UNLV UNIVERSITY LIBRARIES

Journal of Research in Technical Careers

Volume 8 | Issue 1 Article 1

5-6-2024

Using ChatGPT with Novice Arduino Programmers: Effects on Performance, Interest, Self-Efficacy, and Programming Ability

Donald M. Johnson
University of Arkansas

Will Doss
University of Arkansas

Christopher M. Estepp
University of Arkansas

Recommended Citation
Johnson, D. M., Doss, W., & Estepp, C. M. (2024). Using ChatGPT with Novice Arduino Programmers: Effects on Performance, Interest, Self-Efficacy, and Programming Ability. *Journal of Research in Technical Careers, 8* (1). https://doi.org/10.9741/2578-2118.1152

Using ChatGPT with Novice Arduino Programmers: Effects on Performance, Interest, Self-Efficacy, and Programming Ability

Cover Page Footnote

This work was supported, at least in part, by the USDA National Institute of Food and Agriculture, Hatch project 1024473 and the University of Arkansas, Division of Agriculture.

Journal of Research in Technical Careers
May 2024, Vol. 8, No. 1

Using ChatGPT with Novice Arduino Programmers: Effects on Performance, Interest, Self-Efficacy, and Programming Ability

Donald M. Johnson
University of Arkansas
Will Doss
University of Arkansas
Christopher M. Estepp
University of Arkansas

A posttest-only control group experimental design compared novice Arduino programmers who developed their own programs (self-programming group, *n* =17) with novice Arduino programmers who used ChatGPT 3.5 to write their programs (ChatGPT-programming group, *n* = 16) on the dependent variables of programming scores, interest in Arduino programming, Arduino programming self-efficacy, Arduino programming posttest scores, and types of programming errors. Students were undergraduates in an introductory agricultural systems technology course in Fall 2023. The results indicated no significant ($p \leq .10$) differences between groups for programming rubric scores ($p = .50$) or interest in Arduino programming ($p = .50$). There were significant differences for Arduino programming self-efficacy, ($p = .03$, Cohen's $d = 0.75$) and Arduino posttest scores, ($p = .03$, Cohen's $d = 0.76$); students in the self-programming group scored significantly higher on both measures. Analysis of students' errors indicated the ChatGPT group made significantly ($p < .01$) more program punctuation errors. These results indicated novice students writing their own programs developed greater Arduino programming self-efficacy and programming ability than novice students using ChatGPT. Nevertheless, ChatGPT may still play an important role in assisting novices to write microcontroller programs.

Keywords: artificial intelligence, Arduino, ChatGPT, microcontrollers, novices, teaching and learning

Introduction

ChatGPT (OpenAI, 2022), the first widely available generative artificial intelligence (AI) technology, was introduced in November 2022 and had over 100 million registered users within two months (Ebert & Louridas, 2023). Generative AI is a term used to describe "machine learning solutions trained on massive amounts of data in order to produce output based on user prompts" (Saetra, 2023, para. 2). Generative AI has been predicted to be a disruptive technology, potentially revolutionizing education, the workplace, and careers (Chow et al., 2023; Saetra, 2023).

The use of generative AI to assist students with their assignments has become a growing concern for those in education with implications for teaching and learning, ethics, and workforce preparedness (Alasadi & Baiz, 2023; Baidoo-Anu & Ansah, 2023;

Chiu, 2023; Su & Yang, 2023). According to Rasul et al. (2023), ChatGPT (OpenAI, 2022), one of several generative AI platforms, has potential benefits for teaching and learning including the ability to facilitate adaptative learning experiences; provide individualized feedback; support for research, writing, and data analytics; provide automated administrative support; and develop innovative assessment activities. On the other hand, the use of ChatGPT in education can create challenges in the areas of ethics, equity, academic integrity, the potential for generating biased or falsified information, increased difficulty in evaluating skill sets of graduates, and assessing student learning outcomes (Rasul et al., 2023). While many benefits and challenges have been identified for using this relatively new technology in educational settings, much is still unknown about how using generative AI affects student learning. Therefore, several authors and educators have suggested further research is needed in this area (Chiu, 2023; Sheehan, 2023; Su & Yang, 2023).

While the use of generative AI in education is still relatively novel, research has emerged investigating the effects of this technology on student learning. Studies have shown that using ChatGPT had positive impacts on student learning (Hakiki, 2023; Li, 2023) including improved self-efficacy, attitudes, intrinsic motivation, and creative thinking. In the field of computer education, research revealed that ChatGPT can provide an adaptive learning experience for students to enhance their learning resulting in improved performance, self-efficacy, and motivation in the context of computer programming (Yilmaz & Yilmaz, 2023).

Computer programming has not extensively been taught in agricultural education; however, the inclusion of microcontrollers as components of agricultural equipment and systems requiring basic programming is becoming more common (Garling, 2013). The future workforce of the agricultural industry is expected to possess a basic knowledge of programming related to microcontrollers (Titoskaya et al., 2019). A microcontroller is a small, integrated circuit device consisting of a microprocessor, memory, and peripherals used for receiving inputs and controlling other parts of an electronic or mechanical system (Keim, 2019). Microcontrollers have an increasingly wide range of agricultural uses including robotics and drone applications in precision agriculture, greenhouse climate and irrigation controls, tractors, and variable rate applicators (Goering et al., 2003; Jude et al., 2022; Kurkute et al., 2018; Liu, 2022; Negrete, 2023, Schumann, 2010).

