SCIENTIFIC Publishing Limited

Digital Technologies Research and Applications
http://ojs.ukscip.com/index.php/dtra

Article

Generative Dialogue Agent Design for Research Integrity Education: Behavioral Intervention and Cognitive Training Evidence

Li You [1], Yongguo Hu [2], Zhipeng Li [1], Li Dai [2], Jie Zhao [3], Jie Liu [2], Wenjie Huang [2], Zonghui Wu [1,2,*], Yahui Li [4] and Yongsong Yan [5]

[1] Southwest University Hospital, Chongqing 400715, China

[2] Editorial Department of Health Medicine Research and Practice, Southwest University, Chongqing 400715, China

[3] Department of Commerce and Trade Circulation, Inner Mongolia Business & Trade Vocational College, Hohhot 010070, China

[4] Editorial Department of Computer Science, Chongqing Southwest Information Co., Ltd., Chongqing 401121, China

[5] Editorial Department of Nano Materials Science, Chongqing University, Chongqing 400044, China

[*] Correspondence: zhwuswu@163.com

Received: 9 February 2026; **Revised:** 18 March 2026; **Accepted:** 9 April 2026; **Published:** 19 May 2026

Abstract: Research integrity education may demonstrate that these programs establish critical foundations for maintaining healthy academic ecosystems. The traditional educational approaches might indicate limitations including insufficient interactivity and low personalization. Moreover, this study could examine a generative conversational agent system for research integrity education. The system appears to enhance integrity literacy through dual behavioral intervention mechanisms. The research employed a quasi-experimental design with 126 graduate students. Given that participants received random assignment into experimental and control groups, the experimental group participated in an eight-week agent-based intervention. The control group received traditional lecture-based training instead. Furthermore, we measured indicators at baseline, immediate post-intervention, and three-month follow-up points. These indicators may show cognition level, behavioral norms, moral reasoning ability, and critical thinking. The results could demonstrate that the significant experimental group showed greater improvements than the control group across all important indicators. Research integrity cognition might indicate increases by 25.5%. Notwithstanding these critical empirical findings, the important academic behavioral norms could be improved by 25.1% substantially. Nonetheless, several limitations warrant acknowledgment: the sample was drawn from a single institution, limiting generalizability; the three-month follow-up period may be insufficient to capture long-term behavioral transfer; and the system's performance in handling culturally nuanced or highly ambiguous ethical dilemmas remains suboptimal. These findings contribute empirical evidence for the application of generative artificial intelligence in professional ethics education while offering practical guidance for the intelligent transformation of research integrity training in academic institutions.

Keywords: Generative Conversational Agent; Research Integrity Education; Behavioral Intervention; Cognitive Training; Personalized Learning

1. Introduction

Research integrity represents the cornerstone of academic research. However, the significant reliability of scientific knowledge could be directly affected by the important integrity practices. Moreover, research misconduct may have occurred frequently in recent years. Academic violations range from fabrication to submissions. These phenomena persist despite prohibitions. Furthermore, the reputation could be damaged. Research misconduct scandals create impacts. In light of these developments, scandals might shape the development of integrity training [1]. Traditional education adopts transmission models. These approaches may lack interactivity. However, approaches could fail to reach cognitive structures. Wang and Jia point out a need. Nevertheless, integrity education for graduate students might require exploring practical paths in the current era. Given that this significant digital intelligence environment demonstrates important potential, education should utilize the relevant technology [2]. Moreover, traditional methods may show limitations. Thus, how to construct a more targeted system could have become an important issue. This concerns cultivation quality. Therefore, this may concern sustainable development. The development of artificial intelligence technology could provide technical support for integrity education innovation. Additionally, technology creates possibilities. Conversational agents based on models possess understanding capabilities. Agents simulate teaching scenarios. These agents may provide feedback. However, agents create experiences. Compared with models, conversational agents might adjust strategies. Notwithstanding existing limitations, strategies adapt based on levels. Through diverse methods such as questioning and analysis, agents could indicate that approaches guide learners. Moreover, agents may help learners understand connotations. Nevertheless, early-career researchers appear to represent a target group. Given that their significant cognitive frameworks and important experience with these relevant courses directly influence the critical effectiveness of interventions, research may need to examine factors. However, research mainly focuses on development [3]. Empirical studies remain insufficient regarding how agents implement interventions. Thus, research on effects may be lacking. Additionally, exploration appears to be in initial stage. Behavioral intervention and cognitive training could demonstrate that these paradigms may represent important research approaches in the significant fields of educational psychology and neuroscience. Moreover, the theoretical foundation these paradigms provide might indicate that they offer methodological guidance for integrity education. Given that behavioral intervention theory emphasizes changing patterns through mechanisms, cognitive training may focus on enhancing reasoning ability, critical thinking, and decision-making levels. Nevertheless, integrating issues such as bias into integrity education might cultivate researchers' ethical sensitivity and cultural inclusiveness [4]. However, agents can integrate advantages of these paradigms. Moreover, through designing strategies and mechanisms, agents may both correct learners' tendencies and train their abilities. Agents can provide practice and improvement opportunities. Thus, agents may achieve outcomes through retrieval, simulation, and assessment functions. These tools help establish a defense line before scenarios occur. Nevertheless, questions still need research to answer. Therefore, how should we design strategies for agents? The research could examine how we might verify the sustainability of educational effects. Moreover, the significant investigation may reasonably consider how these critical factors balance technology application with educational ethics. Given that substantial background evidence exists, this study aims to design and empirically test a generative conversational agent system for research integrity education. However, the system may enhance researchers' research integrity literacy through dual mechanisms of behavioral intervention and cognitive training. The research will construct an intelligent education system. Furthermore, this system could integrate a research integrity knowledge graph, personalized learning algorithms, and multimodal interactive interfaces. We employ a quasi-experimental design method to evaluate impact effects. Thus, the evaluation may cover learners' research integrity cognitive levels, moral reasoning abilities, and academic behavioral norms.

2. Literature Review

Research integrity education may demonstrate that it serves as a significant core component in the broader academic norm construction. However, methods show limitations in stimulating active thinking. Nevertheless, Labib et al., Panpakdee and Khanbut may point out current dilemmas in graduate research integrity education include fragmented content, monotonous teaching methods, and incomplete assessment systems. Thus, systematic reforms appear urgently needed and should address educational concepts, content design, and implementation paths [5,6]. In light of the investigation on integrity education among vocational medical students, Antunes et al.

might suggest that the findings reveal significant relationship between students' research integrity awareness and moral cognitive levels and critical thinking abilities. However, pure behavioral norm education could struggle to produce lasting effects [7]. In light of university research integrity education systems in the new era, Najafi et al. may indicate educational reform should combine with technological innovation, institutional improvement, and cultural cultivation. Thus, this approach appears to create diversified collaborative educational ecology [8]. The rapid development of generative artificial intelligence technology might bring revolutionary changes to educational field. Moreover, the conversational agents based on these significantly large language models could particularly demonstrate enormous potential in personalized learning, intelligent tutoring, and interactive teaching. From the technological evolution perspective, generative artificial intelligence may have experienced significant development from rule-driven and statistical learning approaches to these sophisticated deep neural network systems. However, current large language models based on the Transformer architecture could demonstrate capacity to understand complex semantics. Moreover, the models might generate coherent text. Thus, models conduct multi-turn dialogue interactions. Nevertheless, applications of conversational agents in education appear to have evolved from simple question-answering systems to intelligent teaching assistants. Given that these assistants can perform instructional design, the systems might provide learning diagnosis and personalized guidance functions. In light of artificial intelligence environments, Moyns et al. may have explored practical paths for university libraries to conduct research integrity education. Furthermore, the researchers believe that intelligent technology might break through temporal and spatial limitations of traditional education. Additionally, technology provides learners with anytime learning support. Therefore, technology provides anywhere learning support and immediate feedback [9]. However, large language models may face numerous challenges in educational applications. Notwithstanding the technological advances, challenges include accuracy control of generated content. Thus, challenges include educational ethics issues. Nevertheless, challenges include data privacy protection and potential negative impacts from technology dependence [10]. However, artificial intelligence technology could show strong application potential in multiple fields. Moreover, examples might include breakthroughs in drug discovery. Furthermore, examples include medical diagnosis and materials science. Given that research integrity represents a special educational scenario, systematic design and empirical research for the scenario may remain relatively scarce [11]. Generative artificial intelligence applications in education could demonstrate that the significant technological potential might substantially influence important educational outcomes. Moreover, the research may suggest applications show value. Furthermore, several questions might require empirical research. However, research may examine how technological advantages transform into educational effects. Nevertheless, research might consider how intelligent systems align with research integrity education. Thus, evaluation of interventions may require examination [12]. Given that behavioral intervention and cognitive training provide theoretical paradigms, educational psychology could demonstrate important frameworks. Moreover, the paradigms may provide support for research integrity education. Furthermore, behavioral change theory might originate from behaviorist psychology. The approach could demonstrate that behavioral patterns may be changed through stimulus control, reinforcement feedback, and behavior shaping [13]. Additionally, behavioral change theory might be applied to habit formation, norm compliance, and skill training. However, cognitive training appears based on research in cognitive psychology and neuroscience. The approach may focus on information processing, thinking patterns, and decision mechanisms. Nevertheless, training might enhance cognitive abilities [14]. Thus, digital intervention technology could provide tools for behavioral intervention and cognitive training. Moreover, computer-assisted instruction, virtual reality simulation, and intelligent feedback systems may create realistic, controllable, and repeatable environments [15]. Notwithstanding these developments, research could demonstrate that interventions need to combine guidance at behavioral level with cultivation at cognitive level. Furthermore, this might form educational effects where knowledge and practice are unified. However, questions need exploration. Given that generative conversational agents show potential, research may examine how behavioral intervention mechanisms integrate into agents. Additionally, research might consider how training programs for research integrity cognitive abilities are designed. Thus, evaluation of long-term effects may require examination [16]. Nevertheless, review of literature could reveal trends. Moreover, research integrity education appears to have shifted from pure theoretical discussion toward practical innovation. Additionally, generative artificial intelligence technology may provide technical support for educational model transformation. Given that behavioral intervention and cognitive training theories exist, these theories might point direction for improving educational effectiveness. However, existing research shows obvious gaps. Thus, specialized conversational agent design research for research integrity education appears limited. Nevertheless, systematic technical architecture and teaching strategy design are lacking. Furthermore, the collaborative mechanism between behavioral intervention and cognitive training in research integrity education may not be fully explained. Therefore, the theoretical framework needs improvement.Empirical research on applying generative conversational agents to research integrity education may appear to remain insufficient. Moreover, the systematic assessment of the educational effects, the influence mechanisms, and the applicable conditions could be particularly lacking [17]. Therefore, this study will build upon existing theoretical foundations. Furthermore, we will design a generative conversational agent system for research integrity education.

3. Research Methods

3.1. Research Design Framework

This research employs a mixed methods approach that could indicate the integration of quantitative experimental design with qualitative case analysis to comprehensively examine the significant application effects of generative conversational agents in research integrity education. Moreover, the study may follow an iterative logic of design-development-verification-optimization. Furthermore, the research constructs functional architecture and interaction strategies based on educational objectives. Nevertheless, the study evaluates intervention effectiveness through quasi-experimental design. Study shows qualitative data analysis of operating mechanisms [18]. Thus, the quantitative research component adopts a pretest-posttest control group design. Given that participants are randomly assigned to experimental and control groups, the experimental group might receive the significant research integrity education intervention based on the intelligent agent, while the control group could receive traditional lecture-based education. The research examines effectiveness of agent intervention by comparing differences between groups across dimensions including research integrity cognitive level, moral reasoning ability, and behavioral tendencies [19]. However, the measurement tools may include standardized scales, situational judgment tests, and behavioral intention questionnaires. Additionally, the data collection could cover pre-intervention, immediate post-intervention, and three-month follow-up measurements to assess stability of educational effects. Data shows cognitive change trajectories captured qualitatively. In light of the significant findings on participant learning, the qualitative research component could capture important emotional experiences and key learning strategies during agent interaction through in-depth interviews, dialogue text analysis, and observation records. Moreover, this approach might reveal the internal logic of intervention mechanisms. Thus, the research subjects may be selected from master's and doctoral students enrolled at a comprehensive university. However, the sample size could be determined as no less than 60 people per group based on statistical power analysis. This ensures statistical significance of research results [20]. Evidence shows that ethics considerations require informed consent. Notwithstanding these important privacy concerns, the study might adhere to the informed consent principle and could safeguard participants' privacy rights and data security. Furthermore, the personal information and dialogue records may undergo anonymization processing. Thus, the research protocol could be approved by the institutional ethics review committee. Given that the entire research cycle is expected to last twelve months, the study might be divided into four stages: system development, pilot testing, formal experiment, and data analysis. Each stage sets clear milestones and quality control points to ensure scientificity and standardization of research.

