Effects of Proactive Interaction and Instructor Choice in AI-Generated Virtual Instructors for Financial Education

Thanawit Prasongpongchai
Beacon Interface
KASIKORN Business-Technology Group
Nonthaburi, Thailand
thanawit.p@kbtg.tech

Pat Pataranutaporn
MIT Media Lab
Massachusetts Institute of Technology
Cambridge, MA, United States
patpat@media.mit.edu

Auttasak Lapapirojn
KASIKORN Labs
KASIKORN Business-Technology Group
Nonthaburi, Thailand
auttasak.l@kbtg.tech

Chonnipa Kanapornchai
Beacon Interface
KASIKORN Business-Technology Group
Nonthaburi, Thailand
chonnipa.k@kbtg.tech

Joanne Leong
MIT Media Lab
Massachusetts Institute of Technology
Cambridge, MA, United States
joaleong@media.mit.edu

Pichayoot Ouppaphan
KASIKORN Labs
KASIKORN Business-Technology Group
Nonthaburi, Thailand
pichayoot.o@kbtg.tech

Kavin Winson
KASIKORN Labs
KASIKORN Business-Technology Group
Nonthaburi, Thailand
kavin.w@kbtg.tech

Monchai Lertsutthiwong
KASIKORN Labs
KASIKORN Business-Technology Group
Nonthaburi, Thailand
monchai.le@kbtg.tech

Pattie Maes
MIT Media Lab
Massachusetts Institute of Technology
Cambridge, MA, United States
pattie@media.mit.edu

***Abstract*—This research full paper describes a web-based online learning platform that delivers financial literacy lessons via talking head videos of AI-generated personas with two additional core features: LLM-powered proactive chat-based question-and-answer interactivity, and personal choice of the AI instructor from a list of distinct personas. We conducted two comparative studies with a total of 233 Thai students aged 18-25, which aim to 1) investigate the impact of interactivity and instructor selection on the learning experience, and 2) further explore the underlying factors at play with instructor selection by introducing AI instructors' backstories as an extra intervention. We found that enabling interactivity significantly enhanced learning motivation, perceived learning facilitation, engagement, and virtual instructors' humanness compared to the passive setting. Providing learners with a choice of AI instructors provided minimal additional benefit. However, the learner's feeling of relatedness toward the instructor is a significant positive predictor of learning motivation, positive emotion, and agent credibility, while goal alignment with the agent correlates with perceived learning facilitation, and admiration corresponds with perceived agent humanness. These findings underscore the potential of interactive virtual instructors—ones that interactively encourage learners to reflect on the teaching materials throughout the lesson through two-way interaction—in enhancing motivational and experiential aspects of remote education, even if they do not significantly impact comprehension, and the importance of promoting learner's relatedness and goal alignment with the agent in boosting other aspects of the learning experience.**

***Index Terms*—AI-Generated Virtual Instructor, Human-AI Interaction, Personalized AI, AI for Education, Pedagogical Agent, Financial Literacy Education.**

I. Introduction

Online remote education, through means such as e-learning and video-sharing platforms, has become increasingly popular. It is used by both instructors as supplemental material for traditional classes and by learners as additional learning resources. Among online educational content, multiple styles of "virtual instructors," or "pedagogical agents," have been introduced as an alternative mode of delivering lesson content [1]. Prior works have explored designing such virtual agents with varying characteristics, e.g., age, liked or admired public figures, fictional characters, and historical personas [2]–[5]. These characters engage a learner's socio-emotional processes and can aid performance, with learners recognizing emotions in virtual tutors [6] and showing activations in socio-emotional brain regions during learning [7].

At the same time, financial literacy—the set of skills needed to make sound financial decisions, including concepts such as budgeting, saving, investing, and managing credit [8]—remains unfortunately undertaught in most school systems. A disparity between countries globally is also present as the level of financial literacy in populations of emerging economies lags behind more advanced economies. This leaves many individuals worldwide who may lack valuable skills to make important financial decisions [9].

Given these insights, we are motivated to explore using generative artificial intelligence (AI) and large language models (LLMs) to build personalized virtual instructors that could support financial literacy education through increased learners' motivation and engagement. We developed a web-based financial literacy education platform featuring: 1) *LLM-powered two-way interactivity*, where the virtual instructors proactively ask questions during the session and learners can respond and converse with the instructor in a chat interface, and 2) *personalization via instructor selection*, where learners can choose an instructor from a pool of pre-generated agents based on their appearances, goals, and personalities.

This paper presents two studies. The first study investigates the impact of interactivity and instructor selection on the learning experience, while the second further probes into instructor selection and the resulting learners' sense of autonomy and relatedness by introducing backstories for virtual instructors before learners make their instructor choices.

This paper offers two novel contributions: 1) An AI-powered financial literacy education system with added two-way question-and-answer interactivity, and personalization through instructor selection, and 2) Results of two comparative studies across a total of 233 participants investigating the effects of said interactivity and personalization, as well as learners' perceptions toward the AI instructor, on learning motivation, experience, and comprehension.

II. Related Work

Our work on understanding the role of AI-generated interactive virtual instructors is situated within three different contexts of research as follows:

A. Learning Motivation and Personalization

Learner-centricity and personalization in education have recently gained increasing popularity among researchers, educators, and education policymakers [10], [11]. Within the area of personalized learning, *context personalization* is a technique where content is aligned with one's interests and background. It is a common strategy that has been shown to impact students' attention and engagement by leveraging situational interests [12]–[15], the starting phase of developing interests according to the four-phase model of interest development [16]. One related framework, Self-Determination Theory (SDT) [17], also suggests basic human psychological needs of autonomy, competence, and relatedness to be crucial components of motivation in education and other settings [18]. Research has also shown the effect of choice on students' motivation through an increased sense of autonomy [19], [20].

Past efforts have sought to inject personal details into traditional materials, requiring extensive manual effort and failing to scale broadly [21]–[24]. With its capacity for generating on-demand text dynamically customized to virtually any topic, generative AI could overcome these limitations. Yet, its potential applications have not been widely explored. In contrast to manual approaches to personalization, generative AI offers a scalable method for dynamically producing tailored content by directly interacting with learners to incorporate their details across topics [15]. Key challenges remain surrounding developing best practices for generating maximally motivating narratives [25]. We intend to explore these issues to pave the path forward for AI-enabled personalized learning.

