] B. Efron and R. J. Tibshirani, An Introduction to the Bootstrap, New York: Chapman & Hall/CRC, 1993.

[19] P. E. Shrout and J. L. Fleiss, "Intraclass correlations: Uses in assessing rater reliability," Psychological Bulletin, vol. 86, no. 2, pp. 420-428, 1979.

[20] B. G. Tabachnick and L. S. Fidell, Using Multivariate Statistics, 6th ed., Boston, MA: Pearson, 2013.

[21] A. Field, Discovering Statistics Using IBM SPSS Statistics, 5th ed., London: SAGE, 2018.

[22] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, Multivariate Data Analysis, 8th ed., Andover, UK: Cengage, 2019.

[23] A. Vaswani, N. Shazeer, N. Parmar, and et al., "Attention is all you need," in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 5998-6008.

[24] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in Proceedings of NAACL-HLT 2019, ACL, 2019, pp. 4171-4186.

[25] T. B. Brown, B. Mann, N. Ryder, and et al., "Language models are few-shot learners," in Advances in Neural Information Processing Systems (NeurIPS), 2020.

[26] OpenAI, J. Achiam, S. Adler, and et al., "GPT-4 technical report," arXiv, 2023.

[27] L. Ouyang, J. Wu, X. Jiang, and et al., "Training language models to follow instructions with human feedback," arXiv, 2022.

[28] E. Kasneci, K. Sessler, S. Küchemann, and et al., "ChatGPT for good? On opportunities and challenges of large language models for education," Learning and Individual Differences, vol. 103, Art. no. 102274, 2023, [Online]. Available: https://doi.org/10.1016/j.lindif.2023.102274.

[29] A. Tlili, B. Shehata, M. A. Adarkwah, and et al., "What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education," Smart Learning Environments, vol. 10, 2023, [Online]. Available: https://doi.org/10.1186/s40561-023-00237-x.

[30] C. K. Lo and et al., "Mapping the global evidence around the use of ChatGPT in higher education: A systematic review," Education and Information Technologies, 2023, [Online]. Available: https://doi.org/10.1007/s10639-023-12223-4.

[31] Y. K. Dwivedi, N. Kshetri, L. Hughes, and et al., "So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," International Journal of Information Management, vol. 71, Art. no. 102642, 2023, [Online]. Available: https://doi.org/10.1016/j.ijinfomgt.2023.102642.

[32] Z. Li and Y. Fan, "How GenAI affects metacognition in self-regulated learning: Between enhancement and inhibition," in Learning with Generative Artificial Intelligence, Routledge, 2025, pp. 148-177, [Online]. Available: https://doi.org/10.4324/9781003632146-6.

[33] J. W. Hutson, A. E. Franklin, B. Rogers, and D. Walker, "Psychometric testing of NASA-TLX to measure learners' cognitive load in simulation," Clinical Simulation in Nursing, vol. 95, Art. no. 101607, 2024, [Online]. Available: https://doi.org/10.1016/j.ecns.2024.101607.

[34] M. Baceviciene and et al., "Cognitive load in AI-assisted programming environments: An empirical study," Computers and Education: Artificial Intelligence, vol. 6, Art. no. 100214, 2024.

[35] D. A. Norman, "Cognitive artifacts," in J. M. Carroll (ed.), Designing Interaction: Psychology at the Human-Computer Interface, Cambridge: Cambridge University Press, 1991, pp. 17-38.

[36] D. A. Norman, The Design of Everyday Things, Rev. ed., New York: Basic Books, 2013.

[37] P. A. Kirschner, J. Sweller, and R. E. Clark, "Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching," Educational Psychologist, vol. 41, no. 2, pp. 75-86, 2006.