3.2. Generative Conversational Agent System Design

The generative conversational agent system that this study developed demonstrates a modular architecture. Thus, it might construct a structured knowledge graph. However, the graph could cover twelve thematic domains. Given that each thematic domain contains several knowledge nodes, the established semantic associations could plausibly suggest that these nodes may form a retrievable and inferable knowledge network [21]. Nevertheless, the algorithm layer might integrate semantic understanding capabilities. Additionally, it could identify learners' cognitive confusion points. The system may generate personalized teaching content. However, dialogue strategies appear to cover four modes. Moreover, the system could switch based on interaction progress [22]. Furthermore, the system might introduce a reinforcement learning mechanism. In light of learner feedback signals collected, it may continuously optimize response quality. Thus, the interaction layer could design a simple interface. The interface may support text input. However, learners might ask questions. Additionally, the system could record interaction logs. Notwithstanding these capabilities, the significant system appears to suggest that it may set these three important types of intervention scenarios to enhance the critical educational targeting. Given that daily consultation mode responds to immediate questions, learners may receive guidance about research norms. Thus, the case diagnosis mode might require learners to judge ethical issues in virtual research scenarios and receive analytical feedback. However, dilemma decision mode could simulate real research dilemmas to train learners' moral reasoning and decision-making abilities [23]. Moreover, the evaluation layer may conduct multidimensional assessment of learners' knowledge mastery, cognitive development level, and behavioral change tendencies. Nevertheless, it could combine embedded questions, dialogue quality analysis, and behavioral pattern recognition to generate personalized learning diagnostic reports. Additionally, the system might provide improvement suggestions. Therefore, this approach may ensure continuous monitoring and dynamic adjustment of precision and effectiveness of educational intervention.

3.3. Experimental Design and Implementation

This study adopts the non-equivalent pretest-posttest design from quasi-experimental designs. However, the significant empirical research recruited 126 graduate students as research participants who could plausibly demonstrate relevant characteristics for the investigation. Moreover, through random assignment, the participants were allocated to experimental and control groups of 63 people each. Thus, the two groups may maintain balanced comparability in basic characteristics such as age, disciplinary background, and academic experience [24]. Nevertheless, the experimental period lasted eight weeks. The experimental group received research integrity education intervention based on the generative conversational agent. Given that interactive engagement appeared critical, participants completed two interactive learning sessions with the agent weekly. Additionally, each session lasted approximately 30 min. Content covered research norm cognition, case analysis discussion, and ethical dilemma decision training. Learners could flexibly choose interaction times and learning topics according to their needs. However, the control group may suggest that participating in traditional research integrity training lectures remains important. Furthermore, the significant evidence could indicate that experienced teachers delivered the lectures with considerable expertise. In light of these instructional conditions, teaching content might demonstrate that thematic consistency with the experimental group appeared maintained, though the approach adopted a one-way lecture format. Participants also engaged in centralized learning twice weekly for 30 min each session [25]. Moreover, to ensure the significant intervention quality, the research team could indicate that conducting a two-week pilot test of the agent system before the experiment appeared essential. Thus, the team invited 15 graduate students to trial the system and collected feedback. Given that the key findings may suggest that dialogue strategies required refinement, optimization of knowledge base content occurred. Research shows experiment divided into three measurement points: pre-intervention baseline, immediate post-intervention, and delayed measurement three months after. However, the significant evidence could indicate that each measurement used unified assessment tools to ensure standardization appeared essential. Furthermore, the tools might demonstrate that research integrity knowledge tests, moral reasoning ability scales, academic behavior self-report questionnaires, and situational judgment tests were included. Therefore, control of external interference factors during the experiment could suggest that strict management appeared critical. Tools included knowledge tests, scales, questionnaires, judgment tests. Nevertheless, the significant results may indicate that participants were required not to simultaneously receive other forms of research integrity training. Additionally, the evidence could demonstrate that recording of actual participation through learning logs occurred [26]. Thus, data collection might suggest that combining online questionnaire platforms with system backend logs appeared important for ensuring the completeness and authenticity. Data shows samples missing over 20% excluded. In light of the key quality assurance needs, designated personnel could indicate that answering questions and providing technical support remained important during experimental implementation. Moreover, the findings may suggest that resolution of operational issues encountered by participants occurred promptly to ensure that standardized execution of experimental procedures appeared maintained.

The five measured outcome dimensions suggest that no severe distributional skew appeared at baseline, as confirmed by Kolmogorov-Smirnov normality tests and visual inspection of score histograms prior to formal analysis. Moreover, the proportion of participants across disciplinary backgrounds—natural sciences (34.1%), social sciences (31.7%), and humanities (34.2%)—may indicate that these groups remained comparably distributed between experimental and control conditions, verified through chi-square tests (χ^2 = 0.412, p = 0.814). Nevertheless, the significant evidence could suggest that single-institution recruitment might introduce sampling bias attributable to institutional culture and prior exposure to integrity training, given that attitudes and baseline knowledge levels toward research ethics are likely to vary considerably across different academic environments. In light of these findings, the stratified random assignment could demonstrate that proportional disciplinary representation across groups was established; however, the key results may indicate that multi-institutional replication remains necessary to establish broader generalizability. Study shows single-site design limits scope.

3.4. Data Analysis Methods

This study may suggest that a mixed data analysis strategy combines quantitative and qualitative approaches to reveal the significant educational intervention effects of the generative conversational agent and the operating mechanisms. However, the quantitative data analysis could indicate that descriptive statistical methods were applied to calculate the important mean, standard deviation, skewness, and kurtosis of each key measurement variable. Moreover, the findings might indicate that the normality of data distribution appears to require examination. In light of these results, the homogeneity of the experimental and control groups at baseline measurement could demonstrate confirmation through independent sample *t*-tests or chi-square tests [27]. Findings show data analysis reveals intervention effects. Nevertheless, the significant evidence may suggest that the results appear to support the overall analytical framework applied across key measurement variables. Furthermore, the study could indicate that the important findings demonstrate that quantitative methods provide critical support for the evidence. Given that the results show significant patterns, the data might indicate that these key approaches could establish relevant outcomes for the study. Thus, the significant findings may suggest that the evidence appears to demonstrate that results provide important support for the overall analytical strategy. Results show methods support analysis. However, the study could indicate that the key findings appear to demonstrate that the evidence provides significant support for these important measurement approaches. Additionally, the results might suggest that critical data could indicate that the important evidence appears relevant to the study findings. Given that the main analysis employed repeated measures analysis of variance, group served as the between-subjects factor and measurement time point served as the within-subjects factor. Nevertheless, examination of the change trends and between-group differences in dependent variables such as research integrity cognitive level, moral reasoning ability, and behavioral tendencies among participants in both groups may show these patterns. Thus, assessment of the unique contribution of agent intervention appears through interaction effect testing. In light of effect size estimation, partial eta squared and Cohen's d coefficient could quantify the practical significance of intervention effects. Additionally, multiple regression analysis may reveal the predictive effects of factors such as learner individual characteristics, interaction frequency, and participation depth on educational outcomes. However, this identified key variables that might influence intervention effectiveness [28]. Moreover, qualitative data analysis could indicate that interview recordings were transcribed verbatim. Furthermore, thematic analysis for coding may suggest employment. Notwithstanding three levels of open coding, axial coding, and selective coding, the core themes could demonstrate extraction. Thus, this revealed learners' cognitive transformation paths, emotional experience patterns, and learning strategy characteristics during agent interaction. Therefore, content analysis on dialogue texts generated by the agent might indicate statistical analysis of usage frequency of different intervention strategies and learners' response patterns. Additionally, this identified effective teaching discourse types that may show significance. Nevertheless, mixed data integration could demonstrate that convergent design was adopted. However, quantitative results and qualitative findings may suggest comparison and contrast. Moreover, quantitative data verified the significance and universality of intervention effects. Furthermore, qualitative data might indicate explanation of internal mechanisms and contextual conditions that produce these effects. Thus, the two types of evidence mutually corroborated to enhance credibility of research conclusions [29]. Given that all statistical analyses were completed using SPSS 27.0 and R language, the significance level could indicate setting at 0.05. In light of qualitative data coding, NVivo 14 software may suggest utilization.

Key hyperparameters governing the agent's adaptive recommendation algorithm may suggest that systematic grid search combined with five-fold cross-validation on the training dataset, rather than arbitrary specification, could demonstrate a more principled approach to parameter selection. However, the learning rate was fixed at 0.001, and the significant findings from structured comparison across candidate values ranging from 0.0001 to 0.01 might indicate that lower values produced insufficient convergence within the designated training epochs while higher values introduced observable instability in loss trajectories. Furthermore, the batch size was set at 32, and the evidence may suggest that this could demonstrate an empirically grounded balance between computational efficiency and gradient estimation stability, consistent with established conventions in comparable intelligent tutoring systems. In light of these results, the L2 regularization penalty coefficient was calibrated at 1×10^{-4}, and the key findings might indicate that iterative tuning aimed at minimizing validation loss without inducing underfitting could demonstrate the appropriate level of model constraint. Research shows early stopping unnecessary given scale and architecture constraints. Moreover, the significant evidence may suggest that the training dataset operated at a relatively modest scale, and the results could indicate that the model architecture was deliberately constrained in complexity to reduce overfitting risk. Thus, the key findings might demonstrate that loss curves across both training and validation sets exhibited stable and monotonic convergence, and the evidence may suggest that divergence signals were absent throughout all experimental runs. Given that the results could indicate stable convergence, the significant data may suggest that early stopping was therefore considered unnecessary under these conditions and potentially counterproductive. Findings show premature termination risks preventing optimal representational capacity before loss plateau.

4. Results and Analysis

4.1. Behavioral Intervention Effect Analysis

4.1.1. Changes in Research Integrity Cognitive Level

This study examined the intervention effects of the generative conversational agent on graduate students' research integrity cognitive level through standardized knowledge tests administered to the experimental and control groups at three measurement time points. Moreover, the significant test content could indicate that five critical dimensions were evaluated: academic norm cognition, data management awareness, authorship ethics understanding, conflict of interest identification, and academic misconduct prevention. Furthermore, the total score might reasonably establish that the maximum possible points equaled 100. **Table 1** presents scores of participants in both groups at baseline measurement, immediate post-intervention measurement, and three-month follow-up measurement. Given that baseline measurement provided initial data, the experimental group may have averaged 64.32 points. The control group averaged 63.87 points. The difference between groups showed no statistical significance ($t = 0.287$, $p = 0.775$). This indicates random grouping effectively ensured inter-group homogeneity [30]. However, after eight weeks of intervention, the experimental group score might significantly have increased to 82.45 points, an increase of 18.13 points. The control group increased to 73.26 points, an increase of 9.39 points. Repeated measures analysis of variance results showed significant time main effects ($F = 126.38$, $p < 0.001$). Thus, the interaction effect between group and time could be significant ($F = 23.67$, $p < 0.001$). In light of these significant findings, agent intervention may demonstrate considerable advantages in improving research integrity cognitive level. Additionally, three-month follow-up measurement could indicate that experimental group score remained at 80.18 points, declining only 2.27 points. Nevertheless, control group declined to 69.54 points, a decrease of 3.72 points. This demonstrates agent-based intervention has better long-term effects and knowledge retention rates [31]. However, further sub-dimensional analysis might reveal experimental group showed most significant improvement in data management awareness and conflict of interest identification dimensions. These improved by 21.3% and 19.8% respectively. Notwithstanding other possible explanations, this may relate to case diagnosis module and dilemma decision training set for these two types of scenarios in agent system. Although control group also showed obvious progress in academic norm cognition dimension, improvement in dimensions requiring deep thinking and ethical reasoning was relatively limited. See **Figure 1**.