The application of microcontrollers in agriculture is an emerging topic that can be taught at both the high school and college levels (Global Teach Ag Network, 2024; Johnson et al., 2022). At the high school level, basic DC electrical system concepts are evaluated with programmable controllers at the National FFA Agricultural Technology and Mechanical Systems Career Development Event (National FFA Organization, 2023). A common tool used to teach both novice and advanced students how to use and program microcontrollers is the Arduino UNO (Al-Abad, 2017; Herger & Bodarky, 2015). This technology has been adopted by educators because of its potential for positive educational impacts (Lee, 2020). Studies have shown that students using Arduinos reported positive attitudes toward learning about and programming microcontrollers (Arslan & Tanel, 2021; Johnson et al., 2022; Johnson et al., 2023). Results regarding students' confidence, however, have been mixed, as Johnson et al. (2022, 2023) and

Yilmaz and Yilmaz (2023) found increased programming self-efficacy among students, but no significant increase in self-efficacy was found by Arslan and Tanel (2021). While Arduino software is considered relatively user-friendly, novice users may encounter difficulties due to their unfamiliarity with computer programming (Thomas et al., 2011).

According to the literature, the use of ChatGPT has the potential to improve students' performance, self-efficacy, and motivation in computer programming (Yilmaz & Yilmaz, 2023). In the specific context of agricultural education, would using ChatGPT influence student performance when using Arduinos to teach microcontroller programming? How might using it affect student interest and self-efficacy in the subject? The gap in research connecting these two topics and the lack of generative AI research within the context of agricultural education necessitates research comparing students who use ChatGPT to help them write an Arduino program to students who write their own program without assistance from AI. Results from this study can help determine the feasibility of using generative AI to enhance interdisciplinary teaching within agricultural education.

Theoretical Framework

To determine the effects ChatGPT could have on performance, interest, and self-efficacy, we must examine how experiences theoretically impact these variables. The intersection of social cognitive theory (Bandura, 1986), self-efficacy theory (Bandura, 1977), and Roberts' (2006) model of experiential learning served as the theoretical frameworks for this study and provided insight into how experiences impact learning. Social cognitive theory seeks to explain cognitive learning through the reciprocal interactions of personal, behavioral, and environmental factors (Bandura, 1986). Personal factors can include characteristics such as self-efficacy, values, and outcome expectations. Behavioral factors have been characterized as choice of activities, effort, persistence, and achievement. Lastly, environmental factors include feedback, instruction, opportunities for self-evaluation, and rewards (Schunk & DiBenedetto, 2020).

Self-efficacy has been defined as a person's confidence in their ability to perform a specific task or behavior (Bandura, 1977). Because self-efficacy is contextual (Smith et al., 2006), individuals with high self-efficacy would be confident in their ability to complete a specific task while individuals with low self-efficacy would be less confident in the same task. Self-efficacy theory suggests mastery, vicarious, and social persuasion experiences each influence a person's self-efficacy toward a task or behavior. Mastery experiences occur when an individual successfully accomplishes a behavior or task and tend to have the greatest influence on an individual's self-efficacy. Accordingly, Smith et al. (2006) suggested that repeated failure of a task can have negative impacts on task-specific self-efficacy. Vicarious experiences occur when an individual witnesses someone similar to themselves complete a behavior or task successfully (Bandura, 1977). Social persuasion experiences occur when another person, such as a teacher, expresses confidence in the individual's ability to successfully complete a behavior or task. Social persuasion experiences tend to have the least impact on self-efficacy (Bandura, 1977).

Experiential learning theory also lends insight into how experiences affect the learning process. According to Roberts (2006), the process of experiential learning is "cyclical in nature and requires an initial focus of the learner, followed by interaction with the phenomenon being studied, reflecting on the experience, developing generalizations, and then testing those generalizations" (p. 27). Roberts posited that learning begins with an experience, which must then be reflected upon in order for students to make generalizations. Learners then use their new knowledge in subsequent experiences leading to further experimentation in an on-going pattern.

This study applied social cognitive, self-efficacy, and experiential learning theories to test the learning impacts of using ChatGPT in the context of a college level agricultural systems technology course where students participated in a three-day lesson on microcontrollers and Arduino programming. Half the students used ChatGPT to program an Arduino UNO microcontroller to operate light-emitting diodes (LEDs) on a breadboard, while the rest programmed Arduino UNO to perform the same task without the use of ChatGPT. The programming activity served as a mastery experience where all students in the study were allowed to program their Arduino until it executed the correct blinking LED sequence on their breadboard. A breadboard operating with the correct blinking LED sequence constituted a successful mastery experience within Bandura's (1977) self-efficacy theory, theoretically improving students' programming self-efficacy.

Since the programming experience allowed students to continually make corrections to their program until it executed the correct LED sequence, the mastery experience included a reflective experiential learning component. Students who wrote their own programs could interact with the programming software and the Arduino microcontroller to determine if they could obtain the desired results on their breadboard. They could reflect on the results, develop a generalization, and try again by testing their generalization. This process could be continued until success was achieved, as recommended by Robert's (2006) experiential learning model. Students who used ChatGPT to assist them in writing their program had the opportunity to alter their prompt in ChatGPT to give them different results and allow them to keep trying for a successful experience. The interaction with ChatGPT served as a social component (environmental factor) within Bandura's (1986) social cognitive theory, influencing persistence (behavioral factor) to create a successful program, and ultimately impacting self-efficacy (personal factor).

