Learning Waste Management from Interactive Quizzes and Adaptive GPT-guided Feedback

Qiming Sun[1], Shih-Yi Chien[2] and I-Han Hsiao[1,*]

[1]Santa Clara University, 500 El Camino Real, Santa Clara, California, USA
[2]National Chengchi University, Taipei, Taiwan

Abstract
Our research team has designed an educational technology called Waste Genie (WG) to support waste management learning. WG provides a series of technological features to engage informal learning, including interactive waste-sorting quizzes, a waste scanner to detect waste objects and the associated waste bin labels, and virtual carbon credits to help users quantify and visualize the environmental impacts. However, waste sorting alone is an immense and complicated problem. The recycling regulations vary by the city, the county, or even by the organization. We engineer an AI agent that utilizes GPT to provide adaptive feedback to support sorting. We hypothesize that the chatbot implementation is insufficient, because the users may experience limited vocabulary or have difficulty expressing their questions and concerns. The adaptive feedback is provided based on users' waste sorting attempts. A study was conducted with 54 participants to examine the effects of environmental awareness and waste management learning. The results showed increased waste sorting accuracy and efficiency, improved sustainability awareness, and positive effects of AI-generated feedback. Overall, the study demonstrated the feasibility of integrating the off-the-shelf AI agent to enhance educational technology and effectively support waste management learning.

Keywords
Educational technology, sustainability learning, behavioral patterns

1. Introduction

In a world with escalating environmental concerns, it is important to foster sustainable waste management practices. Recently, researchers and organizations have undertaken various efforts to support sustainability through improved waste management. For instance, an assortment of Artificial Intelligence techniques have been deployed to facilitate waste sorting: deep learning algorithms to accurately classify trash images [1, 2, 3]; smart bins and sensors to help waste collection facilities sort garbage [4, 5] and various programs [6, 7] undertaken by local agencies or working groups to encourage participation in energy saving or carbon neutrality activities. In our research group, we are dedicated to researching technological solutions to support waste management learning.

With the rapid growth of the Large Language Model (LLM) across fields in the past year, there is mounting evidence to illustrate LLM-based applications and benefits [8, 9] albeit carrying risks. A common LLM application is to integrate a chatbot (i.e. ChatGPT) to exchange conversations to build the context and yield questions and answers. In our targeted field, waste management learning, the body of waste knowledge is increasingly complex. The composition of our daily wastes usually spans across categories of recycling, landfill, compost, etc., and poses challenges for us to dispose of them appropriately. Moreover, the recycling regulations vary by the city, the county, or even by the organization. It further complicates the proper practices of waste management. We hypothesize that the chatbot implementation is insufficient, because the users may experience limited vocabulary or have difficulty expressing their questions and concerns. Therefore, in this work, we engineer an AI agent that adapts to the users' actions during waste sorting practices and leverages the power of GPT to provide adaptive feedback.

In this work, we design and conduct studies to investigate the following research questions:

1. **Impact of AI-Guided Assistance**: How does the integration of an AI agent impact the overall user experience?
2. **Effectiveness in Learning**: To what extent does the GPT-powered feedback contribute to users' increased knowledge and awareness of waste management?
3. **Usability and User Engagement**: How do the users perceive the overall Waste Genie usability and to what extent do they engage in the WG platform?

2. Related Work

2.1. LLMs in Educational Technologies

Recent advances in large language models (LLMs) have opened up new possibilities for educational content generation. Several research studies have demonstrated the benefits of employing LLMs. For instance, students felt they learned and perceived less difficulty in solving math problems with the help of LLMs [8]; OpenAI Codex was used to create programming exercises and the quality of the generated questions was found both novel and sensible [9]. Similar research also showed that the reading comprehension questions generated by LLMs could surpass the quality of those written by humans [10]. LLM-based systems were also used for supporting teachers' preparation of lectures, such as generating mini-lectures [11], designing curriculum [12], generating lecture metadata [13]. Furthermore, studies also evaluated the effectiveness of LLMs like ChatGPT in providing feedback and hints for digital learning games [14] and programming assignments [15], and have compared the learning gains between ChatGPT and human tutor-generated hints for algebra problems [16]. Our work uses the language model as a virtual agent that supplies textual information as feedback to guide students through their interactive waste-sorting progress.

Educational Datamining '24 Human-Centric eXplainable AI in Education and Leveraging Large Language Models for Next-Generation Educational Technologies Workshop Joint Proceedings, July 13, 2024, Atlanta, GA
*Corresponding author.
qsun4@scu.edu (Q. Sun); sychien@nccu.edu.tw (S. Chien); ihsiao@scu.edu (I. Hsiao)
0009-0000-1609-5217 (Q. Sun); 0000-0002-1888-3951 (I. Hsiao)

2.2. Adaptive Feedback in Learning

Adaptive feedback dynamically tailors the instructional guidance to individual learners based on their progress and needs, reflecting a careful design consideration of all available information [17]. Adaptive learning systems model students' learning style, prior knowledge, goals, and preferences [18]. Research showed that students' perception of helpfulness and reflection improved when they received personalized feedback on their essays from a natural language processing pipeline [19]. The use of adaptive AI in intelligent science stations under mixed reality environments helped enhance children's STEM learning effects without sacrificing their enjoyment [20]. Community peer feedback within an educational content annotation system was also shown to improve the quality of student-authored content [21]. Additionally, research demonstrated the efficacy of adaptive learning systems compared to traditional instruction settings, resulting in improved mathematics scores for students [22].