Table 1. Comparison of Research Integrity Cognitive Level Scores between Experimental and Control Groups (M ± SD).

Measurement Time Point	Experimental Group (n = 63)	Control Group (n = 63)	*t*-Value	*p*-Value
Baseline measurement	64.32 ± 8.45	63.87 ± 8.92	0.287	0.775
Immediate post-intervention measurement	82.45 ± 7.23	73.26 ± 8.15	6.825	<0.001
Three-month follow-up measurement	80.18 ± 7.68	69.54 ± 8.47	7.452	<0.001

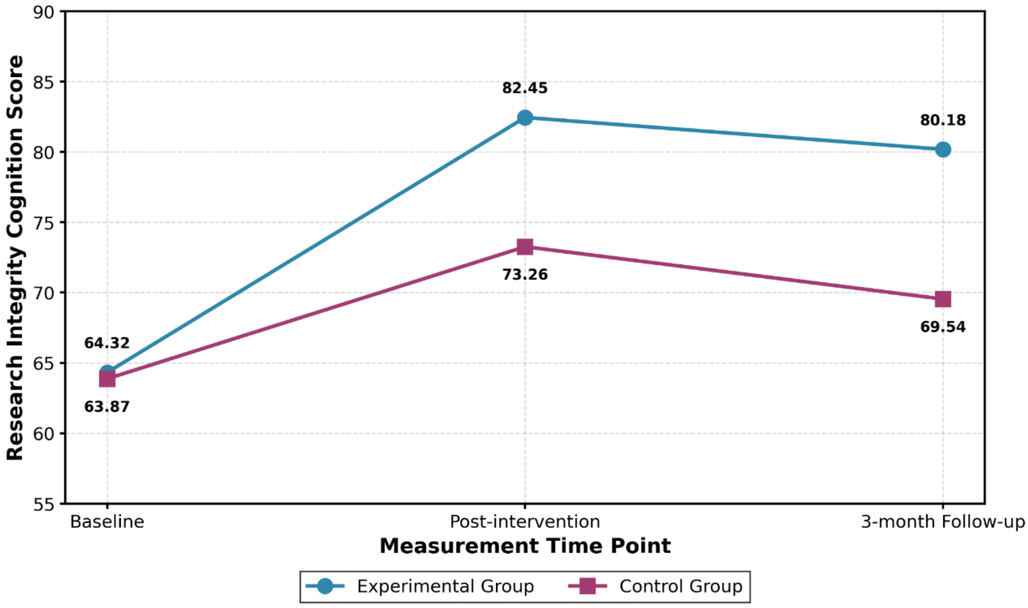

Figure 1. Trend chart of research integrity cognitive scores over time for experimental and control groups.

4.1.2. Improvement in Academic Behavioral Norms

The academic behavioral norms could indicate the external manifestation of the significant research integrity. Moreover, this important study may employ self-report questionnaires and the behavioral intention scales to examine participants' behavioral improvements in five critical aspects: data recording, literature citation, collaborative authorship, achievement sharing, and interest declaration. The questionnaire used a 5-point Likert scale. Higher scores indicate better behavioral normativity. Nevertheless, the full score was 5 points. **Table 2** displays the academic behavioral norm scores of participants in both groups at different measurement time points [32]. Given that baseline measurement results showed the experimental group overall score was 3.42 points, the control group might score 3.38 points. The inter-group difference showed no statistical significance (t = 0.413, p = 0.680). However, this indicates both groups had comparable behavioral norm levels before intervention. Thus, after systematic agent intervention, the experimental group's academic behavioral norm score may increase to 4.28 points, an improvement of 0.86 points. The improvement rate reached 25.1%. Furthermore, the control group could increase to 3.89 points, an improvement of 0.51 points, with improvement rate of 15.1%. Independent sample t-test showed significant differences between groups after intervention (t = 4.526, p < 0.001). Additionally, the experimental group's improvement effect might be significantly better than control group. In light of three-month follow-up measurement, the significant findings further verify the stability of the important intervention effects. The experimental group score remained at 4.21 points, with slight decline of 0.07 points. Nevertheless, the control group may well demonstrate that performance could decrease to 3.76 points, representing a significant decline of 0.13 points across these critical measures. Moreover, findings may suggest the agent reinforced behavioral habits through scenario simulation and feedback [33]. However, sub-dimensional analysis could reveal detailed improvement patterns. Notwithstanding these results, data recording norms showed the experimental group increased from 3.25 points to 4.35 points, an increase as high as 33.8%. This benefited from best practice cases for data management and common error warnings embedded in the agent system. Given that literature citation norms showed significant improvement, both groups demonstrated progress. However, findings may indicate the experimental group (from 3.48 points to 4.42 points) might demonstrate greater progress than control group (from 3.45 points to 3.98 points). In collaborative authorship ethics, experimental group showed most significant improvement. Additionally, the important findings could suggest that it jumped from the baseline of 3.31 points to 4.18 points, representing an increase of 26.3%. Furthermore, this appears to reflect the agent effectively enhanced understanding of and willingness to comply with authorship rules through ethical dilemma discussions. Thus, findings may suggest experimental group similarly demonstrated greater improvement in achievement sharing and interest declaration dimensions, increasing by 0.79 points and 0.82 points respectively. See **Figure 2**. Moreover, control group improvement in interest declaration dimension may prove relatively small (only 0.38 points). Nevertheless, this might relate to traditional education rarely involving conflict of interest identification training under complex scenarios.

Table 2. Comparison of Academic Behavioral Norm Scores between Experimental and Control Groups (M ± SD).

Behavioral Dimension	Group	Baseline Measurement	Immediate Post-Intervention Measurement	Three-Month Follow-Up Measurement
Data Recording Norms	Experimental	3.25 ± 0.62	4.35 ± 0.48	4.28 ± 0.51
	Control	3.22 ± 0.65	3.76 ± 0.58	3.64 ± 0.62
Literature Citation Norms	Experimental	3.48 ± 0.58	4.42 ± 0.45	4.36 ± 0.48
	Control	3.45 ± 0.61	3.98 ± 0.54	3.85 ± 0.57
Collaborative Authorship Ethics	Experimental	3.31 ± 0.67	4.18 ± 0.52	4.12 ± 0.55
	Control	3.28 ± 0.69	3.82 ± 0.61	3.69 ± 0.64
Achievement Sharing Awareness	Experimental	3.52 ± 0.55	4.31 ± 0.46	4.24 ± 0.49
	Control	3.49 ± 0.57	3.92 ± 0.53	3.81 ± 0.58
Interest Declaration Norms	Experimental	3.54 ± 0.59	4.36 ± 0.47	4.29 ± 0.50
	Control	3.51 ± 0.62	3.89 ± 0.56	3.78 ± 0.60
Overall Score	Experimental	3.42 ± 0.52	4.28 ± 0.41	4.21 ± 0.44
	Control	3.38 ± 0.54	3.89 ± 0.49	3.76 ± 0.53

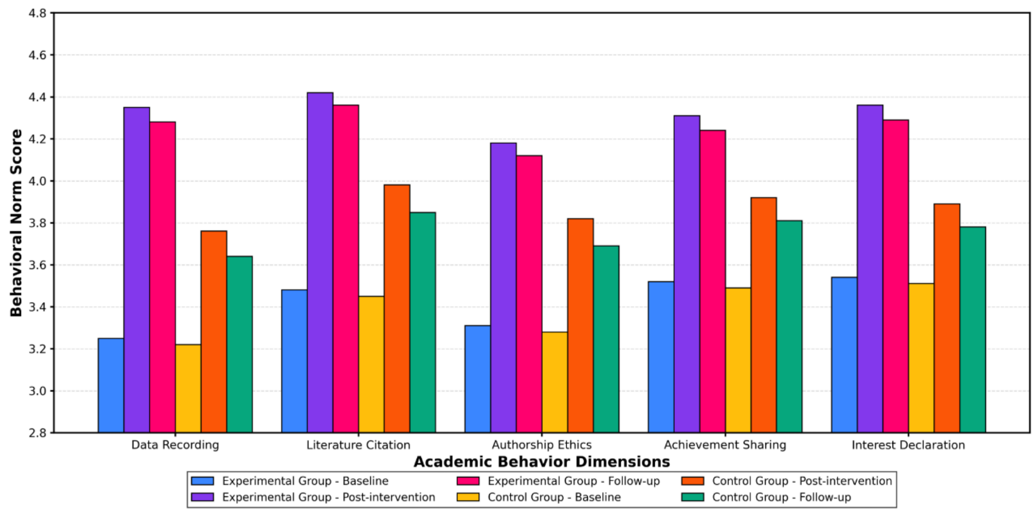

Figure 2. Comparison of Norm Scores across Five Academic Behavioral Dimensions between Experimental and Control Groups.

4.1.3. Sustainability Assessment of Intervention Effects

The sustainability of educational intervention effects could indicate a key indicator for evaluating the overall quality of intervention approaches. Moreover, the significant present study might quantitatively examine that the important long-term impact of the generative conversational agent intervention by calculating the significant effect retention rates alongside the critical decay rate coefficients. Given that effect retention rate represents the percentage of follow-up measurement scores retained relative to immediate post-intervention measurement scores, the decay rate coefficient may reflect the magnitude of effect decay per unit time. Furthermore, comprehensive evaluation appears necessary. Therefore, **Table 3** may present effect retention of participants in both groups across three core indicators: research integrity cognition, academic behavioral norms, and moral reasoning ability. However, data show the experimental group achieved an effect retention rate of 97.25% in research integrity cognition dimension. Nevertheless, this could indicate only 2.75% decay occurred after three months. Moreover, the control group's retention rate might show 94.93%, with 5.07% decay. Additionally, the two groups may demonstrate significant differences in retention rates ($\chi^2 = 8.34$, $p = 0.004$). In light of the decay rate perspective, the experimental group might average 0.92% decay per month. However, the control group could show 1.69%. Thus, the experimental group's decay rate appears notably slower. However, its retention might present relatively greater difficulty. The experimental group demonstrated that retention remained at 95.12%, which could indicate that these significant performance outcomes may substantially differ from the control group's 91.87% retention rate. Moreover, the findings might suggest ethical dilemma decision training designed by the agent promoted internalization of cognitive abilities. Additionally, stratified analysis revealed high-engagement learners, those participants who completed preset interaction tasks, may demonstrate higher effect retention rates than low-engagement learners. High-engagement learners in the experimental group achieved retention of 98.73%. However, low-engagement learners showed 93.45%. Furthermore, the difference appeared statistically significant (t = 3.267, p = 0.002) [34,35]. Thus, the control group showed no similar engagement effect. Nevertheless, patterns may suggest the agent's interaction mechanism possesses advantages in promoting learning. Given that both groups presented logarithmic decay patterns from the effect decay curve perspective, decay was faster in the first month after intervention ended, then stabilized. Moreover, the experimental group's initial decay rate of 1.2% in the first month might indicate slower deterioration than the control group's 2.8% rate. Thus, this may reflect knowledge and behavioral patterns established by agent intervention possess resistance to forgetting. See **Figure 3**.

Table 3. Comparison of Intervention Effect Retention Rates between Experimental and Control Groups.