Purpose and Objectives. The purpose of this study was to compare novice Arduino programmers who wrote their own programs (self-programming group) with those who used ChatGPT version 3.5 to write their programs (ChatGPT programming group) on (a) programming task scores, (b) interest in Arduino programming, (c) Arduino programming self-efficacy, (d) Arduino programming ability, and (e) programming errors. The specific objectives were to:

1. Determine if there was a significant ($p \leq .10$) difference in laboratory programming task rubric scores between students who wrote their own Arduino programs (self-programming group) and students who used ChatGPT (ChatGPT-programming group).
2. Determine if there was a significant ($p \leq .10$) difference in (a) interest in Arduino programming, (b) Arduino programming self-efficacy, or (c) Arduino programming posttest scores between students who wrote their own Arduino

programs (self-programming group) and those who used ChatGPT (ChatGPT-programming group).

3. Determine if there was a significant ($p \leq .10$) difference between groups (self-programming or ChatGPT-programming) on types of errors made on the Arduino programming posttest.

Methods

The accessible sample for this study consisted of students (n = 44) enrolled in one introductory agricultural systems technology course at the University of Arkansas during the Fall 2023 semester. After IRB approval, 43 students consented to participate in the study. These students were randomly assigned to the self-programming (n = 21) and ChatGPT-programming (n = 22) groups using the RANDBETWEEN function in Excel. After removing students who did not complete all research activities (f = 6) and those who reported previous Arduino programming experience (f = 4), data from 33 students were used in analysis, with approximately equal numbers in the self-programming (n = 17) and ChatGPT-programming (n = 16) groups. The self-programming group had fewer females (f = 3, 17.7%) compared to the ChatGPT-programming group (f = 6, 37.5%). A slight majority of students in both groups were freshmen or sophomores (self-programming group = 52.9%; ChatGPT group = 56.3%).

A potential limitation of experimental research in college classrooms is small samples and the resultant lack of statistical power (McGrath, 2016). Spatz (2019) defined statistical power as the probability of rejecting the null hypothesis when it is false in the population, and Cohen (1998) recommended a minimum statistical power of .80. One method of increasing statistical power, often recommended for small sample exploratory studies such as this, is to increase the alpha level (Baguley, 2004). With 33 subjects and an alpha level of .10, our statistical power at the large effect was .73, .86, and .71 for the Mann-Whitney U tests, the overall MANOVA, and the *post-hoc* Bonferroni t-tests, respectively. Thus, we recognize low statistical power as a potential limitation of our study; readers should consider this limitation in interpreting the results.

Research Design. This study employed a posttest-only control group experimental design as described by Campbell and Stanley (1963). According to Campbell and Stanley this design controls all threats to internal validity.

Experimental Procedures. During the 12th week of the fall 2023 semester, students were randomly assigned to two groups (1 and 2) and membership in each group was shared with students via class announcements and two email notifications. However, students were not informed of the specific tasks or conditions for either group. Student access to free ChatGPT 3.5 accounts was also confirmed prior to the study.

The study was conducted during the Monday, Wednesday, and Friday class meetings (50-minutes each) during the 13th week of the Fall 2023 semester. On Monday, pairs of students were provided with a package containing an Arduino UNO microcontroller and breadboard, one 240-ohm resistor, one light emitting diode (LED), pin connector wires, and a paper mock-up of the Arduino programming environment. Students then participated in an illustrated lecture introducing Arduino UNO microcontrollers, their uses in agriculture, simple resistor and LED circuits, and basic Arduino programming. The lecture incorporated the same four hands-on practice tasks described by Johnson et al. (2022; 2023): (a) point to the primary components of the Arduino UNO, (b) identify resistors and LEDs and identify the anode (+) and cathode (-) terminals of the LED, (c) breadboard a simple resistor and LED circuit between a specific digital pin and a ground pin on the Arduino UNO, and (d) write an Arduino program (in pencil on the paper mock-up of the Arduino programming environment) to cause the LED to blink repeatedly with a 1-second delay. The programming component of the lecture emphasized the two primary sections of an Arduino program (void setup and void loop), and the three basic Arduino statements (pinMode, digitalWrite, and delay) and syntax (wording, capitalization, and punctuation) necessary to accomplish the hands-on practice task. The lecture concluded with a brief demonstration of how ChatGPT could be used to write the Arduino program. Johnson et al. (2022, 2023) found that successful completion of simple hands-on practice tasks provided students with positive mastery and vicarious experiences and increased self-efficacy.