B. AI-Generated Characters

Recent progress in generative artificial intelligence has made creating animated virtual characters that imitate the appearance of other personalities become feasible. Work has been done to explore using these techniques to create virtual instructors and storytellers with the likeness of contemporary, historical, or fictional personalities [3], [5]. This method offers significant prospects for the dynamic creation and interaction with such characters, which could be effectively utilized across educational platforms, wellness interfaces, and entertainment mediums [5], [26], [27]. These representations are algorithmically generated using machine learning techniques [28], [29], allowing for the production of highly realistic representations. Moreover, AI-generated exchanges utilizing generative models and large language models (LLMs) have made simulated conversations possible, further enhancing the realism of these digital avatars [30], [31]. This LLM adoption has greatly expanded integrating humans with computers, moving far past human-computer interaction into the developing area of human-AI interaction [32]–[36].

C. Virtual Pedagogical Agents

With the increase in remote learning, "virtual instructors" or "pedagogical agents" have been developed to facilitate video-based teaching. These digitally synthesized instructors, encompassing 2D, 3D, and other formats, provide an alternative to real human teachers [1]. Research shows significant learning improvements with virtual instructors. Learners can effectively recognize emotions from virtual tutors [6], with instructor appearance influencing motivation and outcomes – a young, attractive persona yielded superior engagement over an elderly, less attractive counterpart [2]. Brain imaging during learning with a virtual instructor revealed enhanced activity in social processing regions and better performance, highlighting socio-emotional processes in play [7]. Recent work has explored customizing virtual instructors based on admired individuals to boost engagement. In a study of 134 students, an AI-generated instructor resembling someone the learner likes/admires increased motivation and positivity despite similar test performance [4]. Separately, "living memories" of historical figures like Da Vinci created from their journals were more engaging for learning than reading the raw texts [3]. These insights on virtual characters for education motivate further research for pedagogical agents.

III. System Design

We developed a financial literacy e-learning web platform featuring AI-generated agents as virtual instructors to evaluate the potential of 1) *proactive two-way interactivity*: the ability for the instructors to ask learners questions throughout the lesson and follow up on their responses, as well as allowing learners to pose questions they might have to the instructor, and 2) *personal instructor selection*: the ability for learners to select a preferred instructor from a set of pre-defined personas. These features aim to boost learner's participation and thereby improve their motivation and their learning experience.

A. Functional Design

The overall flow of the learning platform is depicted in Fig. 1. Learners start by selecting an AI-generated persona to act as their "AI virtual instructor," then proceed to the learning environment where the lesson is delivered through a pre-rendered talking head video of the selected virtual instructor, accompanied by a slideshow. Throughout the lesson, the AI instructor asks users questions at certain points specified in the script. Learners can respond via a pop-up chat interface, to which the instructor dynamically responds and asks follow-up questions according to prompts embedded in the script. Moreover, learners can also use the "Raise Hand" button to ask questions and get dynamic answers generated by a GPT-4-based LLM via OpenAI API. This addition of virtual instructor-initiated interactions is designed to be proactive to foster bidirectional engagement throughout the lessons rather than relying on learners to come up with questions themselves.

The talking head video creation pipeline is based on [5]. First, a lecture script for the virtual instructor is written and converted into voice through an online text-to-speech service. Separately, a headshot of the virtual instructor is generated using a diffusion-based image generation tool. Finally, the headshot is then animated with a driving video for facial motion [37] and lip-synced with the voice track using Wav2Lip [38]. Additionally, a mark-up file detailing timestamps for slideshow pages, pauses for learner's answers, and custom LLM prompts are written to be played back with the video (Fig. 2). This pipeline can be compiled into a tool to enable non-technical experts to create engaging interactive lessons.

For the chat interaction, the LLM is initially prompted with lesson content and instructions to act as the selected virtual instructor, to respond in a short, conversational tone, not to answer about other topics to avoid inaccuracies, and to instruct the user to press the "Back to Lesson" button when satisfied with the learners' answers. Additional custom prompts in the lesson script are injected into the LLM at specified points as also seen in Fig. 2.

B. Personalized Virtual Instructors

Intending to boost personalization, we generated three agents, *Luna*, *Captain*, and *Ura*, to serve as selectable virtual instructors (see Fig. 2). Each agent is designed to have different personalities: *Luna* is a gentle and cheerful female instructor, *Captain* is a confident and serious male instructor, and *Ura* is a playful and friendly cartoon elephant instructor.

The lesson scripts are crafted to be conversational to facilitate the two-way interactions. While the overall content stays the same across the three characters, each persona's script is slightly different in their tone to highlight their personalities. Brief explanations of the chat and hand-raise features were added to the scripts as ice-breakers to familiarize students with these features before diving into the main content.

C. Lesson Content

For the purpose of evaluating the platform, the script for the AI instructors was adapted from the Six Jars Method for basic money management, created by T. Harv Eker [39]. The method proposed a framework for budgeting one's income into six parts or "jars," including Necessities, Financial Freedom, Long-Term Savings, Education, Play, and Give, with certain recommended proportions. We chose this topic because of its simplicity which fits well as an introduction to personal finance for our target student demographic. The lessons in this research were written in Thai to best serve our target audience.

IV. Study 1

The first study aimed to evaluate the impact of interactivity and instructor selection on comprehension, learning experience, and perceptions toward the virtual instructors.

A. Methodology

To assess the impact of the chat-based interactivity and instructor selection, we designed a randomized control study with three experimental conditions: *A) Passive Virtual Instructor*, where learners are assigned *Luna* as their instructor with chat functionality removed. This acts as our control condition, *B) Interactive Virtual Instructor*, where *Luna* teaches the lesson with the chat interactivity enabled, and *C) Personally Selected Interactive Virtual Instructor* where chat interactivity and the ability to select a preferred instructor are enabled.

We recruited 96 Thai students aged 18-25 and randomly assigned them to the three conditions. After taking the Six Jars Method lesson through our platform, participants took a 10-point assessment quiz to measure their comprehension and completed questionnaires comprising surveys measuring learning motivation and positive emotion (adapted from [4]), as well as established metrics for pedagogical agents adapted from [40], [41], quantifying learners' perceptions on learning facilitation, engagement, agent credibility, and agent's human-likeness. These metrics were measured by averaging the responses to multiple 5-point Likert rating questions for each subscale. Each session lasted approximately 15 minutes.