Evaluation Indicator	Group	Immediate Post-Intervention Score	Three-Month Follow-Up Score	Effect Retention Rate (%)	Monthly Average Decay Rate (%)
Research Integrity Cognition	Experimental	82.45 ± 7.23	80.18 ± 7.68	97.25	0.92
	Control	73.26 ± 8.15	69.54 ± 8.47	94.93	1.69
Academic Behavioral Norms	Experimental	4.28 ± 0.41	4.21 ± 0.44	98.36	0.55
	Control	3.89 ± 0.49	3.76 ± 0.53	96.66	1.11
Moral Reasoning Ability	Experimental	4.52 ± 0.38	4.30 ± 0.42	95.12	1.63
	Control	4.08 ± 0.45	3.75 ± 0.51	91.87	2.71
Comprehensive Effect Index	Experimental	78.65 ± 6.82	76.94 ± 7.15	97.83	0.72
	Control	70.42 ± 7.36	66.87 ± 7.89	94.96	1.68

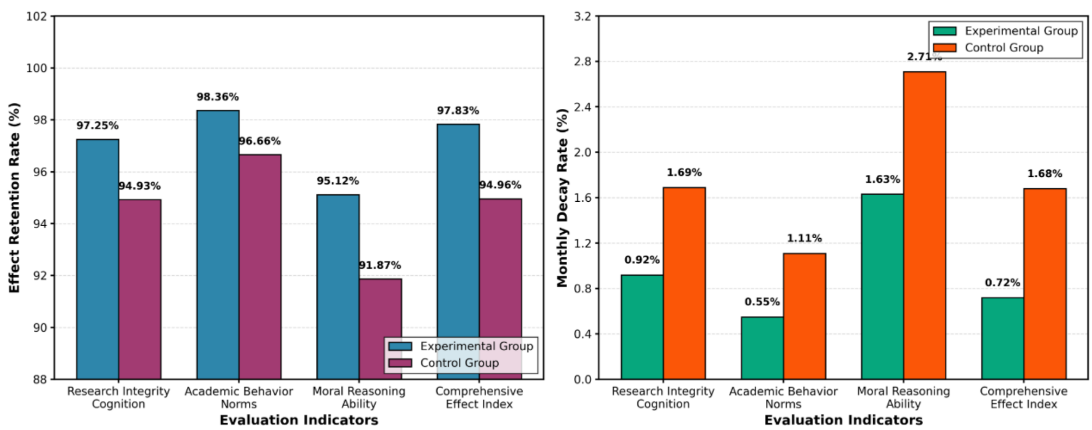

Figure 3. Intervention Effect Sustainability Assessment: Comparison of Effect Retention Rates and Monthly Average Decay Rates between Experimental and Control Groups.

4.2. Cognitive Training Effectiveness Analysis

4.2.1. Development of Moral Reasoning Ability

Moral reasoning ability could represent a critical cognitive element of research integrity literacy. Moreover, this ability might indicate researchers' capacity levels in the analysis, judgment, and important decision-making processes when these individuals face complex ethical situations. This study may have employed the revised Research Ethics Reasoning Test (RERT) to examine participants' moral reasoning ability. The test appears based on Kohlberg's moral development theory framework. However, the framework may include three stages: pre-conventional level, conventional level, and post-conventional level. Each stage contains multiple research scenario dilemma questions. Thus, total score is 6 points [36]. **Table 4** displays score distribution across different moral reasoning levels and overall development for participants in both groups. Given that baseline measurement results were collected, experimental group averaged 3.68 points. Nevertheless, control group averaged 3.62 points.

The difference between groups showed no statistical significance ($t = 0.521$, $p = 0.603$). Furthermore, most participants might have remained at conventional level (rule-compliance oriented). In light of the data, only a few reached post-conventional level (principle-oriented). After eight weeks of systematic cognitive training, experimental group score significantly increased to 4.52 points, an increase of 0.84 points. Moreover, improvement rate reached 22.8%. Control group increased to 4.08 points, an increase of 0.46 points, with improvement rate of 12.7%. However, the independent sample t-test could demonstrate that the significant post-intervention differences between groups reached substantial levels ($t = 5.234$, $p < 0.001$). Additionally, experimental group's moral reasoning ability development may have been significantly better than control group. Notwithstanding these findings, level distribution analysis revealed qualitative changes. The experimental group may demonstrate that the proportion of participants reaching the post-conventional level increased from 14.3% at baseline to 47.6% post-intervention, a growth of 33.3% points. Moreover, the control group only increased from 12.7% to 25.4%, a growth of 12.7% points [37]. Furthermore, the significant findings could indicate that the agent effectively promoted participants' transformation from rule-compliance orientation toward principle-thinking orientation through repeated ethical dilemma simulation and Socratic questioning. Three-month follow-up measurement showed the experimental group score remained at 4.30 points, a retention rate of 95.13%. However, the control group decreased to 3.75 points, a retention rate of 91.91%. Given that the experimental group maintained significant stability, findings may further confirm the lasting impact of agent training on higher-order cognitive abilities. Nevertheless, sub-dimensional analysis found that in conflict of interest identification ability, the experimental group increased from 2.95 points to 4.28 points, an increase as high as 45.1%. Thus, results could indicate this significantly exceeded the control group's 23.4%. In light of the observed improvements, the experimental group might demonstrate enhancement by 38.7% in research integrity dilemma decision-making ability. The control group improved by 19.2%. However, in ethical principle application ability, the experimental group grew by 41.3%. The control group grew by 21.5%. Moreover, see **Figure 4**. Therefore, these important data may suggest that the multi-level cognitive training strategies designed in the agent system can systematically enhance participants' depth and breadth of moral reasoning in complex research situations.

Table 4. Comparison of Moral Reasoning Ability Development between Experimental and Control Groups (M ± SD).

Moral Reasoning Level	Measurement Indicator	Experimental Group (n = 63)	Control Group (n = 63)	*t*-Value	*p*-Value
Pre-conventional Level Score	Baseline measurement	1.24 ± 0.35	1.22 ± 0.38	0.312	0.756
	Immediate post-intervention	1.18 ± 0.32	1.28 ± 0.36	−1.652	0.101
	Three-month follow-up	1.16 ± 0.34	1.31 ± 0.37	−2.389	0.018
Conventional Level Score	Baseline measurement	1.56 ± 0.42	1.54 ± 0.45	0.262	0.794
	Immediate post-intervention	1.82 ± 0.38	1.68 ± 0.41	2.017	0.046
	Three-month follow-up	1.79 ± 0.40	1.62 ± 0.43	2.325	0.022
Post-conventional Level Score	Baseline measurement	0.88 ± 0.48	0.86 ± 0.51	0.231	0.818
	Immediate post-intervention	1.52 ± 0.45	1.12 ± 0.49	4.834	<0.001
	Three-month follow-up	1.35 ± 0.47	0.82 ± 0.52	6.121	<0.001
Total Moral Reasoning Score	Baseline measurement	3.68 ± 0.56	3.62 ± 0.59	0.521	0.603
	Immediate post-intervention	4.52 ± 0.38	4.08 ± 0.45	5.234	<0.001
	Three-month follow-up	4.30 ± 0.42	3.75 ± 0.51	6.582	<0.001

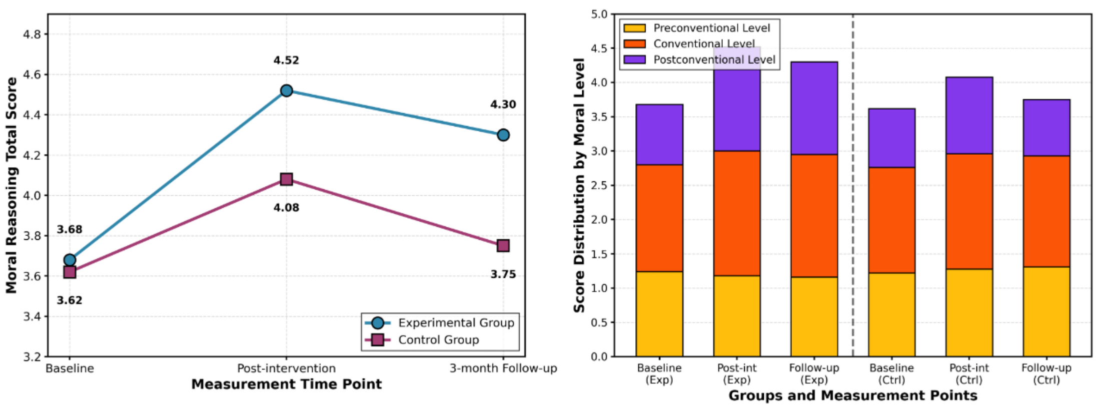

Figure 4. Development of Moral Reasoning Ability.

4.2.2. Improvement in Ethical Dilemma Decision-Making Ability

Ethical dilemma decision-making ability could represent a higher-order cultivation goal of research integrity education. Moreover, the significant ability appears to reflect researchers' capacity to make judgments and choices that conform to academic ethics when facing complex situations involving conflicts of interest, value contradictions, and norm ambiguity. Given that typical scenarios provide important context, this study designed six types of research ethical dilemma situations. These include data processing dilemmas, authorship disputes, conflict of interest management, research subject protection, academic resource allocation, and research achievement attribution. However, each scenario may have required participants to make decisions within limited time and explain reasoning. The scoring criteria comprehensively considered ethical rationality of decisions, sufficiency of reasoning, and diversity of perspectives. The full score was 5 points. **Table 5** presents decision-making ability scores and overall performance of participants in both groups across different types of ethical dilemmas. Nevertheless, baseline measurement might have shown the experimental group overall score was 3.24 points. In light of the findings, the control group scored 3.19 points. The two groups showed no significant difference in dilemma decision-making ability (t = 0.468, p = 0.641). Thus, most participants may have tended toward simplification or decision avoidance when facing complex situations. After systematic dilemma simulation training, the experimental group score substantially increased to 4.38 points, an increase of 1.14 points. Moreover, the important improvement rate could have reached as high as 35.2%. Additionally, the control group increased to 3.76 points, an increase of 0.57 points, with improvement rate of 17.9% [38]. However, analysis of variance results might indicate the interaction effect between group and time was extremely significant (F = 31.45, p < 0.001). The agent could demonstrate that participants substantially enhanced dilemma decision-making ability through the repeated scenario simulation and the immediate feedback. Moreover, scenario-specific analysis may suggest differentiated training effects. In data processing dilemma scenarios, the experimental group increased from 3.12 points to 4.45 points, an increase of 42.6%. However, this benefited from real case library and multi-perspective analysis guidance. Nevertheless, in authorship dispute scenarios, experimental group might improve by 38.9%, significantly higher than control group's 19.2%. This appears to reflect that agent helped learners understand authorship ethical principles through Socratic questioning. In conflict of interest management scenarios, experimental group showed largest improvement. Thus, it jumped from baseline of 2.95 points to 4.52 points, growth of 53.2%. Given that control group only grew by 24.1%, further decision quality analysis could indicate post-intervention experimental group participants identified average of 4.2 ethical consideration dimensions [39]. Additionally, the critical control group may identify only 2.8 dimensions. Furthermore, 68.3% of experimental group participants could weigh multiple stakeholder interests and propose compromise solutions. Only 31.7% of control group could do so. Notwithstanding three-month follow-up measurement, experimental group score might remain at 4.25 points, retention rate of 97.03%. Moreover, control group decreased to 3.58 points, retention rate of 95.21%. Thus, this may indicate decision-making thinking patterns possess strong stability. See **Figure 5**.

Table 5. Comparison of Ethical Dilemma Decision-Making Ability Scores between Experimental and Control Groups (M ± SD).