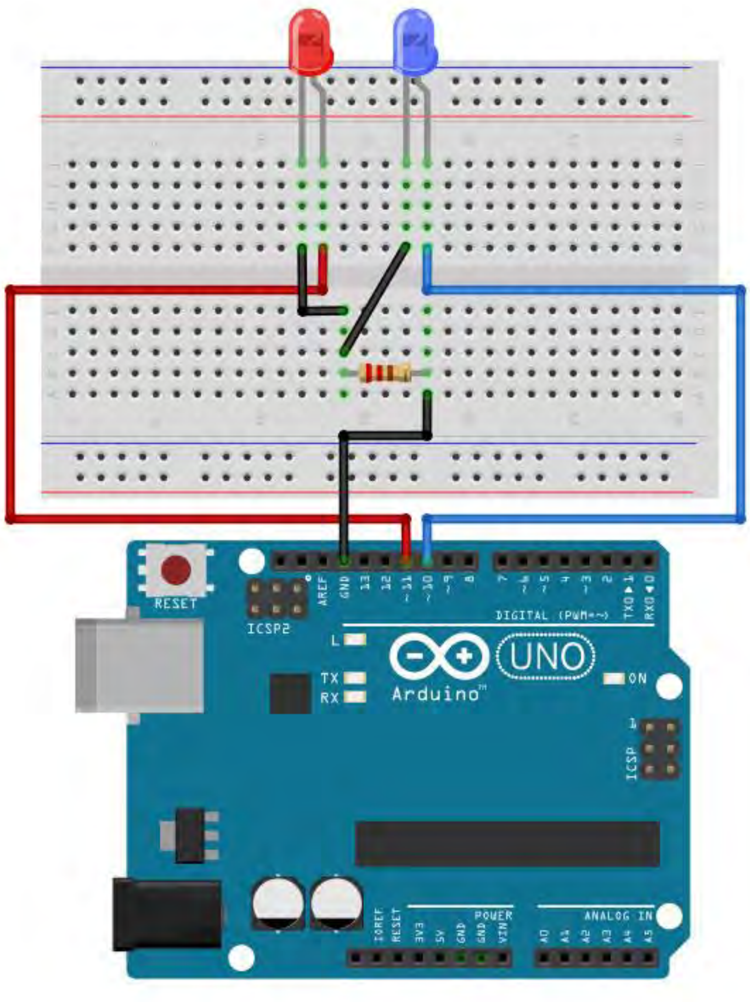
On Wednesday students reported to a college computer laboratory to complete the Arduino programming task. Each student was provided with an Arduino UNO connected to a desktop computer running the Arduino programming environment, an identical pre-breadboarded circuit (Figure 1), and a single-page reference sheet showing a pictorial drawing and the program developed in class on Monday. A slide was projected showing the two groups (1 and 2) and the students in each group. Students were instructed to open the online course management system and then to open the assignment for their group (1 or 2). The programming assignment was the same for both groups and required students to develop Arduino programs that would cause the LEDs to blink in the following sequence. The program could be written using the three basic Arduino statements (pinMode, digitalWrite, and delay) introduced in the lecture.

- Blue LED turns “ON” for 1.0 second
- Blue LED turns “OFF” for 1.0 second
- Blue LED turns “ON” for 1.0 second
- Blue LED turns “OFF” for 1.0 second
- Red LED turns "ON” for 1.0 second
- Red LED turns “OFF” for 1.0 second

Repeat the sequence

Figure 1

Arduino UNO and Breadboarded Circuit for Laboratory Activity



Students assigned to Group 1 (self-programming) were instructed to use their knowledge of Arduino programming to write a program to cause the LEDs to blink in the indicated sequence. Students assigned to Group 2 (ChatGPT-programming) were instructed to query ChatGPT to write a program to accomplish the same task by adding to the stem provided: *"Write the most simple and basic Arduino program so that an Arduino UNO will [in your own words complete the prompt so that ChatGPT will write a program that causes the LEDs to blink as described]*." The ChatGPT group was instructed to formulate their own queries, not to simply copy the desired sequence into the ChatGPT message window; students were required to copy their final ChatGPT query into the online assignment form. Both groups were instructed to paste their final Arduino programs into the online assignment form for grading. Students were not informed of the instructional treatment being received by the other group and appeared to be unaware of what process the other group was using to write their programs.

On Friday students received their graded rubric for their Arduino programs and were debriefed on the laboratory activity. A PowerPoint slide showing a pictorial drawing of the Arduino laboratory circuit and a correctly written Arduino program was displayed and discussed. The debriefing concluded with a brief (10 minute) mini lesson on applications of microcontrollers in agriculture and on additional coursework offered for interested students. After the debriefing, students completed the survey instrument measuring Arduino interest and Arduino programming self-efficacy, then completed the Arduino programming posttest.

Instrumentation. Data were collected using a survey instrument, a posttest, and two scoring rubrics. The survey instrument, adapted from Johnson et al. (2022), contained three sections. The first section assessed students' interest in learning about Arduino programming using a 13-item summated Likert scale (1 = *strongly disagree* and 5 = *strongly agree*) with a coefficient alpha reliability of .91. Johnson et al. based this section on an original instrument developed by Gable and Roberts (1983). The second section measured Arduino programming self-efficacy using an 8-item summated Likert scale (1 = *very unconfident* and 5 = *very confident*) with a coefficient alpha reliability of .71. This section was based on an original instrument developed by Kittur (2020). The third section contained three items about academic classification, gender identity, and previous Arduino programming experience. Johnson et al. reported the original instrument was evaluated by three experts in engineering education who were informed of the research objectives and subjects and judged the instrument to have face and content validity.

The Arduino programming posttest was similar to the laboratory activity. A drawing presented an Arduino UNO and a breadboard with one orange and one red LED circuit connected. The desired operation of each LED was described, and students were provided with a paper mock-up of the Arduino programming environment. Students were instructed to write an Arduino program to achieve the desired circuit operation using correct commands and syntax as if they were typing directly into the Arduino programming environment. Students were not allowed to use any reference materials in completing the posttest.

The course instructor used two rubrics, based on those used by Johnson et al. (2022), to evaluate the student Arduino programs developed during the hands-on laboratory activity and the programming posttest. Both rubrics contained dichotomously scored (incorrect = 0 and correct = 1) items; the laboratory activity rubric consisted of 17 items and the programming posttest rubric contained 52 items. Scores for both rubrics were converted to a percentage correct basis.