Data analysis was conducted using one-way analysis of variance (ANOVA) to compare the learning experience metrics between the three conditions. Dependent variables include test scores, learning motivation, positive emotion, perceived learning facilitation, agent credibility, engagement, and agent humanness. Benjamini-Hochberg correction was performed to control false discovery rate [42]. Statistically significant main effects were followed up with Tukey's post-hoc tests to assess each condition pair.

B. Results

This study aims to analyze two main factors: 1) how the chat-based interactivity affects learning when compared to the passive setting, and 2) how the personal selection of AI virtual instructors influences the learning experience.

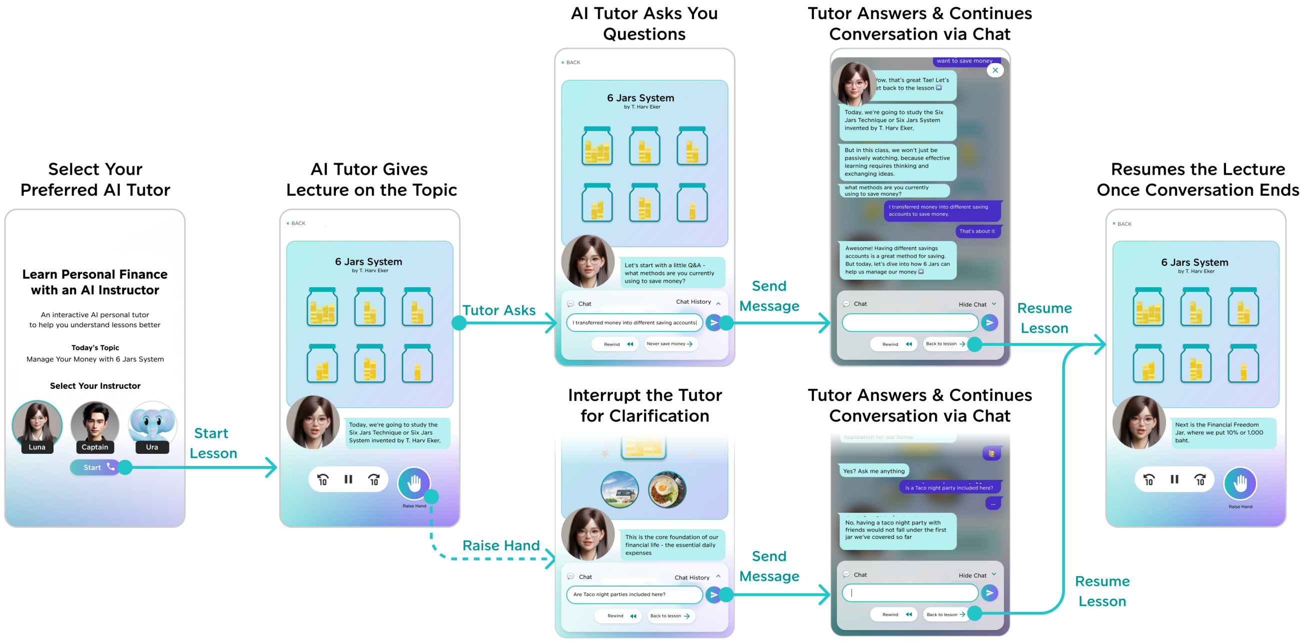

Fig. 1. The overall system of AI-generated virtual instructors for financial literacy.

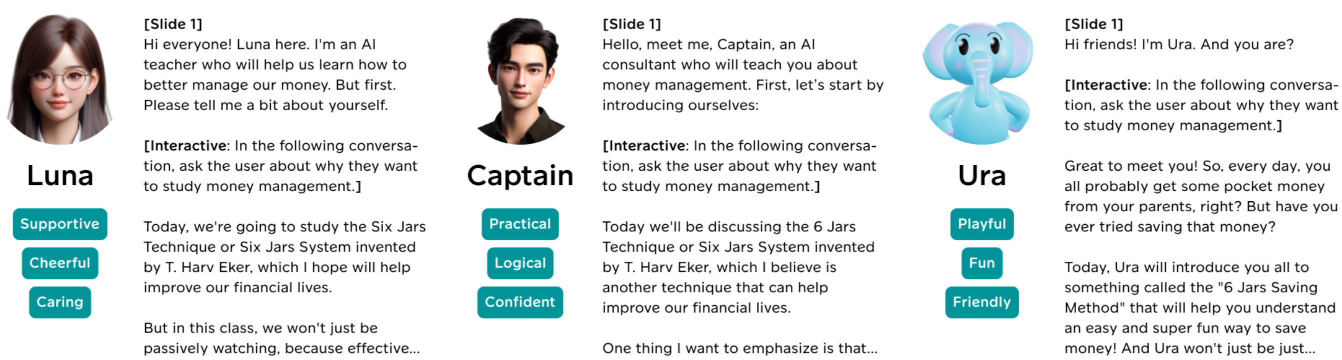

Fig. 2. Examples of scripts and in-lesson embedded prompts. Each persona's script is infused with their distinct personality while delivering similar content.

The results in Table I and Fig. 3 show that the added LLM-powered interactivity enhances learning motivation, perceived learning facilitation, engagement, and perceived agent humanness, while personalization via instructor selection provided little extra benefit. Nevertheless, both interventions did not significantly alter content comprehension levels as quantified by test scores from the assessment quiz.

The post-hoc Tukey's tests revealed significant increases in multiple learner experience metrics between the Passive and Interactive Virtual Instructor conditions, including learning motivation (p = .003), learning facilitation (p = .003), engagement (p = .004), and humanness (p = .005). No significant improvements in positive emotion and agent credibility between the two groups were found ($p > .05$). Additionally, an increase in agent humanness was also found between the Passive Virtual Instructor and the Personally Selected Interactive Virtual Instructor conditions (p = 0.02).

Regarding the minimal added benefit of instructor choice despite this having been demonstrated in prior literature [19], one explanation could be that the options and the way they are presented might not provide enough information for learners to make a sufficiently meaningful choice. Hence, further investigation was needed to examine the effects of instructor choice on the learning experience.