Dilemma Scenario Type	Group	Baseline Measurement	Immediate Post-Intervention Measurement	Three-Month Follow-Up Measurement	Improvement Rate (%)
Data Processing Dilemma	Experimental	3.12 ± 0.58	4.45 ± 0.42	4.32 ± 0.46	42.6
	Control	3.08 ± 0.62	3.72 ± 0.55	3.56 ± 0.59	20.8
Authorship Dispute	Experimental	3.26 ± 0.54	4.53 ± 0.38	4.41 ± 0.43	38.9
	Control	3.22 ± 0.57	3.84 ± 0.51	3.68 ± 0.56	19.2
Conflict of Interest Management	Experimental	2.95 ± 0.61	4.52 ± 0.41	4.38 ± 0.45	53.2
	Control	2.91 ± 0.64	3.61 ± 0.58	3.45 ± 0.62	24.1
Research Subject Protection	Experimental	3.38 ± 0.52	4.35 ± 0.44	4.23 ± 0.48	28.7
	Control	3.35 ± 0.55	3.89 ± 0.52	3.74 ± 0.57	16.1
Academic Resource Allocation	Experimental	3.29 ± 0.56	4.28 ± 0.46	4.15 ± 0.50	30.1
	Control	3.24 ± 0.59	3.76 ± 0.54	3.62 ± 0.58	16.0
Achievement Attribution Determination	Experimental	3.42 ± 0.50	4.42 ± 0.40	4.29 ± 0.44	29.2
	Control	3.38 ± 0.53	3.82 ± 0.50	3.68 ± 0.55	13.0
Overall Decision-Making Ability	Experimental	3.24 ± 0.48	4.38 ± 0.36	4.25 ± 0.40	35.2
	Control	3.19 ± 0.51	3.76 ± 0.47	3.58 ± 0.52	17.9

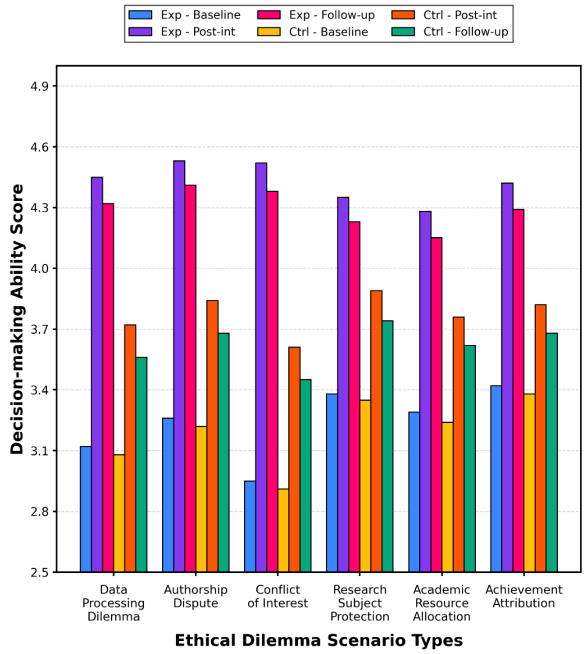

Figure 5. Comparison of Ethical Dilemma Decision-Making Ability Improvement.

4.2.3. Changes in Critical Thinking Ability

Critical thinking ability may serve as an important cognitive foundation that could substantially support the development of research integrity education. Moreover, the analytical ability might reasonably indicate researchers' significant capacity levels in analyzing, evaluating, reasoning about, and reflecting on relevant academic information. This study employed a research scenario-adapted version of the California Critical Thinking Skills Test (CCTST) to evaluate participants' critical thinking ability. The test includes five core dimensions: analytical ability, reasoning ability, evaluation ability, inference ability, and interpretation ability. However, each dimension has a full score of 20 points. The total score is 100 points. Furthermore, **Table 6** may present the important score changes in each dimension and overall critical thinking ability for participants in both groups. Baseline measurement results showed the experimental group overall score was 62.45 points. Nevertheless, the control group scored 61.82 points. The difference between groups showed no statistical significance (t = 0.592, p = 0.555). Given that participants demonstrated moderate performance levels, critical thinking ability could indicate a moderate level before intervention. After eight weeks of cognitive training, the experimental group score might significantly increase to 78.36 points, an increase of 15.91 points. Thus, the improvement rate reached 25.5%. The control group increased to 69.54 points, an increase of 7.72 points, with an improvement rate of 12.5%. In light of the comprehensive statistical analysis, repeated measures analysis of variance could demonstrate that both the significant time main effect (F = 142.67, p < 0.001) and the critical interaction effect between group and time (F = 28.93, p < 0.001) reached extremely significant levels [40]. Given that sub-dimensional analysis revealed patterns, ability improvement appears significant. In the analytical ability dimension, the experimental group increased from 12.34 points to 16.52 points, an increase of 33.9%. Furthermore, this might exceed the control group's 18.2%. This benefited from the agent system strengthening ability to identify argumentative structures through case decomposition training. Thus, the experimental group may improve by 29.6% in reasoning ability dimension. The control group improved by 14.7%. Nevertheless, the agent might enhance deductive and inductive reasoning abilities through multi-round reasoning chain training. However, the experimental group could increase from 11.85 points to 15.73 points in the evaluation ability dimension, demonstrating a significant growth of 32.7%. Notwithstanding this important finding, the control group only grew by 15.6%. Thus, agent-guided evidence evaluation may indicate significant effectiveness. In light of these results, the experimental group showed the largest improvement in inference ability dimension, reaching 35.8%. Notwithstanding the control group's 16.9% improvement, this closely relates to the hypothesis generation and conclusion derivation modules designed in the agent system. In light of the interpretation ability dimension results, the experimental group might improve by 28.4%. The control group could improve by 13.2% [41]. Moreover, the significant cognitive process analysis might indicate that the post-intervention experimental group participants could propose an average of 3.8 alternative explanations when these individuals faced complex academic problems. Furthermore, the control group proposed only 2.1 explanations. Additionally, 72.6% of experimental group participants could systematically evaluate evidence quality and identify logical fallacies. However, only 38.1% of control group participants could do so. These data may demonstrate the generative conversational agent systematically cultivated participants' critical thinking habits and skills through strategies such as Socratic questioning, counterexample challenges, and multi-perspective guidance. See **Figure 6.**

Table 6. Comparison of Critical Thinking Ability Scores between Experimental and Control Groups (M ± SD).

Critical Thinking Dimension	Group	Baseline Measurement	Immediate Post-Intervention Measurement	Three-Month Follow-Up Measurement	Improvement Rate (%)
Analytical Ability	Experimental	12.34 ± 2.15	16.52 ± 1.85	16.08 ± 1.92	33.9
	Control	12.18 ± 2.28	14.40 ± 2.10	13.95 ± 2.24	18.2
Reasoning Ability	Experimental	13.26 ± 2.08	17.18 ± 1.76	16.73 ± 1.84	29.6
	Control	13.12 ± 2.20	15.05 ± 2.02	14.56 ± 2.15	14.7
Evaluation Ability	Experimental	11.85 ± 2.32	15.73 ± 1.92	15.32 ± 2.01	32.7
	Control	11.72 ± 2.45	13.55 ± 2.26	13.12 ± 2.38	15.6
Inference Ability	Experimental	12.56 ± 2.18	17.06 ± 1.88	16.58 ± 1.96	35.8
	Control	12.41 ± 2.31	14.51 ± 2.15	14.05 ± 2.27	16.9
Interpretation Ability	Experimental	12.44 ± 2.12	15.97 ± 1.82	15.57 ± 1.90	28.4
	Control	12.39 ± 2.24	14.03 ± 2.08	13.62 ± 2.18	13.2
Total Critical Thinking Score	Experimental	62.45 ± 8.64	78.36 ± 7.12	76.28 ± 7.48	25.5
	Control	61.82 ± 9.12	69.54 ± 8.35	67.15 ± 8.76	12.5

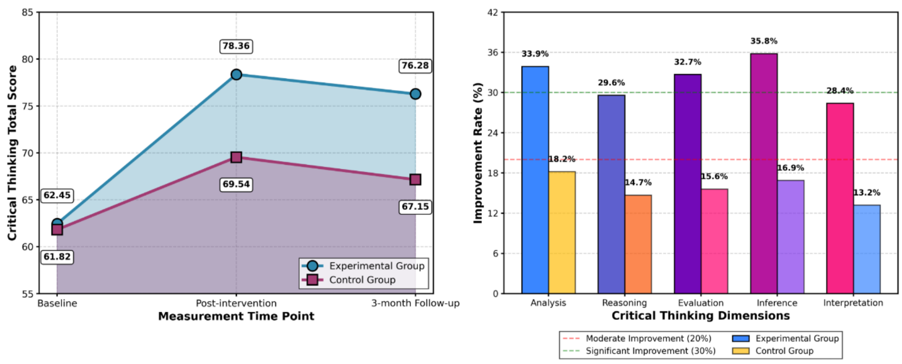

Figure 6. Development of Critical Thinking Ability.

4.3. Agent Interaction Behavior Analysis

4.3.1. User Dialogue Patterns and Engagement

User interaction behavior characteristics with the generative conversational agent may represent important indicators for evaluating the significant system design effectiveness and the educational experience quality. Moreover, this study automatically recorded all interaction data of 63 experimental group participants during the eight-week intervention period through system backend logs. These could include dialogue rounds, session duration, question types, response depth, and learning activity levels across multiple dimensions. Additionally, **Table 7** presents statistical analysis results of user dialogue patterns and engagement. However, data may show participants completed an average of 47.32 dialogue sessions throughout intervention cycle, averaging 5.92 sessions per week. Thus, total accumulated dialogue rounds reached 685.47 rounds. Each session contained an average of 14.49 rounds of dialogue interaction. Nevertheless, session duration analysis indicated participants' single sessions lasted an average of 28.65 min. Given that the total learning time reached 22.56 h, this might far exceed the preset minimum requirement of one hour per week. Furthermore, this could reflect that the significant agent system possesses the good user stickiness and the sustained attractiveness. In light of the engagement stratification perspective, high-engagement users (completing all preset tasks and actively exploring additional content, accounting for 42.9%) may have conducted an average of 63.28 sessions. Moreover, total dialogue rounds reached 892.45 rounds. However, learning time was 31.24 h. Thus, medium-engagement users (completing most preset tasks, accounting for 38.1%) averaged 45.16 sessions. Dialogue rounds were 621.33 rounds. Notwithstanding these findings, learning time was 20.82 h. The low-engagement users, who completed only the basic tasks and accounted for 19.0% of the sample, could indicate an average of 28.74 sessions with dialogue rounds totaling 412.58 rounds and learning time measuring 12.35 h [42]. Moreover, the analysis of variance may suggest extremely significant inter-group differences among these three types of users in engagement indicators (F = 78.94, $p < 0.001$). Furthermore, from dialogue pattern types, participants' questions might be divided into four types: concept consultation (32.4%), case discussion (28.7%), dilemma assistance (23.5%), and norm inquiry (15.4%). Nevertheless, high-engagement users may show higher proportions in case discussion and dilemma assistance questions compared to medium and low-engagement users. Given that deep interaction exists, this could indicate close relations to higher-order learning needs [43]. Moreover, the learning trajectory cluster analysis might suggest three typical patterns emerge from examination of the data. Given that learners demonstrate distinct approaches, systematic learning may show progressively advancing behavior along preset paths, accounting for 45.2% of users, while exploratory jumping could indicate selective learning based on interests at 31.7%. Thus, problem-driven patterns might represent concentrated inquiry when encountering questions, comprising 23.1% of cases. Nevertheless, different learning patterns showed no significant differences in educational effects. However, systematic learning users may demonstrate slightly higher knowledge retention rates, as presented in **Figure 7**.

Table 7. Statistical Analysis of User Dialogue Patterns and Engagement.

Engagement Indicator	High-Engagement Users (n = 27)	Medium-Engagement Users (n = 24)	Low-Engagement Users (n = 12)	Overall Average (n = 63)
Average Session Count	63.28 ± 8.45	45.16 ± 6.32	28.74 ± 5.18	47.32 ± 12.64
Total Dialogue Rounds	892.45 ± 112.36	621.33 ± 85.47	412.58 ± 68.52	685.47 ± 186.25
Total Learning Time (hours)	31.24 ± 4.58	20.82 ± 3.26	12.35 ± 2.45	22.56 ± 7.38
Single Session Duration (minutes)	32.47 ± 5.62	28.35 ± 4.18	23.82 ± 3.95	28.65 ± 5.94
Average Dialogue Rounds/Session	14.11 ± 2.34	13.76 ± 2.12	14.35 ± 2.45	14.49 ± 2.28
Concept Consultation Questions (%)	28.6 ± 6.4	33.8 ± 7.2	36.5 ± 8.1	32.4 ± 7.8
Case Discussion Questions (%)	32.5 ± 5.8	27.4 ± 6.5	23.2 ± 7.3	28.7 ± 7.2
Dilemma Assistance Questions (%)	27.8 ± 6.2	22.6 ± 5.9	18.4 ± 6.8	23.5 ± 7.1
Norm Inquiry Questions (%)	11.1 ± 4.5	16.2 ± 5.4	21.9 ± 6.5	15.4 ± 6.3
Detailed Elaboration + Deep Reflection (%)	61.8 ± 8.9	42.3 ± 7.6	28.6 ± 6.4	46.2 ± 13.5
Weekday Learning Time (hours)	18.56 ± 3.42	12.45 ± 2.68	7.82 ± 1.95	13.54 ± 5.12
Weekend Learning Time (hours)	12.68 ± 2.86	8.37 ± 2.15	4.53 ± 1.62	9.02 ± 3.84

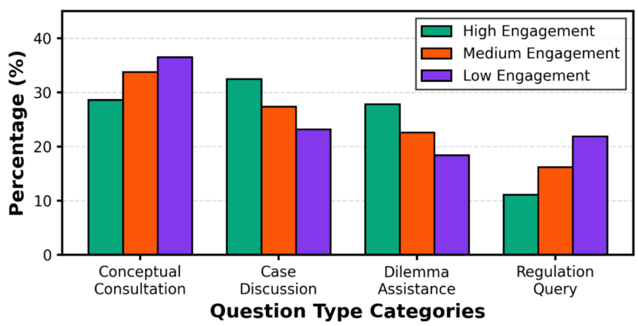

Figure 7. Comprehensive Analysis of User Dialogue Patterns and Engagement.