Data Analysis. Data were analyzed using SAS 9.4. For objectives one and three, nonparametric Mann-Whitney *U* tests were used to determine if there were significant differences between groups (self-programming vs. ChatGPT-programming) for scores on the laboratory programming task or types of programming errors. For objective two, a one-way multivariate analysis of variance (MANOVA) was used to determine if significant ($p \leq .10$) differences existed in group means for the posttest measures of (a) interest in learning about Arduino, (b) Arduino programming self-efficacy, or (c) Arduino programming test scores. Bonferroni *t*-tests were used *post hoc* to identify dependent variables on which the groups differed significantly while maintaining the overall experiment-wise error rate at the .10 level. The *alpha* level for all statistical tests was set at .10 *a priori*.

Nonparametric Mann-Whitney *U* tests were used for objectives one and three because the data for these objectives did not meet the assumption of homogeneity of group variances required for parametric tests (Field & Miles, 2010). Before MANOVA testing (objective two), the data were examined to identify outliers and tested for violation of the assumption of homogeneity of covariance matrices. Two outliers (both in the self-programming group with low scores on the posttest) were identified. Following suggestions by Field and Miles (2010), the MANOVA analysis was conducted both with and without the outliers included. These analyses resulted in consistent results for both the MANOVAs and the *post hoc* Bonferroni *t*-tests; therefore, the two outliers were retained and reported in the analysis. The Box's *M* test results, $\chi^2(6) = 2.15$, $p = .91$, indicated the MANOVA assumption of homogeneity of variances was met.

Results

Objective One. Objective one was to determine if there was a significant ($p \leq .10$) difference in laboratory programming task rubric scores between students who wrote their own Arduino programs (self-programming group) and students who used ChatGPT (ChatGPT-programming group). The mean rubric scores for the Arduino laboratory programming activity were 94.8% (SD = 6.5%) for the self-programming group and 90.4% (SD = 17.2) for the ChatGPT-programming group. The sample mean for the self-programming group was 4.9% higher than the sample mean for the ChatGPT-programming group. However, a higher percentage (68.8%) of the ChatGPT-programming group made a perfect score on the programming activity compared to the self-programming group (41.2%). The results of the Mann-Whitney *U* test (U = 117.00, p = .50) indicated no significant difference between groups for laboratory programming rubric scores.

All students in the self-programming group wrote programs using only the three statements taught and practiced in the lecture. For the ChatGPT-programming group, despite the prompt to write the "most simple and basic Arduino program," all programs included one additional statement, the *const int* statement, to assign descriptive names (such as LED1 and LED2) to the digital pins controlling the LEDs. The most common errors made by the self-programming group were relatively minor and included failure to include comments to document the program (f = 5) and omitting one or more delay statements (f = 4) from the program. The most common errors for the ChatGPT-programming group were omitted or incorrect delay statements (f = 4) and, more seriously, omitted sections of code required to control one of the LEDs (f = 4). Evaluation of the ChatGPT queries indicated these programming errors resulted primarily from incomplete or incorrectly worded ChatGPT queries. For example, one student query, "Write the most simple and basic Arduino program so that an Arduino UNO will activate a blue LED for 2 seconds, turn off the LED for 2 seconds, and repeat the pattern," omitted mention of the red LED and specified the wrong delay period, resulting in a rubric score of 47.1%.

Considering the small sample used in this study, the effect size for laboratory programming activity rubric scores was calculated even though the difference between groups for this variable was not statistically significant ($p > .10$). The resulting effect size of 0.13 indicated that if larger sample studies find a significant difference in the population for Arduino programming rubric scores between groups for this task, the magnitude of this difference is likely to be negligible to small (Cohen, 1988).

Objective Two. The second objective was to determine if there was a significant ($p \leq .10$) difference in (a) interest in Arduino programming, (b) Arduino programming self-efficacy, and (c) Arduino programming posttest scores between students in the self-programming group and students in the ChatGPT-programming group. Observed means for interest, self-efficacy, and posttest scores were higher for the self-programming group than for the ChatGPT-programming group (Table 1). Scores for the self-programming group were 3.8% higher for interest, 11.7% higher for self-efficacy, and 20.5% higher for posttest scores.

Table 1

Descriptive Statistics for Interest, Self-efficacy and Test Scores, by Group

		Dependent Variable					
		Interest [a]		Self-efficacy [b]		Posttest [c]	
Group	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Self-programming	17	3.86	0.58	3.63	0.53	86.9%	20.6%
ChatGPT-programming	16	3.72	0.61	3.25	0.50	72.1%	15.2%

[a] Measured on a summated 13-item scale where 1 = *strongly disagree* and 5 = *strongly agree*. [b] Measured on a summated 8-item scale where 1 = *very unconfident* and 5 = *very confident*. [c] Percent correct on a 52-item scoring rubric.

The results of a one-way MANOVA indicated a significant multivariate effect for group on one or more dependent variables, $F(3, 29) = 4.21$, $p = .01$. *Post hoc* Bonferroni *t*-tests indicated significant differences between groups for self-efficacy, $t(31) = 2.14$, $p = .03$), and posttest scores, $t(31) = 2.17$, $p = .03$. There was no significant difference between groups for interest, $t(31) = 0.69$, $p = .50$).

Cohen's *d* effect sizes (Cohen, 1988) were calculated to quantify the magnitude of the group differences for the two significant dependent variables self-efficacy and posttest scores. With Cohen's *d*s of 0.75 and 0.76, group membership had a medium effect on both self-efficacy and posttest scores, respectively, for novice Arduino programmers. According to Cohen (1988), a medium effect size represents "an effect likely to be visible to the naked eye of the careful observer" (p. 156) and is the typical effect observed in most fields of social science research.