V. Study 2

In response to the results of the first experiment and the hypothesis that the given instructor choice might not have been adequately meaningful to the learners, we conducted a second study by adding a backstory for each instructor as an intervention and probing the learners' sentiment toward their virtual instructors and their sense of autonomy and relatedness, which are key components of motivation according to the Self-Determination Theory [18].

TABLE I
MEASURED MEAN AND STANDARD DEVIATION (IN PARENTHESES) FOR EACH LEARNER EXPERIENCE METRIC ACROSS THE THREE CONDITIONS IN STUDY 1, WITH RESULTS FROM THE F-TEST FOR STATISTICAL SIGNIFICANCE *(in italics)*.

Condition	Test Scores (Out of 10)	Post-Lesson Metrics (1-5 Likert) Learning Motivation	Positive Emotion	Learning Facilitation	Engage-ment	Agent Credibility	Agent Humanness
A) Passive	8.48	3.75	3.80	3.77	3.63	4.07	2.88
	(1.23)	(0.65)	(0.80)	(0.72)	(0.79)	(0.76)	(1.03)
B) Interactive	8.13	4.41	4.19	4.45	4.29	4.32	3.69
	(1.36)	(0.83)	(0.78)	(0.83)	(0.67)	(0.96)	(1.09)
C) Personalized, Interactive	8.21	4.15	3.97	4.05	3.85	4.07	3.55
	(1.45)	(0.82)	(0.88)	(0.80)	(0.89)	(0.88)	(0.88)
F(2, 93)	*0.60*	*5.99*	*1.82*	*5.91*	*5.71*	*0.87*	*6.02*
p-value	*0.551*	*0.012**	*0.235*	*0.009***	*0.008***	*0.492*	*0.024**

*p<0.05 level (two-tailed), **p<0.01, ***p<0.001, p-values are adjusted with Benjamini-Hochberg correction.

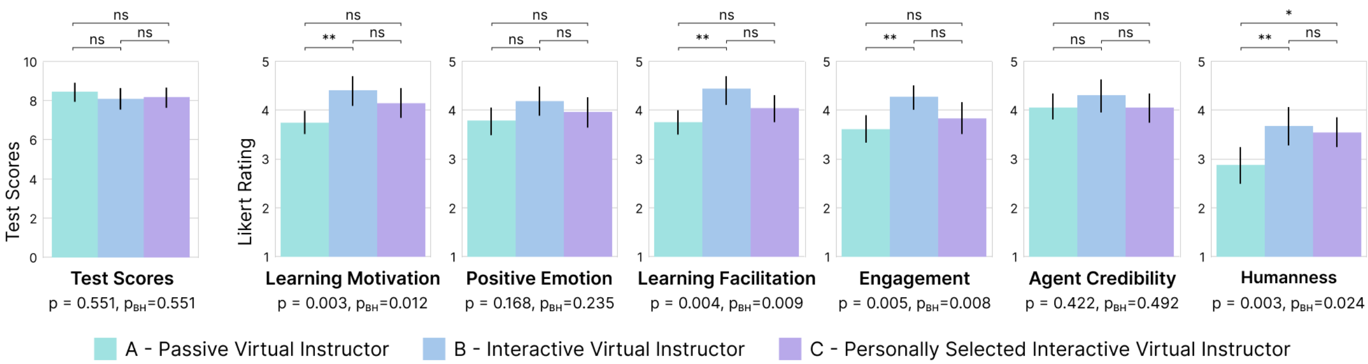

Fig. 3. Measured learning experience metrics from Study 1.

A. Methodology

Self-Determination Theory (SDT) suggests learner motivation stems from autonomy, competence, and relatedness to the learning environment [18]. Research also shows virtual instructors based on liked or admired personas can boost motivation and engagement [4]. Thus, to enhance autonomy and relatedness, we created 30-second introductory videos for each character (Fig. 4). These videos feature backstories, personal goals, and distinct personalities: *Luna* focuses on happiness through financial literacy, *Captain* on wealth, and *Ura* on having fun. Each video includes personality-reflecting music and visuals depicting each agent's "lifestyle" (Fig. 5).

To evaluate the effect of the added introductions as well as investigate how learner's perception toward the choice affects the learning experience, we conducted another study with a 2×2 factorial design featuring two interventions: "Choice," the presence of the ability to personally select an instructor (as opposed to being randomly assigned) and "Introduction," the presence of the instructor backstory videos. This results in four between-group conditions: *A) Control* where the learners were assigned a random character as the instructor without being shown their introduction video, *B) Introduction Condition* with randomly assigned instructor accompanied by the backstory video, *C) Choice Condition* where the learners select an instructor by only seeing the instructors' appearances and names but not the introduction videos. (This condition is identical to

Fig. 4. Pre-instructor selection screen in Study 2 featuring instructor introduction videos.

the Personally Selected Interactive Virtual Instructor in Study 1), and *D) Introduction* $*$ *Choice Condition* where learners watched the introduction videos of the three agents in random order before making the selection.

Once participants selected or were assigned an instructor, they completed a survey with 7-point Likert scale questions measuring the following subscales: 1) their liking toward the AI instructor ("I like this AI instructor"), 2) their admiration toward the instructor ("I admire this AI instructor"), 3) whether

Fig. 5. Self-introduction scripts and sample imagery used in the introductory videos for each virtual character in Study 2.

their goals align with the instructor's stated goals ("I agree with this AI instructor's goals"), 4) their sense of autonomy in the selection, and 5) their perceived relatedness toward the instructor. Questions for 4) and 5) were adapted from the autonomy and relatedness subscales of the Intrinsic Motivation Inventory (IMI) [18], [43]. For this study, these five subscales are considered "post-(instructor) selection metrics".

Then, the participants took the Six Jars Method lesson with the interactive agent (with the LLM-based question-and-answer feature enabled). This was followed by an assessment quiz and surveys measuring their learning experience including learning motivation, positive emotion, perceived learning facilitation, agent credibility, engagement, and agent humanness. These quiz and surveys are the same as in Study 1. We will refer to these metrics as "post-lesson metrics."

The participants of this second study consisted of additional 137 recruited Thai students aged 18-25 who were randomly assigned to the aforementioned four conditions.

Three groups of multiple linear regression analyses were conducted to investigate the relationship between each intervention or metric on other metrics: 1) between each post-selection metric and the interventions (including the interaction term "introduction $*$ choice"), 2) between each post-lesson metric and the interventions (including the interaction term), and 3) between each post-lesson metric and the post-selection metrics (no interaction terms). The resulting regression coefficients β were used to determine important factors that act as predictors of other measurements. Benjamini-Hochberg correction was also applied to this set of measurements with 71 hypotheses in total among these three groups.