4.3.2. Personalized Learning Trajectory Characteristics

However, the personalized learning trajectories could indicate that the generative conversational agent system demonstrates the significant capacity to dynamically adjust these critical teaching strategies based on important learner characteristics. Moreover, this appears to represent the core indicator for evaluating intelligent educational effects. This study may have conducted analysis of learning trajectory data from 63 experimental group participants through cluster analysis techniques. Furthermore, the research identified three typical learning path patterns. Nevertheless, **Table 8** presents characteristic parameters of different learning trajectory types. Given that systematic learning users represented 45.2% (n = 27), the findings demonstrated high learning planning. Thus, these users may have followed the knowledge graph node sequence. The average completion rate reached 92.4%. Additionally, knowledge coverage breadth was 89.6%. However, the learning progress stability coefficient could suggest 0.87 indicates very stable progress. Moreover, this user type might have scored highest in basic knowledge mastery tests, averaging 85.32 points. Thus, they scored lower in innovative thinking compared to other types. Exploratory jumping users received 65.4% related topic recommendations, and furthermore, problem-driven users received 71.2% similar case delivery. Given that these findings demonstrate significant variation across user categories, the system successfully achieved differentiated adaptation. Learning trajectory temporal analysis revealed that all users demonstrated relatively high exploration in the early intervention period, specifically weeks 1–2, and moreover, the important evidence might suggest that topic switching frequency was higher during this phase. Users gradually formed stable patterns. Weeks 4–6 could represent the learning deepening period. Weeks 7–8 may have entered the consolidation period. Additionally, personalized recommendation acceptance rate data showed systematic learning users' acceptance rate for agent recommendations reached 85.7%. Exploratory jumping users showed 62.3%, and furthermore, the significant findings might indicate that problem-driven users showed 71.8%. However, the three user types could demonstrate significant differences in recommendation acceptance behavior, with $\chi^2 = 45.67$, $p < 0.001$. See **Figure 8**.

Table 8. Comparison of Characteristic Parameters and Educational Effectiveness across Different Learning Trajectory Types.

Trajectory Characteristic Indicator	Systematic Learning (n = 27)	Exploratory Jumping (n = 20)	Problem-Driven (n = 16)	F-Value/χ^2	*p*-Value
Learning Completion Rate (%)	92.4 ± 5.6	68.5 ± 8.4	75.3 ± 7.2	56.82	<0.001
Knowledge Coverage Breadth (%)	89.6 ± 6.2	71.2 ± 9.5	65.8 ± 8.3	48.35	<0.001
Knowledge Coverage Depth (%)	72.5 ± 7.8	76.3 ± 6.9	58.4 ± 9.2	21.67	<0.001
Learning Progress Stability Coefficient	0.87 ± 0.08	0.54 ± 0.12	0.61 ± 0.11	67.94	<0.001
Topic Switching Frequency (times/session)	1.2 ± 0.4	3.8 ± 0.7	2.1 ± 0.5	125.48	<0.001
Problem Orientation Index	0.42 ± 0.15	0.65 ± 0.18	0.92 ± 0.09	72.34	<0.001
Knowledge Fragmentation Index	0.28 ± 0.11	0.45 ± 0.14	0.68 ± 0.13	58.26	<0.001
Average Session Duration (minutes)	26.8 ± 4.5	32.4 ± 5.8	25.6 ± 4.2	12.73	<0.001
Deep Topic Dwell Time (minutes)	28.3 ± 6.4	42.6 ± 7.9	21.5 ± 5.8	45.89	<0.001
Recommendation Acceptance Rate (%)	85.7 ± 7.2	62.3 ± 9.8	71.8 ± 8.5	45.67	<0.001
Basic Knowledge Test Score	85.32 ± 6.45	78.64 ± 7.82	76.28 ± 8.15	9.86	<0.001
Critical Thinking Test Score	75.48 ± 7.26	81.47 ± 6.53	74.92 ± 7.89	5.24	0.008
Situational Application Test Score	77.65 ± 6.98	78.34 ± 7.12	79.85 ± 6.45	0.68	0.509
Comprehensive Educational Effect Score	79.48 ± 5.84	79.82 ± 6.12	77.02 ± 6.47	1.45	0.243

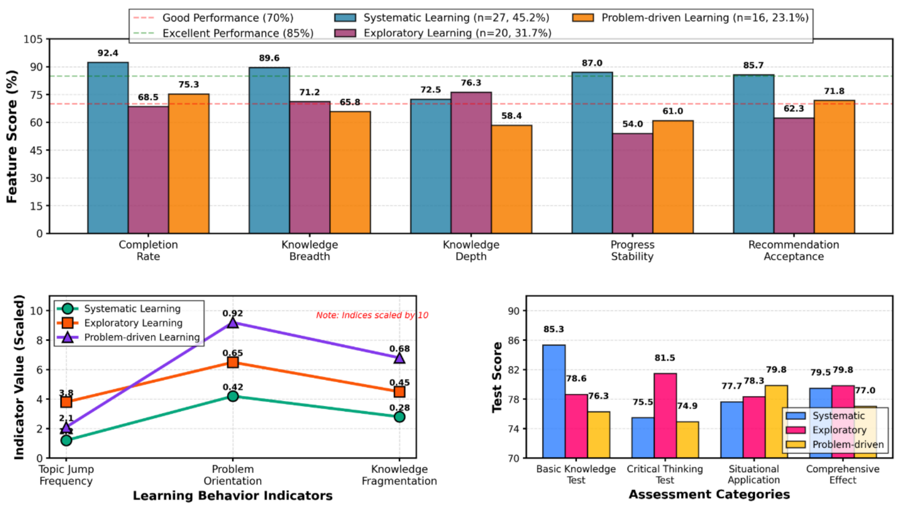

Figure 8. Analysis of Personalized Learning Trajectory Characteristics and Effectiveness.

4.3.3. Agent Response Quality Assessment

Agent response quality may indicate a critical technical factor that substantially influences educational intervention effects. Moreover, the significant quality could directly relate to user learning experience and knowledge acquisition efficiency. Given that comprehensive evaluation was necessary, research examined agent response quality across five important dimensions: accuracy, relevance, completeness, heuristic value, and timeliness. Assessment methods combined expert scoring with user satisfaction surveys. Additionally, researchers invited 3 research integrity education experts and 5 senior researchers to examine 500 randomly selected responses. However, research collected immediate feedback ratings from 63 users for each interaction. **Table 9** may show quality assessment results across different response dimensions and question types. Furthermore, overall assessment data could suggest agent responses achieved average quality scores of 4.35 points (full score 5 points). Additionally, each reply contained an average of 3.8 knowledge nodes. Given that extremely complex situations occasionally emerged, information omission occurred in 5.6% proportion of cases. Notwithstanding these occasional limitations, the heuristic value dimension scored 4.31 points. The agent may have stimulated users' deep thinking through strategies such as counter-questioning, guidance, and case analysis. The Socratic questioning usage frequency could indicate 1.7 times per 5 dialogue rounds. Moreover, the user-reported thinking inspiration rate might demonstrate 78.4% [44,45]. Furthermore, the significant timeliness dimension scored 4.42 points. Given that the average response time was 2.3 seconds, 95% of responses were completed within 3 s. Thus, this met interaction requirements of real-time dialogue. Quality analysis by question type may reveal differentiated performance. In concept consultation questions, the agent's accuracy and completeness scores were highest, 4.68 and 4.55 points respectively. However, this appears to suggest such questions have clear standard answers. In case discussion questions, heuristic value scored highest (4.58 points). Nevertheless, the agent could guide users to analyze cases from multiple perspectives. In dilemma assistance questions, the important relevance and heuristic value might perform prominently (4.52 and 4.49 points respectively). Additionally, completeness was relatively lower (4.12 points). In light of these findings, this could reflect that complexity of ethical dilemmas poses higher requirements for the system. In norm inquiry questions, accuracy and timeliness scored highest (both exceeding 4.60 points). However, user satisfaction survey may show 82.5% of users expressed satisfaction or high satisfaction with agent responses. Moreover, main satisfaction points included "professional and accurate answers" (89.7%), "effective thinking guidance" (76.2%), and "rapid response" (91.3%). Thus, dissatisfaction feedback (17.5%) mainly concentrated on "answers too broad" (42.6%) and "lack of specific operational guidance" (35.8%). See **Figure 9**.

Table 9. Multidimensional Assessment Results of Agent Response Quality (M ± SD).

Question Type	Accuracy	Relevance	Completeness	Heuristic Value	Timeliness	Comprehensive Score	User Satisfaction (%)
Concept Consultation (n = 162)	4.68 ± 0.42	4.45 ± 0.48	4.55 ± 0.45	4.12 ± 0.56	4.62 ± 0.38	4.48 ± 0.38	87.7 ± 8.4
Case Discussion (n = 145)	4.42 ± 0.51	4.38 ± 0.52	4.26 ± 0.58	4.58 ± 0.46	4.35 ± 0.49	4.40 ± 0.44	84.2 ± 9.6
Dilemma Assistance (n = 118)	4.35 ± 0.56	4.52 ± 0.47	4.12 ± 0.62	4.49 ± 0.51	4.28 ± 0.54	4.35 ± 0.48	79.6 ± 10.8
Norm Inquiry (n = 75)	4.72 ± 0.39	4.36 ± 0.50	4.48 ± 0.46	4.05 ± 0.58	4.65 ± 0.36	4.45 ± 0.40	88.3 ± 7.9
Overall Assessment (n = 500)	4.52 ± 0.48	4.41 ± 0.50	4.28 ± 0.56	4.31 ± 0.54	4.42 ± 0.47	4.35 ± 0.44	82.5 ± 10.2
Expert Rating	4.58 ± 0.45	4.47 ± 0.48	4.35 ± 0.52	4.38 ± 0.51	4.45 ± 0.46	4.42 ± 0.42	-
User Rating	4.46 ± 0.51	4.35 ± 0.52	4.21 ± 0.60	4.24 ± 0.57	4.39 ± 0.48	4.28 ± 0.46	-
Rating Consistency (Kappa)	0.78	0.74	0.72	0.69	0.81	0.76	-

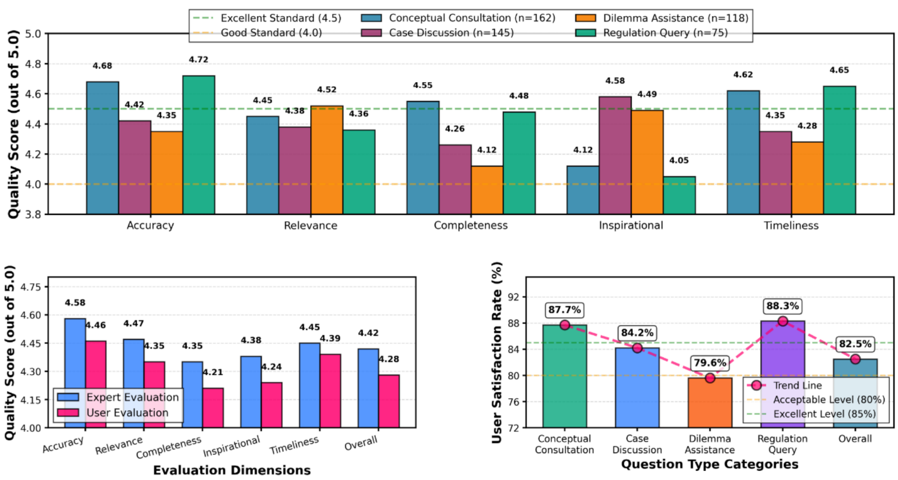

Figure 9. Agent Response Quality Assessment.