Again, considering the small sample size used in this study, the Cohen's *d* for interest was calculated even though the difference between groups for this variable was not statistically significant ($p > .10$). The resulting Cohen's *d* of 0.24 indicated that if larger sample studies find a significant difference exists in the population for Arduino programming interest in similar studies comparing novice students who self-program and those who program using ChatGPT, the magnitude of this difference is likely to be negligible to small (Cohen, 1988).

Objective Three. The final objective was to describe errors on the Arduino programming posttest and to determine if there were significant ($p \leq .10$) differences between groups (self-programming or ChatGPT-programming) for any category of error. Table 2 shows the seven categories of programming errors evaluated by the rubric, the number of opportunities to make each error in writing the program, descriptive statistics for the number of errors made by group, and the ratio of the mean errors made comparing the ChatGPT-programming group to the self-programming group (ChatGPT: Self) for each category of error.

Table 2

Summary of Posttest Programming Errors by Group

		Group				
		Self-programming		ChatGPT-programming		
Error	Error opportunities	*M*	*SD*	*M*	*SD*	Error ratio (ChatGPT: Self)
Required statement is absent or incorrect	10	0.35	1.03	0.25	.066	0.71
Wording of required statement is incorrect	10	2.24	3.20	3.00	3.39	1.34
Incorrect digital pin specified in setup	2	0.35	0.76	0.23	0.66	0.66
Incorrect digital pin specified in loop	2	0.24	0.64	0.25	0.66	1.04
Incorrect capitalization of statements	10	1.52	2.68	2.81	3.03	1.85
Incorrect punctuation of statements	10	2.12	3.05	7.27	3.86	3.43
Incorrect delay duration specified	6	0.12	0.47	0.00	0.99	--

The mean number of errors by group was similar for six of the seven categories; however, the mean punctuation errors for the ChatGPT programming group was 3.43

times higher than for the self-programming group. The primary punctuation error was omitting the required semicolon at the end of each program statement. The results of the Mann-Whitney U test confirmed the frequency of punctuation errors committed by the ChatGPT-programming group was significantly higher than for the self-programming group, $U = 51.00$, $p < .01$. The effect size of 0.55 indicated a medium effect (Cohen, 1988) for group on punctuation errors. There were no other significant ($p \leq .10$) differences between groups for errors by category.

Conclusions, Recommendations, and Implications

Caution should be used in interpreting the results of this study due to the small sample size. However, the results do suggest several tentative conclusions and recommendations for both research and teaching practice in colleges of agriculture. Results from objective one indicated the mean rubric scores for the self-programming and the ChatGPT-programming groups were above 90% on the laboratory programming activity and there was no statistically significant difference in mean scores. Further, the ChatGPT-programming group had a higher percentage of perfect rubric scores (68.8%) than did the self-programming group (41.2%), and errors by the ChatGPT-programming group were primarily the result of incorrect or incomplete queries. This finding raises an important question for teaching Arduino (and similar) programming languages in colleges of agriculture. For agriculture students who may need to use Arduinos and similar microcontrollers only occasionally in their academic and professional careers, would it be more effective to focus on basic programming skills, or on how to write complete and correct ChatGPT (and similar AI chatbot) queries and evaluate the resultant programs? Or is some combination of instruction in basic programming and effective use of ChatGPT more warranted? This is an area that calls for further discussion and research.

Data from objective two revealed that both groups were somewhat interested in learning more about Arduino programming with no significant difference between the two groups. This indicates that either method (self-programming or ChatGPT-programming) can be used to teach Arduino programming to novices without sacrificing student interest. This may be an especially important finding for instructors teaching undergraduates who may occasionally need to use microcontrollers for specific academic and career tasks, but do not require deep expertise in microcontroller programming.

The self-programming group had a significantly higher mean score for Arduino programming self-efficacy than the ChatGPT-programming group (medium effect sizes). This seems to suggest that self-programming should be the preferred method of instruction if the desired outcome is self-efficacy. However, programming self-efficacy was measured with statements concerning students' ability to *write* Arduino programs, and according to Smith et al. (2006), self-efficacy is task specific. Thus, had we used statements concerning students' ability to *use ChatGPT* to write Arduino programs our results might have been different. This warrants further research. Studies from the literature review found the use of ChatGPT improved self-efficacy (Li, 2023; Yilmaz & Yilmaz, 2023). While our study did not assess changes in self-efficacy over time, it is plausible gains in self-efficacy resulting from ChatGPT use still may not be as strong as self-efficacy gains from self-programming; however, further experimental research is needed to test this speculation.

According to Bandura's (1977) self-efficacy theory, a mastery experience impacts self-efficacy. Students in both the self-programming and ChatGPT programming groups engaged in a mastery experience by successfully executing the correct blinking LED sequence, so why might their self-efficacy levels differ? Consistent with Social Cognitive Theory (Bandura, 1986), ChatGPT use could possibly be an environmental factor reciprocally interacting with self-efficacy. Additionally, due to students' potential lack of familiarity with ChatGPT, the addition of ChatGPT may have added complexity to the task, which was not experienced by the self-programming group. Increased difficulty of a task has been shown to affect self-efficacy (Smith et al., 2006) and is another area in which empirical testing is warranted in the context of generative AI use.