B. Results

The objectives of this second study are two-fold: 1) to determine whether the introduced interventions (instructor selection and introduction videos) affect the post-selection and post-lesson metrics, and 2) to determine whether there are underlying relationships between the post-selection and post-lesson metrics. *Post-selection metrics* refer to liking, admiration, goal alignment toward the virtual instructor, perceived autonomy in instructor choice, and relatedness toward the selected instructor while *post-lesson metrics* include test scores, learning motivation, positive emotion, perceived learning facilitation, engagement, agent credibility, and agent humanness.

The means and standard deviations of each resulting metric are displayed in Table II. Regression analyses between the measured metrics (post-selection and post-lesson), and the interventions unveiled the Choice intervention as a significant positive predictor of perceived autonomy ($\beta = 1.23, p = 0.003$) while no other significant changes in these metrics due to the interventions were observed ($p > 0.05$). These are illustrated in Table III and IV.

Regression analyses of post-selection and post-lesson metrics further revealed that relatedness is a significant positive predictor for learning motivation ($\beta = 0.20, p = 0.027$), positive emotion ($\beta = 0.31, p = 0.005$), and agent credibility ($\beta = 0.19, p = 0.023$), as shown in Table V. Additionally, admiration toward the AI instructor is also a significant positive predictor of perceived agent humanness ($\beta = 0.40, p < 0.001$), and goal alignment is a significant positive predictor of perceived learning facilitation ($\beta = 0.19, p = 0.023$). No significant effects between other post-lesson versus post-selection metrics were found ($p > 0.05$).

VI. Discussion and Future Work

The two comparative studies of our AI-generated virtual instructor platform revealed the effects of LLM-powered two-way interactivity and personalized instructors on the learning experience and perception of the virtual agent. Study 1 demonstrated that two-way interactivity between the learner and the AI instructor had significant positive effects on the learning experience, including increased learning motivation, perceived learning facilitation, engagement, and the AI instructor's "humanness" compared to passive formats. This may be due to how the interactivity mimics the dynamics of a classroom, and how it keeps learners engaged by directly asking them questions throughout the lesson, encouraging active engagement with the lesson content.

On the other hand, Study 1 also showed minimal extra benefit from instructor selection, despite prior research showing that AI-generated instructors based on liked or admired figures can enhance the learning experience [4]. Since our AI instructors are not based on real people, we hypothesized that

TABLE II
MEASURED MEAN AND STANDARD DEVIATION (IN PARENTHESES) FOR EACH LEARNER EXPERIENCE METRIC BY CONDITION FROM STUDY 2.

	Post-Selection Metrics (1-7 Likert)					Test	Post-Lesson Metrics (1-5 Likert)					
Condition	Liking	Admiration	Goal Alignment	Autonomy	Relatedness	Scores (0-10)	Learning Motivation	Positive Emotion	Learning Facilitation	Engagement	Agent Credibility	Agent Humanness
A) Control	5.26	4.86	5.60	4.30	4.81	8.26	4.34	4.36	4.39	4.16	4.65	3.75
	(1.75)	(1.73)	(1.59)	(1.44)	(1.10)	(1.27)	(0.79)	(0.85)	(0.69)	(0.87)	(0.55)	(0.95)
B) Introduction	5.81	5.36	6.06	4.53	4.88	7.64	4.27	4.25	4.26	4.33	4.43	3.68
	(1.12)	(1.36)	(1.17)	(1.15)	(0.82)	(1.94)	(0.75)	(0.84)	(0.80)	(0.70)	(0.63)	(0.98)
C) Choice	5.91	5.57	5.89	5.53	4.89	7.60	4.41	4.10	4.37	4.24	4.43	3.81
	(1.22)	(1.54)	(1.28)	(1.12)	(1.14)	(1.88)	(0.65)	(0.95)	(0.71)	(0.77)	(0.66)	(1.05)
D) Intro. $*$ Choice	5.71	5.35	5.94	5.30	5.08	8.13	4.24	4.27	4.23	4.11	4.38	3.57
	(1.37)	(1.54)	(1.29)	(1.33)	(1.13)	(2.00)	(0.87)	(0.88)	(0.87)	(1.11)	(0.82)	(1.25)

TABLE III
REGRESSION COEFFICIENTS (β) AND 95% CONFIDENCE INTERVAL ON POST-SELECTION METRICS WITH INTERVENTION CONDITIONS AS PREDICTORS

	Post-Selection Metrics (1-7 Likert)	Liking	Admiration	Goal Alignment	Autonomy	Relatedness
Intervention Terms	Introduction	0.55	0.50	0.46	0.23	0.07
		$[-0.10, 1.20]$	$[-0.22, 1.23]$	$[-0.17, 1.09]$	$[-0.36, 0.83]$	$[-0.43, 0.56]$
		$(p = 0.347)$	$(p = 0.437)$	$(p = 0.459)$	$(p = 0.672)$	$(p = 0.887)$
	Choice	0.66	0.71	0.29	1.23**	0.09
		$[0.00, 1.31]$	$[-0.02, 1.45]$	$[-0.35, 0.92]$	$[0.63, 1.83]$	$[-0.41, 0.58]$
		$(p = 0.292)$	$(p = 0.282)$	$(p = 0.649)$	$(p = 0.003)$	$(p = 0.884)$
	Introduction $*$ Choice	-0.75	-0.72	-0.41	-0.46	0.12
		$[-1.69, 0.19]$	$[-1.77, 0.33]$	$[-1.31, 0.50]$	$[-1.32, 0.39]$	$[-0.59, 0.83]$
		$(p = 0.371)$	$(p = 0.416)$	$(p = 0.626)$	$(p = 0.577)$	$(p = 0.876)$

*p<0.05 level (two-tailed), **p<0.01, ***p<0.001, p-values are adjusted with Benjamini-Hochberg correction.