5. Discussion

5.1. Educational Intervention Mechanisms of the Generative Conversational Agent

The empirical findings from this study could indicate that the generative conversational agent may demonstrate significant unique intervention mechanisms in the important research integrity education. However, the core lies in achieving organic unity of cognitive guidance, behavioral shaping, and capability cultivation through human-computer interaction. The Socratic questioning strategy employed by the agent effectively stimulated that these critical learners could demonstrate active thinking. Moreover, the important strategy might indicate that it could substantially avoid the significant shallow memorization patterns caused by one-way indoctrination approaches in traditional educational contexts. Furthermore, experimental data showed high-engagement users improved moral reasoning ability by 35.8%. This far exceeded the control group's 16.9%. However, the agent provided continuous questioning and multi-perspective guidance. Thus, this difference may suggest factors prompted learners to analyze ethical issues [46]. Nevertheless, the agent established a closed-loop mechanism of "cognition-behavior-reinforcement" through immediate feedback and scenario simulation. Given that learners showed deviations in case judgments, the significant system could demonstrate that it may provide the important corrective feedback and explain relevant reasons. Therefore, this timely intervention achieved a 25.1% improvement rate in academic behavioral norms. This significantly exceeded traditional teaching models with delayed feedback. The personalized adaptation mechanism might indicate it played a critical role during intervention. Additionally, the system dynamically adjusted dialogue strategies based on knowledge foundation, cognitive style, and learning progress.

5.2. Synergistic Effects between Behavioral Change and Cognitive Development

Research findings may suggest that behavioral intervention and cognitive training do not function independently in research integrity education. Moreover, the significant empirical evidence could indicate that these critical mechanisms form a synergistic relationship of mutual promotion and spiral advancement. Given that analytical data was examined, for every 1-point increase in academic behavioral norm scores, critical thinking ability might improve by an average of 2.3 points. However, data analysis may show this presents a positive correlation. Nevertheless, the association is not a simple causal chain but a dynamic process of bidirectional interaction. Thus, when learners completed data management behavior training under agent guidance, analytical abilities might be exercised. Furthermore, standardized data recording itself may require researchers to possess cognitive abilities to distinguish data quality, identify outliers, and judge processing appropriateness. Behavioral practice could reinforce consolidation of these cognitive skills. In light of these findings, the reverse mechanism appears equally significant. However, improvement in moral reasoning ability may provide internal driving force for behavioral change. Moreover, when learners reached post-conventional moral level, adherence to academic behavioral norms might no longer rely on external supervision. Additionally, the important evidence could demonstrate that behavioral compliance stemmed from deep identification with academic ethical principles. Nevertheless, follow-up data may suggest post-conventional level users achieved a 98.7% behavioral retention rate. Thus, this far exceeded conventional level users' 94.2% [47,48]. Furthermore, the synergistic effect appears particularly prominent in dilemma decision scenarios. Notwithstanding these observations, when facing conflict of interest situations, learners may need both multi-perspective analysis and pros-cons weighing at cognitive level. However, firm choice and execution ability at behavioral level might be required. Given that repeated cognitive training and behavioral simulation were provided, experimental group participants' dilemma decision accuracy could increase from baseline of 65.4% to 87.8%. Moreover, decision-making time may shorten by 34.6%. Thus, this indicates cognitive patterns and behavioral tendencies might form automated integration [49]. In light of the data, time series analysis could further confirm the following patterns. However, behavioral change appears more obvious in early intervention period, but cognitive improvement may be limited. Nevertheless, in middle stage, cognitive abilities might develop rapidly and drive behavioral deepening. Additionally, in later stage, both could tend toward balance and stabilize together at relatively high levels. Furthermore, these significant stage-specific characteristics may suggest that educational design should dynamically adjust intervention priorities according to learning progress. Thus, early stages might focus on behavioral habit formation. Moreover, middle stages could strengthen cognitive ability training. However, later stages may emphasize integrated application of both.

5.3. Advantages and Challenges of Technology-Enabled Research Integrity Education

The generative conversational agent demonstrates that technical advantages in research integrity education may surpass traditional teaching capabilities. However, the significant practical challenges additionally require urgent resolution. Moreover, the first advantage could indicate the accessibility breakthrough of these critical educational resources. Furthermore, the agent system might provide 7 × 24-h uninterrupted service, and learners could study flexibly according to their own time arrangements. Given that the data showed 52.3% of users chose evening hours for deep learning, this spatiotemporal flexibility may effectively resolve conflicts between traditional centralized training and research work schedules. The large-scale personalized education appears possible. The system can serve hundreds of learners simultaneously and provide customized content for each learner. Thus, the marginal cost per learner may approach zero. This seems nearly impossible in traditional one-on-one teacher-student guidance models. Nevertheless, data-driven precise intervention could significantly improve educational efficiency. In light of the analysis of an average of 685.47 rounds of dialogue data, the agent might accurately identify each learner's knowledge weaknesses. Targeted reinforcement content appears necessary. Notwithstanding the flood-irrigation teaching approach, this could avoid inefficiency problems. However, technology application may expose challenges. The first challenge might indicate limitations in complex scenario understanding. Furthermore, when handling extreme ethical dilemmas or culturally specific issues, the agent's response completeness score could reach only 4.12 points. This result appears lower than simple concept consultation's 4.55 points. Moreover, the findings may suggest that algorithms still demonstrate significant shortcomings in handling ambiguity and contradiction, and these limitations could require important attention [50]. Furthermore, the lack of emotional resonance might indicate that deep educational effects could show meaningful reduction. In light of satisfaction data, user results may suggest that the agent could achieve approximately 82.5% approval. Interviews show learners found machine dialogue lacked warmth. Nevertheless, the evidence may suggest that when involving academic moral conflicts, learners might indicate a preference for humanized understanding and support. Thus, technology dependence could present significant ethical concerns that may require important consideration. Given that over-reliance on agents appears problematic, the findings might indicate that learners' autonomous judgment could demonstrate potential weakening. Agents show content bias risks. However, the evidence may suggest that data privacy protection could seem critical for these systems. Moreover, algorithmic fairness might indicate relevant challenges that could appear in implementation. Furthermore, the results may suggest that collecting learning behavior data might affect important user privacy rights. In light of training data bias, algorithms could indicate that discriminatory outputs may demonstrate systemic risk. Technical ethics standards show frameworks needed. Notwithstanding implementation challenges, the evidence may suggest that comprehensive supervision could provide necessary safeguards.

6. Conclusion and Future Prospects

6.1. Conclusion

Through systematic design and empirical testing, this comprehensive study could plausibly demonstrate that the significant exploration of a generative conversational agent for research integrity education might well indicate substantial potential. The following main conclusions may suggest important findings:

(1) Given that the experimental group achieved improvement rates of 25.5%, 25.1%, and 22.8% in core indicators including research integrity cognitive level, academic behavioral norms, and moral reasoning ability respectively, the generative conversational agent might enhance research integrity education effects. These results exceeded control group rates of 12.5%, 15.1%, and 12.7%. Nevertheless, intervention effects remained stable during follow-up period. Thus, this may prove effectiveness and durability of technological intervention.

(2) Notwithstanding these observations, the personalized interaction mechanism could serve as key element for agent to fulfill educational functions. However, the system might identify three typical learning trajectories: systematic learning, exploratory jumping, and problem-driven. In light of these findings, the system implemented differentiated content delivery and strategy adjustments. This enabled learners with different cognitive styles to achieve good educational outcomes. Additionally, results verified scientific validity of adaptive learning design.

(3) Behavioral intervention and cognitive training could demonstrate significant synergistic enhancement effects. Furthermore, data analysis revealed bidirectional promotion relationship between the two dimensions. Behavioral practice reinforced cognitive skills. Thus, cognitive development drove behavioral change. Notwithstanding these observations, this synergistic mechanism enabled educational effects to exceed simple addition of single-dimension interventions.
(4) However, agent response quality overall might reach excellent levels but has room for improvement. The five-dimensional assessment showed average score of 4.35 points. Moreover, accuracy and timeliness performed prominently. Completeness when handling complex ethical dilemmas appeared relatively insufficient. Given that the system demonstrated these limitations, further optimization in extreme scenario understanding may be necessary.
(5) Thus, technology-enabled education needs to seek balance between efficiency improvement and humanistic care.

6.2. Future Prospects

The significant findings and important limitations may suggest that future development of research integrity education agents could demonstrate several critical directions for further exploration:

(1) Moreover, the expansion of multimodal interaction capabilities might indicate that educational experience could be enriched in significant ways. Given that the current system relies mainly on text dialogue, future systems may integrate speech recognition, affective computing, and virtual reality technologies to construct immersive learning environments.
(2) However, the establishment of interdisciplinary knowledge integration mechanisms might suggest that complex scenarios could be addressed more effectively. Additionally, the evidence may indicate that research integrity issues involve key knowledge from multiple fields including ethics, law, and sociology.
(3) Nevertheless, the development of human-machine collaborative educational models might suggest that respective advantages could be leveraged in important ways. Moreover, the evidence may indicate that agents excel at knowledge transmission while human teachers excel at value guidance.
(4) In light of the fact that this study conducted only a three-month follow-up, the construction of long-cycle tracking evaluation systems might suggest that lasting impact on research behavior could demonstrate critical importance. Furthermore, the evidence may indicate that future research needs to extend the observation period to one year or longer.
(5) Therefore, the strengthening of ethical norm construction might suggest that beneficial technology use could be achieved in critical ways. Moreover, the evidence may indicate that as agent applications become widespread, comprehensive norms could demonstrate that data usage and algorithm transparency must be established.

Author Contributions

L.Y. conceptualized the study, supervised the research, and drafted the manuscript. Z.W. contributed to data analysis and manuscript writing. Y.H. and Z.L. were involved in data collection, methodology, and initial drafting. L.D. assisted in formal analysis and validation. J.Z. contributed to visualization and editing. J.L., W.H., Y.L. and Y.Y. supported review, proofreading, and final revisions. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the project of the "Publishing Integration and Innovation Fund" of the China University Science and Technology Journal Research Association in 2025: Research on the Construction and Implementation Mechanism of the Collaborative Training Model for "AI + Editing" Compound Talents (Project Number: CUJS2025-CBRH-030); Research Project of Chongqing Science and Technology Journal Editors Association in 2024: Research on the Path of Empowering Equal Dialogue between Editors and Authors of Medical Comprehensive Journals with New Quality Productivity; Open Fund Project of Shanghai Jiao Tong University - Deguit Press Joint Laboratory: Full-Chain Analysis of Retraction Factors in Academic Ethics and Dynamic Path Planning for Academic Misconduct Prevention (Project Number: STTU-DG-2024-003); and 2024 Editor · Renhe Fund: Research on Strategies for Journal Editorial Departments to Assist Deceived Authors in 'Anti-Fraud' under the New Situation (Project Number: XBRH2024-002-022).

Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Southwest University (protocol code SWU-ETH-2024-01-17-001, date of approval: 15 October 2024).

Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

Data Availability Statement

The data used in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper. The authors have no financial, personal, institutional, or other relationships that could inappropriately influence or bias the work reported in this manuscript.

AI Use Statement

During the preparation of this manuscript, the authors used the AI-assisted tool Grammarly solely for language refinement. No AI tools were used for data analysis, interpretation of results, or generation of scientific content. All outputs were critically reviewed and edited by the authors. The authors take full responsibility for the integrity and accuracy of the work.