The self-programming group scored significantly higher on the Arduino programming posttest compared to the ChatGPT-programming group (medium effect size). This finding was somewhat intuitive in that students who had written a program (self-programming group) scored higher than students who had not written a program (ChatGPT-group). While Yilmaz and Yilmaz (2023) suggested ChatGPT can enhance learning resulting in improved performance over time, our study would indicate its use does not necessarily equate with better performance. Perhaps this is an area worthy of continued investigation and over longer periods of time. Examination of students' errors indicated the only significant difference between groups was punctuation errors, primarily the mistake of omitting the semicolon at the end of statements. Likely future ChatGPT-programming assignments should require students to deliberately examine and reflect on syntax and other nuances of the ChatGPT generated program. This finding also illuminates the concern of assessing student learning outcomes highlighted by Rasul et al. (2023). Accordingly, what should be the intended learning outcome: have students use ChatGPT to effectively write Arduino programs and be able to diagnose and correct errors or write Arduino programs from memory?

Overall, posttest rubric scores for the ChatGPT-programming group (72.1%) provided evidence of learning; however, it was unclear whether this learning was a product of classroom instruction or from developing the ChatGPT-written program, or a combination of the two. Further research should explore the exact source of this learning and how it can be enhanced. According to experiential learning theory (Roberts, 2006), reflection should occur for learning to take place. Therefore, intentional reflection activities (Roberts, 2006) should be incorporated into the ChatGPT programming experience to increase learning and to aid students in more productively using the results of generative AI to perform common programming tasks.

This study provided consistent and intriguing insights for both additional research and for teaching Arduino microcontroller programming to novice college of agriculture students. However, the most important questions are curricular and focus on the purpose and desired outcomes of teaching Arduino programming as artificial intelligence applications become even more prevalent and powerful. Should colleges of agriculture focus on developing students' programming skills or developing students' ability to productively and efficiently use artificial intelligence applications like ChatGPT in completing microcontroller programming tasks, or perhaps some combination of the two? Research such as the present study can inform this important discussion.

Generative AI technologies such as ChatGPT will continue to impact education and careers in all areas including agriculture (Chow et al., 2023; Saetra, 2023). This will present an ongoing challenge and opportunity for faculty members in universities and colleges of agriculture. Educators and researchers must be proactive in developing strategies to positively incorporate this potentially disruptive technology (Saetra, 2023) into the student experience so graduates will be able to productively combine their expertise with generative AI to produce outcomes superior to what either personal expertise or AI alone can accomplish.

References

Al-Abad, M. (2017). A review of embedded systems education in the Arduino age: Lessons learned and future directions. *International Journal of Engineering Pedagogy, 7*(2), 79-93. https://doi.org/10.3991/ijepv7i2.6845

Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education, 100*(8), 2965-2971. https://doi.org/10.1021/acs.jchemed.3c00323

Arslan, K., & Tanel, Z. (2021). Analyzing the effects of Arduino applications on students' opinions, attitude and self-efficacy in programming class. *Education and Information Technologies, 26*, 1143-1163. https://doi.org/10.1007/s10639-020-10290-5

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52-62. https://doi.org/10.61969/jai.1337500

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice Hall.

Baguley, T. (2004). Understanding statistical power in the context of applied research. *Applied Ergonomics, 35*(2004), 73-80. https://doi:10.1016/j.apergo.2004.01.002

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research.* Houghton Mifflin.

Chiu, T. K. F. (2023). The impact of generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2023.2253861

Chow, J. C. L., Sanders, L., & Kay, L. (2023). Impact of ChatGPT on medical chatbots as a disruptive technology. *Frontiers in Artificial Intelligence, 6*(2023). https//doi.org/10.3389/frai.2023.1166014

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Earlbaum Associates.

Ebert, C., & Louridas, P. (2023). Generative AI for software practitioners. *IEEE Software, 40*(4), 30-38.

Field, A., & Miles, J. (2010). *Discovering statistics using SAS.* SAGE.

Gable, R. K., & Roberts, A. D. (1983). An instrument to measure attitude toward school subjects. *Educational and Psychological Measurement, 43*(1), 289-293. https://doi.org/10.1177/001316448304300139

Garling, C. (2013, May 29). *Farm geeks: Learning to love the circuit board*. Modern Farmer. https://modernfarmer.com/2013/05/farm-geeks-learning-to-love-the-circuit-board/

Global Teach Ag Network (2024). *Awesome resources recommended by educators!* https://globalteachagnetwork.psu.edu/educator-recommended-resources/

Goering, C. E., Stone, M. L., Smith, D. W., & Turnquist, P. K. (2003). *Off-road vehicle engineering principles*. https://doi.org/10.13031/orvp.2013

Hakiki, M., Fadli, R., Samala, A. D., Fricticarani, A., Dayurni, P., Rahmadani, K., Astiti, A. D., & Sabir, A. (2023). Exploring the impact of using Chat-GPT on student learning outcomes in technology learning: The comprehensive experiment. *Advances in Mobile Learning Educational Research*, *3*(2), 859-872. https://doi.org/10.25082/AMLER.2023.02.013

Herger, L. M., & Bodarky, M. (2015). Engaging students with open-source technologies and Arduino. *Proceedings of the 2015 Integrated STEM Education Conference*, pp. 27-32. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7119938

Johnson, D. M., Pate, M. L., Estepp, C. M., & Wardlow, G. W. (2023). Using self-efficacy theory to design Arduino instruction for novices: A replication study. *Journal of Research in Technical Careers, 7*(1), 68-84. https://doi.org/10.9741/2578-2118.1125

Johnson, D. M., Pate, M. L., Estepp, C. M., Wardlow, G. W., & Hood, G. T. (2022). Designing Arduino instruction for novice agriculture students: Effects on interest, self-efficacy, and knowledge. *Applied Engineering in Agriculture, 38*(5), 753-761. https://doi.org/10.13031/aea.15031