TABLE IV
REGRESSION COEFFICIENTS (β) AND 95% CONFIDENCE INTERVAL ON POST-LESSON METRICS WITH INTERVENTION CONDITIONS AS PREDICTORS

	Post-Lesson Metrics (1-5 Likert)	Test Scores (Out of 10)	Motivation	Positive Emotion	Learning Facilitation	Engagement	Agent Credibility	Agent Humanness
Intervention Terms	Introduction	-0.65	-0.07	-0.11	-0.13	0.17	-0.22	-0.07
		$[-1.48, 0.19]$	$[-0.44, 0.29]$	$[-0.53, 0.30]$	$[-0.49, 0.23]$	$[-0.24, 0.58]$	$[-0.54, 0.09]$	$[-0.57, 0.43]$
		$(p = 0.402)$	$(p = 0.869)$	$(p = 0.820)$	$(p = 0.688)$	$(p = 0.661)$	$(p = 0.450)$	$(p = 0.892)$
	Choice	-0.60	0.10	-0.28	-0.02	0.09	-0.22	0.08
		$[-1.44, 0.24]$	$[-0.27, 0.46]$	$[-0.70, 0.14]$	$[-0.38, 0.34]$	$[-0.32, 0.51]$	$[-0.54, 0.10]$	$[-0.42, 0.58]$
		$(p = 0.461)$	$(p = 0.827)$	$(p = 0.440)$	$(p = 0.939)$	$(p = 0.853)$	$(p = 0.422)$	$(p = 0.877)$
	Introduction $*$ Choice	1.12	-0.10	0.29	-0.01	-0.32	0.20	-0.15
		$[-0.09, 2.33]$	$[-0.62, 0.43]$	$[-0.31, 0.90]$	$[-0.53, 0.50]$	$[-0.91, 0.27]$	$[-0.25, 0.66]$	$[-0.87, 0.57]$
		$(p = 0.331)$	$(p = 0.880)$	$(p = 0.625)$	$(p = 0.962)$	$(p = 0.592)$	$(p = 0.637)$	$(p = 0.875)$

*p<0.05 level (two-tailed), **p<0.01, ***p<0.001, p-values are adjusted with Benjamini-Hochberg correction.

the way the choices were presented did not provide sufficiently meaningful information to make learners like, admire, or relate to these personas, thus making the decision not adequately meaningful to affect learners' experiences.

Study 2, aimed at addressing this gap, probed underlying factors and experimented with a way to modulate the experience through introduction videos for each AI instructor. Although the results demonstrated that our introduction videos did not yield significant improvements in the learning experience metrics even though the instructor choice did significantly raise perceived autonomy, several observations were found regarding the underlying factors. An explanation could be that the effects of instructor selection are small compared to how learners were exposed to the interactive features throughout the lesson; the learning experience may be more a result of how the entire session itself went rather than the initial choices. This is also evident in that most post-lesson measurements in the second study already yielded high scores, within 4 to 5 on the 5-point Likert scale, even for the control group.

Furthermore, out of the instructor selection-related factors (i.e., learners' liking, admiration, goal alignment, relatedness toward the instructor, and perceived autonomy), we found through multiple regression analysis that relatedness is a significant positive predictor of learning motivation, positive emotion, and agent credibility. Also, goal alignment is a predictor of perceived learning facilitation and admiration is a predictor of perceived agent humanness. These results highlight the importance of the learner's feeling of relatedness, goal alignment, and admiration toward a pedagogical agent in boosting the experience. This led to other possible explanations for why

TABLE V

REGRESSION COEFFICIENTS (β) AND 95% CONFIDENCE INTERVAL ON POST-LESSON METRICS WITH POST-SELECTION METRICS AS PREDICTORS

Post-Selection Metrics Terms (1-7 Likert)	Post-Lesson Metrics (1-5 Likert)	Test Scores (Out of 10)	Motivation	Positive Emotion	Learning Facilitation	Engagement	Agent Credibility	Agent Humanness
	Liking	0.20	0.08	-0.02	-0.01	0.02	0.04	-0.03
		$[-0.17, 0.56]$	$[-0.05, 0.20]$	$[-0.18, 0.14]$	$[-0.13, 0.11]$	$[-0.13, 0.16]$	$[-0.08, 0.15]$	$[-0.20, 0.14]$
		$(p = 0.562)$	$(p = 0.532)$	$(p = 0.881)$	$(p = 0.933)$	$(p = 0.881)$	$(p = 0.745)$	$(p = 0.880)$
	Admiration	-0.26	0.12	-0.03	0.10	0.15	0.08	0.40***
		$[-0.56, 0.04]$	$[0.01, 0.23]$	$[-0.16, 0.10]$	$[-0.00, 0.20]$	$[0.03, 0.27]$	$[-0.01, 0.18]$	$[0.26, 0.54]$
		$(p = 0.354)$	$(p = 0.185)$	$(p = 0.857)$	$(p = 0.297)$	$(p = 0.117)$	$(p = 0.345)$	$(p < 0.001)$
	Goal Alignment	-0.04	0.06	0.11	0.19*	0.16	0.10	0.06
		$[-0.39, 0.30]$	$[-0.06, 0.18]$	$[-0.05, 0.26]$	$[0.07, 0.31]$	$[0.02, 0.30]$	$[-0.01, 0.21]$	$[-0.10, 0.23]$
		$(p = 0.889)$	$(p = 0.626)$	$(p = 0.436)$	$(p = 0.023)$	$(p = 0.180)$	$(p = 0.367)$	$(p = 0.664)$
	Autonomy	-0.08	-0.05	0.08	-0.01	-0.01	-0.08	-0.06
		$[-0.36, 0.20]$	$[-0.15, 0.05]$	$[-0.05, 0.20]$	$[-0.10, 0.09]$	$[-0.12, 0.10]$	$[-0.17, 0.01]$	$[-0.19, 0.08]$
		$(p = 0.829)$	$(p = 0.662)$	$(p = 0.483)$	$(p = 0.941)$	$(p = 0.923)$	$(p = 0.365)$	$(p = 0.654)$
	Relatedness	0.29	0.20*	0.31**	0.18	0.16	0.19*	0.08
		$[-0.07, 0.65]$	$[0.08, 0.33]$	$[0.15, 0.47]$	$[0.05, 0.30]$	$[0.02, 0.30]$	$[0.08, 0.31]$	$[-0.09, 0.25]$
		$(p = 0.368)$	$(p = 0.027)$	$(p = 0.005)$	$(p = 0.051)$	$(p = 0.193)$	$(p = 0.023)$	$(p = 0.675)$

*p<0.05 level (two-tailed), **p<0.01, ***p<0.001, p-values are adjusted with Benjamini-Hochberg correction.

the introductions were not enough to establish relatedness and significant effects on other outcomes, namely that they may be too short of an exposure, the characters may look too artificial, and that the introduction video did not involve two-way engagement. This calls for further investigation into the types of intervention that would successfully boost relatedness and thereby the learning experience.