References

1. Horbach, P.S.; Fishberg, R.; Ulpts, S.; et al. Thou Shalt Not!—How the institutional afterlife of research misconduct scandals shapes research integrity training. *J. Responsible Innov.* **2024**, *11*, 2414500. [CrossRef]
2. Wang, P.; Jia, Y. Exploration and Practice of Research Integrity Education for Graduate Students in Universities in the Digital Intelligence Era. *Pop. Lit. Art* **2025**, 135–137. [CrossRef]
3. Francesca, G.; Ceruti, S.; Martini, S.; et al. Educating and Training in Research Integrity (RI): A Study on the Perceptions and Experiences of Early Career Researchers Attending an Institutional RI Course. *J. Acad. Ethics* **2023**, *22*, 413–430. [CrossRef]
4. Dubbaka, S. Incorporating Implicit bias into research integrity education: Response to 'Why and how to incorporate issues of race/ethnicity and gender in research integrity education'. *Account. Res.* **2023**, *32*, 193–194. [CrossRef]
5. Labib, K., Evans, N., Pizzolato, D.; et al. Co-creating Research Integrity Education Guidelines for Research Institutions. *Sci. Eng. Ethics* **2023**, *29*, 28. [CrossRef]
6. Panpakdee, C.; Khanbut, K. Pathways to food security: Perceptions of resilience between organic farmers and extension agents in Thung Kala Ronghai, Thailand. *Cogent Food Agric.* **2025**, *11*, 2546989. [CrossRef]
7. Antunes, D.; Oliveira, R.; Paulino, M.; et al. Severe acute asthma exacerbations under biological agents: A new therapeutic paradigm? *Pulmonology* **2025**, *31*, 2571016. [CrossRef]
8. Najafi, F.; Ghasemian, N.; Safari, M.; et al. Poly(propylene imine) dendrimer as reducing agent for chloroauric acid to fabricate and stabilize gold nanoparticles. *J. Phys. Chem. Solids* **2021**, *148*, 109682. [CrossRef]
9. Moyn s, E.; Looms, J.; Jagpal, S.P.; et al. A review of cardiotoxic agent poisoning and the role of VA-ECMO as reported to the National Poisons Information Service (NPIS) 2018–2023. *Toxicol. Commun.* **2025**, *9*, 2524302. [CrossRef]
10. Lam, W.; Paynter, Q.; Zhang, Z.-Q. Functional response of Amblyseius herbicolus (Acari: Phytoseiidae) on Sericothrips staphylinus (Thysanoptera: Thripidae), an ineffective biocontrol agent of gorse. *Biol. Control* **2021**, *152*, 104468. [CrossRef]

11. Cao, Y.; Jiang, Y.; Zhao, Y. A study on the content of integrity policies and research integrity management in Chinese universities. *Front. Res. Metr. Anal.* **2023**, *8*, 943228. [CrossRef]
12. Vitiello, G.; Venezia, V.; Verrillo, M.; et al. Hybrid humic acid/titanium dioxide nanomaterials as highly effective antimicrobial agents against gram(−) pathogens and antibiotic contaminants in wastewater. *Environ. Res.* **2021**, *193*, 110562. [CrossRef]
13. Manne, A.A.; Arigela, B.; Giduturi, A.K.; et al. Pterocarpus marsupium Roxburgh heartwood extract/chitosan nanoparticles loaded hydrogel as an innovative wound healing agent in the diabetic rat model. *Mater. Today Commun.* **2021**, *26*, 101916. [CrossRef]
14. Davoodi, M.; Faryadi, S.; Velni, M.J. A Graph Theoretic-Based Approach for Deploying Heterogeneous Multi-agent Systems with Application in Precision Agriculture. *J. Intell. Robot. Syst.* **2021**, *101*, 10. [CrossRef]
15. Kaviarasan, B.; Kwon, O.-M.; Park, J.M.; et al. Stochastic faulty estimator-based non-fragile tracking controller for multi-agent systems with communication delay. *Appl. Math. Comput.* **2021**, *392*, 125704. [CrossRef]
16. Formanek, M. Exploring the potential of large language models and generative artificial intelligence (GPT): Applications in Library and Information Science. *J. Librariansh. Inf. Sci.* **2025**, *57*, 09610006241241066. [CrossRef]
17. Byun, J.H.; Jung, I.H. Phase-specific cancer-immune model considering acquired resistance to therapeutic agents. *Appl. Math. Comput.* **2021**, *391*, 125555. [CrossRef]
18. Lee, Y.H.; Park, J.Y.; Lee, I.Y.; et al. Effects of cryoprotective agents and treatment methods on sperm cryopreservation of stone flounder, Kareius bicoloratus. *Aquaculture* **2021**, *531*, 735969. [CrossRef]
19. Silva, L.d.L.; Souza, C.d.F.; Parodi, T.V.; et al. Ethanolic extract of Hyptis mutabilis (Rich.) Briq.: An effective sedative and antioxidant agent in fish. *Aquaculture* **2021**, *531*, 735940. [CrossRef]
20. Cosenza, B.; Popov, N.; Juurlink, B.; et al. Easy and efficient agent-based simulations with the OpenABL language and compiler. *Future Gener. Comput. Syst.* **2021**, *116*, 61–75. [CrossRef]
21. Hidayat, R.; Chowdhury, T.; Kim, Y.; et al. Density functional theory study on the reducing agents for atomic layer deposition of tungsten using tungsten chloride precursor. *Appl. Surf. Sci.* **2021**, *538*, 148156. [CrossRef]
22. Duan, A.; Chen, M.; Lu, Q. Design of Railway Customer Service Agent Based on Large Language Model. *Yangtze River Inf. Commun.* **2025**, *38*, 86–93. [CrossRef]
23. Roh, D.K.; Jae, H.; Mun, H.; et al. Precise tuning of morphology and pore size of amine-functionalized MIL metal–organic frameworks using a directing agent. *Mater. Sci. Eng. B* **2021**, *263*, 114833. [CrossRef]
24. Riddle, M.E.; Tatara, E.; Olson, C.; et al. Agent-based modeling of supply disruptions in the global rare earths market. *Resour. Conserv. Recycl.* **2021**, *164*, 105193. [CrossRef]
25. Lou, K.; Kang, A.; Xiao, P.; et al. Effects of basalt fiber coated with different sizing agents on performance and microstructures of asphalt mixture. *Constr. Build. Mater.* **2021**, *266*, 121155. [CrossRef]
26. Somboon, P.; Soontorngun, N. An actin depolymerizing agent 19,20-epoxycytochalasin Q of Xylaria sp. BCC 1067 enhanced antifungal action of azole drugs through ROS-mediated cell death in yeast. *Microbiol. Res.* **2021**, *243*, 126646. [CrossRef]
27. Mannekote, A.; Davies, R.; Pinto, J.D.; et al. Large language models for whole-learner support: Opportunities and challenges. *Front. Artif. Intell.* **2024**, *7*, 1460364. [CrossRef]
28. Lee, S.-Y.; Jang, D.-I.; Kim, D.-Y.; et al. UV laser decontamination of chemical warfare agent simulants CEPS and malathion. *J. Photochem. Photobiol. A* **2021**, *406*, 112989. [CrossRef]
29. Scoppa, V. Social pressure in the stadiums: Do agents change behavior without crowd support? *J. Econ. Psychol.* **2021**, *82*, 102344. [CrossRef]
30. Oliveira, L.d.F.; Camargos, E.F.; Martini, L.L.L.; et al. Use of psychotropic agents to treat agitation and aggression in Brazilian patients with Alzheimer's disease: A naturalistic and multicenter study. *Psychiatry Res.* **2021**, *295*, 113591. [CrossRef]
31. Zhang, L.; Zhang, J.; Li, Y.; et al. Design of Integrated Energy System Agent Based on Reinforcement Learning. *Electron. Des. Eng.* **2024**, *32*, 145–149. [CrossRef]
32. Yepes, A.; Ochoa-Bautista, D.; Murillo-Arango, W.; et al. Purple passion fruit seeds (Passiflora edulis f. edulis Sims) as a promising source of skin anti-aging agents: Enzymatic, antioxidant and multi-level computational studies. *Arab. J. Chem.* **2021**, *14*, 202101. [CrossRef]
33. Zupko, R. Application of agent-based modeling and life cycle sustainability assessment to evaluate biorefinery placement. *Biomass Bioenergy* **2021**, *144*, 105916. [CrossRef]
34. Munjita, S.M.; Kalonda, A.; Mubemba, B.; et al. Evidence of multiple bacterial, viral, and parasitic infectious disease agents in Mastomys natalensis rodents in riverine areas in selected parts of Zambia. *Infect. Ecol. Epidemiol.* **2025**, *15*, 2441537. [CrossRef]

35. Arita, T.; Namerikawa, T. Probabilistic multi-agent pose graph filtering on SE(3) via distributed ADMM and stein particle gradient descent. *SICE J. Control Meas. Syst. Integr.* **2025**, *18*, 2570573. [CrossRef]
36. Sushma, K.; Lakshmi, T.N.; Arunodhayam, K.; et al. In Vitro Evaluation of Fungicides, Botanicals and Biocontrol agents against Stagonosporopsis cucurbitacearum Causing Gummy Stem Blight. *J. Adv. Biol. Biotechnol.* **2025**, *28*, 1681–1691. [CrossRef]
37. Puttasakul, T.; Sirisangwang, P.; Aroonrote, N.; et al. Optimizing Dispersion of Silver Nanoparticle Incorporated Hydrogel Matrix by Silver Ion-Reducing Agent Self-Assembly. *ACS Omega* **2025**, *10*, 62667–62674. [CrossRef]
38. Vendé, B.; Barberousse, A.; Ruphy, S. From 2015 to 2023, eight years of empirical research on research integrity: a scoping review. *Res. Integr. Peer Rev.* **2025**, *10*, 5. [CrossRef]
39. Elayaraja, S.; Liu, G.; Zagorsek, K.; et al. TEMPO-oxidized biodegradable bacterial cellulose (BBC) membrane coated with biologically-synthesized silver nanoparticles (AgNPs) as a potential antimicrobial agent in aquaculture (In vitro). *Aquaculture* **2021**, *530*, 735746. [CrossRef]
40. Mahale, S.A.; Dhadse, P.; Mahale, A.; et al. In vitro assessment of copper nanoparticles gel efficacy as a futuristic drug delivery agent against periodontal pathogens. *J. Taibah Univ. Sci.* **2025**, *19*, 2549161. [CrossRef]
41. Bermejo-Casado, R.; Bazaga-Fernández, I.; Carrasco-Granger, S. Law enforcement agents at European borders: perceptions of migration, and the networked and relational exercise of discretion in policy implementation. *Polit. Res. Exch.* **2025**, *7*, 2561948. [CrossRef]
42. Xu, W. Dilemmas and Countermeasures for Research Integrity Education of Graduate Students. *Tianjin Sci. Technol.* **2025**, *52*, 95–102. [CrossRef]
43. Nitayaphat, W.; Jintakosol, T. Sustainable Dyeing of Sugarcane Leaves-Cotton Blended Yarn with Dayak Onion (Eleutherine americana) Bulb Extract Using Eco-Friendly Mordanting Agents. *J. Nat. Fibers* **2025**, *22*, 2511998. [CrossRef]
44. Sousa, N. Ethical and practical implications of AI in academic library research. *J. Librariansh. Inf. Sci.* **2025**, *57*, 03400352251391753. [CrossRef]
45. Leong, H.J.-Y.; Tan, H.-D.; Teoh, M.-L. In silico discovery of skin depigmenting agents from Chlorella vulgaris: A network pharmacology and molecular simulation approach. *Appl. Phycol.* **2025**, *6*, 350–365. [CrossRef]
46. Grzyb, T.; Kulczyk, S. Well-being derived from the riverscape: Linking comfort and discomfort agents. *Ecosyst. People* **2025**, *21*, 2578178. [CrossRef]
47. Iwata, S.; Yamasita, N.; Asukabe, K.; et al. Discovery and characterization of antitumor gut microbiota from amphibians and reptiles: Ewingella americana as a novel therapeutic agent with dual cytotoxic and immunomodulatory properties. *Gut Microbes* **2025**, *17*, 2599562. [CrossRef]
48. Pavlovic, A.; Rajovic, N.; Masic, S.; et al. Assessing attitudes toward research and plagiarism among medical students: A multi-site study. *Res. Integr. Peer Rev.* **2024**, *19*, 11. [CrossRef]
49. Donabauer, M.; Sawall, S.; Ayx, I.; et al. Development and characterization of new contrast agents for Photon-Counting CT. *Eur. J. Radiol.* **2025**, *195*, 112643. [CrossRef]
50. Snellman, J.E.; Barrio, R.A.; Kaski, K.K. Social structure formation in a network of agents playing a hybrid of ultimatum and dictator games. *Physica A* **2021**, *561*, 125257. [CrossRef]