2.3. EdTech for Waste Management

There have been drastic technological advancements in improving waste management these days. Research ranges from (a) the power of deep learning to identify wastes in medical fields [23], to sort plastics [24], to recognize recycle materials [3], and to categorize them in diverse settings [25]; (b) immersive or robotic technology to support waste management, such as visualizing the impact of wastes in Augmented Reality (AR) [26]; Smart waste bins with sensors to promote waste categorization [7]; Our team also ventured into the realm of AR and leveraged its potential to guide individuals in proper waste disposal [27, 28]; (c) social and/or interactive gaming to educate different groups of audience with, such as ROBOTE [29], PEAR [30], FoodFights [31], HotDish [32]; (d) adaptive technology to enhance waste management awareness, for instance, Social recipes recommenders to reduce food waste [33]; recycling suggestions upon purchasing at the vending machines [34].

3. Methodology

3.1. Research Platform

Waste Genie (WG) is a web-based educational technology, designed to support waste management learning through reading and practicing waste sorting [35]. It is developed using Flutter with a NoSQL database that is used to store all the user profiles, quiz data, statistics, and logs. There are many innovative features engineered in WG to support learning, such as a waste scanner for recognizing waste objects and the corresponding waste bin labels, etc. However, to focus on the investigation of GPT-powered feedback in waste sorting and learning, we set up the experimental environment in WG with only the interactive sorting quizzes and the *behind-the-scene* AI feedback agent. The homepage (Fig. 1a) presents interactive sorting quizzes and utilizes infinite linear scrolling to access all the quiz content in the system. The design emulates a popular culture for bite-size content consumption, such as social media applications (i.e. Instagram).

The objective of sorting wastes is to classify each waste object into one of four categories: Recycle, Landfill, Compost, and Hazard. Each quiz consists of three waste items (Fig. 1b). Users interact with the quizzes by dragging and dropping the category labels onto the waste objects. The corrective feedback (correct or incorrect) will be indicated by the visual elements (green check mark and red x mark). Meanwhile, the AI agent is actively engaged when any incorrectness is detected. A quiz is completed when all of the objects are labeled correctly (Fig. 1c). Upon completion, users will be notified with an estimated amount of virtual carbon credits. An educational tip section will also be compiled and made available once the sorting is accomplished. The tip is an additional layer of information where the users can consolidate their learning by reading more detailed information on how to properly process the waste items featured in the quiz. It is intended to promote deeper reflective learning in the practices. Finally, a leaderboard (Fig. 1d) of virtual carbon credits is accessible for the users to comprehend and visualize the real-world impact of the properly sorted wastes. Two primary metrics, *mileage-travel-by-car* and *tree-days equivalent* (the amount of CO2 sequestered by a mature tree per day), are employed as the quantitative estimations to represent the significance of the carbon saved.

3.2. Technology Infrastructure

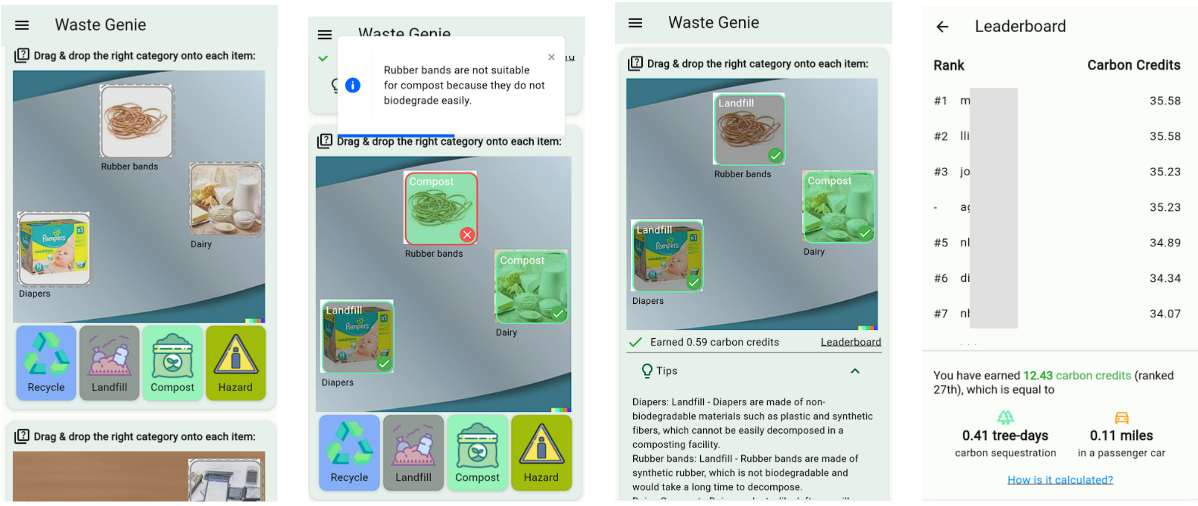
To investigate the effects of GPT-guided feedback for waste management learning in WG, we specifically engineered the AI agent to support feedback adaptation in the context. Fig. 2 demonstrates the overall architecture of our research platform - Waste Genie.