Both studies also noted no differences in test scores across groups. This could be due to participants' initial high performance and/or the lesson and quiz materials being too easy, as can be observed through the high mean score in all groups. This necessitates calibration of the content's difficulty level in future studies to potentially reveal clearer differences in comprehension outcomes. Alternatively, a pre-post study design might also help highlight the differences between groups.

Several limitations of the results should also be recognized, including the limited sample size and demographics of the participants, the specific lesson content domain, and brief interactions with the AI virtual instructors. Future research should address these factors to enhance the generalizability of the results. For example, future studies could expand the participant profiles, increase the sample size, and/or extend these techniques—including LLM-based interactivity and instructor selection—to other educational fields beyond financial literacy. Additionally, conducting longitudinal studies and examining potential adverse effects along with possible mitigation strategies would also be beneficial.

Lastly, it is crucial to consider that these AI systems should be seen more as complementary tools for teaching than replacements for human teachers. While they can assist educators in areas often underrepresented in traditional education systems, such as financial literacy, unique and invaluable human elements of teaching, such as empathy, ethical guidance, and real-world experience sharing are still highly essential and beneficial to learners. Therefore, the goal is to harmonize AI's capabilities with human qualities, creating an enriched learning environment that prepares students for not only the technical but also the interpersonal aspects of the real world.

VII. CONCLUSION

Our development and analysis of an AI-driven financial literacy education platform featuring AI-generated virtual instructors along with two comparative studies revealed significant advantages of interactive virtual instruction, enhanced with LLMs for dynamic question-and-answer capabilities, in terms of boosting learners' motivation, engagement, perceived learning facilitation, and the agent's humanness, compared to a non-interactive approach. On the other hand, personalization via the selection of a generated virtual instructor, even with our intervention of short introductory videos, showed minimal extra benefit in learning outcomes despite learners' heightened sense of autonomy. In addition, we have also demonstrated the importance of relatedness as an underlying predictor for several learning outcomes including learners' motivation, positive emotion, and agent credibility. These results highlight the potential for incorporating interactivity into virtual instruction as a way to enrich remote learning experience. Though comprehension levels may not necessarily be altered, it could make learning more engaging and emotionally fulfilling for learners. With the ongoing evolution of online learning, deeper explorations into optimizing interactive virtual instructors could lead to even more substantial improvements in learner engagement and motivation.

VIII. ACKNOWLEDGMENT

Statistical support was provided by data science specialist Josh Cetron, at the Institute for Quantitative Social Science, Harvard University.

REFERENCES

[1] N. L. Schroeder, O. O. Adesope, and R. B. Gilbert, "How effective are pedagogical agents for learning? a meta-analytic review," *Journal of Educational Computing Research*, vol. 49, no. 1, pp. 1–39, 2013.

[2] Y. Shiban, I. Schelhorn, V. Jobst, A. Hörnlein, F. Puppe, P. Pauli, and A. Mühlberger, "The appearance effect: Influences of virtual agent features on performance and motivation," *Computers in Human Behavior*, vol. 49, pp. 5–11, 2015.

[3] P. Pataranutaporn, V. Danry, L. Blanchard, L. Thakral, N. Ohsugi, P. Maes, and M. Sra, "Living memories: Ai-generated characters as digital mementos," in *Proceedings of the 28th International Conference on Intelligent User Interfaces (IUI)*, 2023, pp. 889–901.

[4] P. Pataranutaporn, J. Leong, V. Danry, A. P. Lawson, P. Maes, and M. Sra, "Ai-generated virtual instructors based on liked or admired people can improve motivation and foster positive emotions for learning," in *IEEE Frontiers in Education Conference (FIE)*. IEEE, 2022, pp. 1–9.

[5] P. Pataranutaporn, V. Danry, J. Leong, P. Punpongsanon, D. Novy, P. Maes, and M. Sra, "Ai-generated characters for supporting personalized learning and well-being," *Nature Machine Intelligence*, vol. 3, no. 12, pp. 1013–1022, 2021.

[6] A. P. Lawson, R. E. Mayer, N. Adamo-Villani, B. Benes, X. Lei, and J. Cheng, "Recognizing the emotional state of human and virtual instructors," *Computers in Human Behavior*, vol. 114, p. 106554, 2021.

[7] W. Li, F. Wang, R. E. Mayer, and T. Liu, "Animated pedagogical agents enhance learning outcomes and brain activity during learning," *Journal of Computer Assisted Learning*, 2021.

[8] H. Tomášková, H. Mohelská, and Z. Němcová, "Issues of financial literacy education," *Procedia-Social and Behavioral Sciences*, vol. 28, pp. 365–369, 2011.

[9] L. Klapper, A. Lusardi, and P. Van Oudheusden, "Financial literacy around the world," *World Bank. Washington DC: World Bank*, vol. 2, pp. 218–237, 2015.

[10] H. Alamri, V. Lowell, W. Watson, and S. L. Watson, "Using personalized learning as an instructional approach to motivate learners in online higher education: Learner self-determination and intrinsic motivation," *Journal of Research on Technology in Education*, vol. 52, no. 3, pp. 322–352, 2020.

[11] M. L. Bernacki, M. J. Greene, and N. G. Lobczowski, "A systematic review of research on personalized learning: Personalized by whom, to what, how, and for what purpose (s)?" *Educational Psychology Review*, vol. 33, no. 4, pp. 1675–1715, 2021.

[12] R. Reber, E. A. Canning, and J. M. Harackiewicz, "Personalized education to increase interest," *Current directions in psychological science*, vol. 27, no. 6, pp. 449–454, 2018.

[13] C. Walkington and M. L. Bernacki, "Personalization of instruction: Design dimensions and implications for cognition," *The Journal of Experimental Education*, vol. 86, no. 1, pp. 50–68, 2018.

[14] K. A. Renninger and S. E. Hidi, *The power of interest for motivation and engagement*. Routledge, 2015.