Jude, P. S. V., Kathirvel, C., Kumar, S. V., & Kumar, R. M. (2022). Multi-tasking robot using microcontroller for agriculture application. *Journal of Physics: Conference Series, 2335*. https://iopscience.iop.org/article/10.1088/1742-6596/2325/1/012029/pdf

Keim, R. (2019, March 25). *What is a microcontroller? The defining characteristics and architecture of a common component.* All About Circuits. https://www.allaboutcircuits.com/technical-articles/what-is-a-microcontroller-introduction-component-characteristics-component/

Kittur, J. (2020). Measuring the programming self-efficacy of electrical and electronics engineering students. *IEEE Transactions in Engineering, 63*(3), 216-223. https://doi.org/10.1109/TE.2020.2975342

Kurkute, S. R., Deore, B. D., Kasar, P., Bhamare, M., & Sahane, M. (2018). Drones for smart agriculture: A technical report. *International Journal for Research in Applied Science & Engineering Technology, 6*(4), 341-346. https://www.researchgate.net/profile/Swapnil-Kurkute/publication/325199681_Drones_for_Smart_Agriculture_A_Technical_Report/links/5caf2eb3a6fdcc1d498c7329/Drones-for-Smart-Agriculture-A-Technical-Report.pdf

Lee, E. (2020). A meta-analysis of the effects of Arduino-based education in Korean primary and secondary schools in engineering education. *European Journal of Educational Research, 9*(4), 1503-1512. https://doi.org/10.12973/eu-jer.9.4.1503

Li, H. (2023). Effects of a ChatGPT-based flipped learning guiding approach on learners' courseware project performances and perceptions. *Australasian Journal of Educational Technology*, *39*(5), 40–58. https://doi.org/10.14742/ajet.8923

Liu, Y. (2022). Smart greenhouse monitoring and controlling based on NodeMCU. *International Journal of Advanced Computer Science and Applications, 13*(4), 597-600. https://media.proquest.com/media/hms/PFT/1/I5GeP?_s=JRsPtB3MGQNPWp6SDZ5eYzuvzAc%3D

McGrath, A. (2016). Searching for significance in the scholarship of teaching and learning and finding none: Understanding non-significant results. *Teaching and Learning Inquiry*, *4*(2), 150–155. https://doi.org/10.20343/teachlearninqu.4.2.12

National FFA Organization. (2023). *Agricultural technology and mechanical systems CDE handbook*. https://ffa.app.box.com/s/s2hy37znu69xwns788799w38kctwxqdc

Negrete, J. C. (2023). Arduino in agriculture and the teaching of mathematics. *Horticulture International Journal, 7*(3), 80-83. https://medcraveonline.com/HIJ/HIJ-07-00280.pdf

OpenAI. (2022). ChatGPT (Version 3.5) [Computer software]. https://www.openai.com/

Rasul, T., Sumesh, N., Kalendra, D., Robin, M., Santini, F. D. O., Ladeira, W. J., Sun, M., Day, I., Rather, R. A., & Heathcote, L. (2023). The role of ChatGPT in higher education: Benefits, challenges, and future research directions. *Journal of Applied Learning & Teaching, 6*(1), 41-56. https://doi.org/10.37074/jalt.2023.6.1

Roberts, T. G. (2006). A philosophical examination of experiential learning theory for agricultural educators. *Journal of Agricultural Education, 47*(1), 17-29. https://doi.org/10.5032/jae.2006.01017

Saetra, H. S. (2023). Generative AI: Here to stay, but for good? *Technology in Society, 75*. https://doi.org/10.1016/j.techsoc.2023.102372

Schumann, A. W. (2010). Precise placement and variable rate fertilizer application technologies for horticultural crops. *HortTechnology, 20*(1), 34-40. https://doi.org/10.21273/HORTTECH.20.1.34

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology, 60*. https://doi.org/10.1016/j.cedpsych.2019.101832

Sheehan, T. (2023, September 11). *Generative AI in education: Past, present, and future.* EDUCAUSE Review. https://er.educause.edu/articles/sponsored/2023/9/generative-ai-in-education-past-present-and-future

Smith, S. A., Kass, S. J., Rotunda, R. J., & Schneider, S. K. (2006). If at first you don't succeed: Effects of failure on general and task-specific self-efficacy and performance. *North American Journal of Psychology, 8*(1), 171-182.

Spatz, C. (2019). *Exploring statistics: Tales of distributions.* Outcrop Publishing.

Su, J., & Yang, W. (2023). Unlocking the power of ChatGPT: A framework for applying generative AI in education. *ECNU Review of Education, 6*(3), 355-366. https://doi.org/10.1177/20965311231168423

Thomas, M. K., Ge, X., & Greene, B. A. (2011). Fostering 21st century skill development by engaging students in authentic game design projects in a high school computer programming class. *Journal of Educational Computing Research, 44*(4), 391-408. https://doi.org/10.2190/EC.44.4.b

Titovskaya, N. V., Titovskaya, T. S., & Titovskii, S. N. (2019). The prospects of microcontroller application in the agriculture digitalization. *IOP Conference Series: Earth and Environmental Science, 315*. https://iopscience.iop.org/article/10.1088/1755-1315/315/3/032011

Yilmaz, R., & Yilmaz, F. G. K. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence, 4*. https://doi.org/10.1016/j.caeai.2023.100147

Acknowledgments

This work was supported, at least in part, by the USDA National Institute of Food and Agriculture, Hatch project 1024473 and the University of Arkansas, Division of Agriculture.