[15] J. Leong, P. Pataranutaporn, V. Danry, F. Perteneder, Y. Mao, and P. Maes, "Putting things into context: Generative ai-enabled context personalization for vocabulary learning improves learning motivation," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–15.

[16] S. Hidi and K. A. Renninger, "The four-phase model of interest development," *Educational psychologist*, vol. 41, no. 2, pp. 111–127, 2006.

[17] R. M. Ryan and E. L. Deci, *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford publications, 2017.

[18] E. L. Deci and R. M. Ryan, "Self-determination theory: A macrotheory of human motivation, development, and health." *Canadian psychology/Psychologie canadienne*, vol. 49, no. 3, p. 182, 2008.

[19] D. I. Cordova and M. R. Lepper, "Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice." *Journal of educational psychology*, vol. 88, no. 4, p. 715, 1996.

[20] G. Schraw, T. Flowerday, and S. Lehman, "Increasing situational interest in the classroom," *Educational Psychology Review*, vol. 13, pp. 211–224, 2001.

[21] S. Høgheim and R. Reber, "Supporting interest of middle school students in mathematics through context personalization and example choice," *Contemporary Educational Psychology*, vol. 42, pp. 17–25, 2015.

[22] ——, "Eliciting mathematics interest: New directions for context personalization and example choice," *The Journal of Experimental Education*, vol. 85, no. 4, pp. 597–613, 2017.

[23] H.-Y. Ku, C. A. Harter, P.-L. Liu, L. Thompson, and Y.-C. Cheng, "The effects of individually personalized computer-based instructional program on solving mathematics problems," *Computers in human behavior*, vol. 23, no. 3, pp. 1195–1210, 2007.

[24] C. L. López and H. J. Sullivan, "Effect of personalization of instructional context on the achievement and attitudes of hispanic students," *Educational Technology Research and Development*, vol. 40, no. 4, pp. 5–14, 1992.

[25] U.S. Department of Education, "Reimagining the role of technology in education: 2017 national education technology plan update," 2017. [Online]. Available: https://tech.ed.gov/files/2017/01/NETP17.pdf

[26] M. O. Riedl and V. Bulitko, "Interactive narrative: An intelligent systems approach," *AI Magazine*, vol. 33, no. 3, pp. 67–77, 2012.

[27] P. Ammanabrolu and M. O. Riedl, "Situated language learning via interactive narratives," 2021. [Online]. Available: http://arxiv.org/abs/2103.09977

[28] S. Xu, G. Chen, Y.-X. Guo, J. Yang, C. Li, Z. Zang, Y. Zhang, X. Tong, and B. Guo, "Vasa-1: Lifelike audio-driven talking faces generated in real time," *arXiv preprint arXiv:2404.10667*, 2024.

[29] M. Hasan, C. Ozel, S. Potter, and E. Hoque, "Sapien: affective virtual agents powered by large language models," in *2023 11th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*. IEEE, 2023, pp. 1–3.

[30] H. Wan, J. Zhang, A. A. Suria, B. Yao, D. Wang, Y. Coady, and M. Prpa, "Building llm-based ai agents in social virtual reality," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, 2024, pp. 1–7.

[31] A. Shoa, R. Oliva, M. Slater, and D. Friedman, "Sushi with einstein: Enhancing hybrid live events with llm-based virtual humans," in *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*, 2023, pp. 1–6.

[32] T. Ueno, Y. Sawa, Y. Kim, J. Urakami, H. Oura, and K. Seaborn, "Trust in human-ai interaction: Scoping out models, measures, and methods," in *CHI Conference on Human Factors in Computing Systems Extended Abstracts*, 2022, pp. 1–7.

[33] D. Wang, E. Churchill, P. Maes, X. Fan, B. Shneiderman, Y. Shi, and Q. Wang, "From human-human collaboration to human-ai collaboration: Designing ai systems that can work together with people," in *Extended abstracts of the 2020 CHI conference on human factors in computing systems*, 2020, pp. 1–6.

[34] Q. Yang, A. Steinfeld, C. Rosé, and J. Zimmerman, "Re-examining whether, why, and how human-ai interaction is uniquely difficult to design," in *Proceedings of the 2020 chi conference on human factors in computing systems*, 2020, pp. 1–13.

[35] S. Amershi, D. Weld, M. Vorvoreanu, A. Fourney, B. Nushi, P. Collisson, J. Suh, S. Iqbal, P. N. Bennett, K. Inkpen *et al.*, "Guidelines for human-ai interaction," in *Proceedings of the 2019 chi conference on human factors in computing systems*, 2019, pp. 1–13.

[36] M. Lee, P. Liang, and Q. Yang, "Coauthor: Designing a human-ai collaborative writing dataset for exploring language model capabilities," in *Proceedings of the 2022 CHI conference on human factors in computing systems*, 2022, pp. 1–19.

[37] A. Siarohin, S. Lathuilière, S. Tulyakov, E. Ricci, and N. Sebe, "First order motion model for image animation," in *Conference on Neural Information Processing Systems (NeurIPS)*, December 2019.

[38] K. R. Prajwal, R. Mukhopadhyay, V. P. Namboodiri, and C. Jawahar, "A lip sync expert is all you need for speech to lip generation in the wild," in *Proceedings of the 28th ACM International Conference on Multimedia*, ser. MM '20. New York, NY, USA: Association for Computing Machinery, 2020, p. 484–492. [Online]. Available: https://doi.org/10.1145/3394171.3413532

[39] T. H. Eker, *Secrets of the Millionaire Mind*. Harper Business, 2005.

[40] N. L. Schroeder, W. L. Romine, and S. D. Craig, "Measuring pedagogical agent persona and the influence of agent persona on learning," *Computers & Education*, vol. 109, pp. 176–186, 2017.

[41] A. Baylor and J. Ryu, "The api (agent persona instrument) for assessing pedagogical agent persona," in *EdMedia+ innovate learning*. Association for the Advancement of Computing in Education (AACE), 2003, pp. 448–451.

[42] Y. Benjamini and Y. Hochberg, "Controlling the false discovery rate: a practical and powerful approach to multiple testing," *Journal of the Royal statistical society: series B (Methodological)*, vol. 57, no. 1, pp. 289–300, 1995.

[43] R. M. Ryan, "Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory." *Journal of personality and social psychology*, vol. 43, no. 3, p. 450, 1982.