3.2.1. Waste items data set and data preparation

Interactive waste sorting quizzes are the essential learning content in WG. A quiz comprises two key elements, the waste items and their corresponding category labels. The association between the item and the label constitutes the corrective information for proper sorting (A.K.A. feedback for the users). In this study, we sourced 96 waste items from California's recycling guidelines [36] and the EPA's directions on universal waste programs [37]. These waste items' category labels serve as the "ground truth" to evaluate the correct solution to the quizzes. Most importantly, ground-truth information provides the context to the AI agent in shaping the feedback guidance, which is deliberately used to prevent the risk of any potential AI hallucination for giving erroneous information. Note that the waste items are just words without visuals.

3.2.2. Quiz generation

To supply unlimited interactive quizzes in WG, we first exploited the option of generative AI quizzes, so users can refresh the content feed to demand more content whenever they want. Many text-to-image models have demonstrated the capability of textual guidance in creating visual content, such as DALL·E 2 [38], DreamBooth [39]. For instance, a textual input for the model to generate the image can be *"one recyclable aluminum can, one compostable banana peel and one landfill candy wrapper randomly positioned in the image and no overlapping one another"*. However, to preserve the identity of the objects, the textual input usually requires detailed specifics, for instance, the name and the condition of the waste item, the information of the correct category label, and the objects' relation and position information.



(a) Waste Genie Homepage. (b) Interactive quiz with real-time AI assistance. (c) Quiz after completion. (d) Leaderboard

Figure 1: Waste Genie application

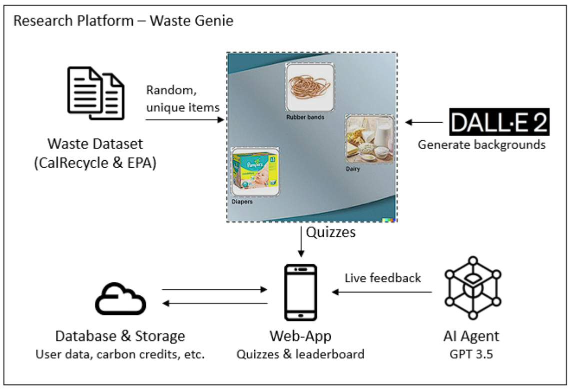


Figure 2: Architecture of Waste Genie.

In this paper, since we have already collected the ground truth labels of the waste items, we just need to assemble the ground truth information with the associated imagery without generating everything artificially from scratch.

Thus, in our study, each quiz used was generated by randomly selecting three items from our pre-collected dataset (Section 3.2.1.1). The selected waste items were firstly associated with corresponding images crawled from Google open image search. Next, The waste items were placed on a canvas with a clear background generated by DALL·E 2 [38]. The metadata of each quiz, including the waste objects (including the images), their categories, and the coordinates on the background, is stored in the database for rendering to the front-end representation in the WG. The platform is designed to be able to cater to random quizzes in real-time for practice. In this user study, we controlled the setup and pre-generated 30 quizzes in advance for each participant, with the same content displayed in the same sequence. For the reason to ensure the participants experience the same WG setup and fair comparisons in the behavioral analysis.

3.2.3. Integration of AI agent

The AI agent, powered by GPT-3.5 [1], plays a pivotal role in guiding the users during waste sorting. When a user makes a mistake, the AI agent receives the context of the current quiz, including the item that causes the user's confusion, the user's erroneous choice, and the correct answer. It is then instructed to generate a concise explanation to help the user move towards the correct answer. For instance, when a user misclassifies a recyclable paper bag as "landfill", the AI agent will be instructed that the user is confused about labeling a paper bag to recycle, and falsely labeled it as a landfill. The agent is asked to generate a concise explanation in a few words, and to avoid direct disclosure of the answer, to guide the user to correct their mistake. In this way, the agent will offer insights into the item's materials and the environmental impacts of disposing of them so that users can make another choice based on the clues they are offered. Additionally, a more detailed tip will be summarized and made available after the completion of the quiz. The AI agent will also compile a comprehensive and informative summary as a tip (Figure 1c), to reflect and illustrate all the information about the waste items in the quiz, their corresponding categories, and nevertheless, the users' sorting processes.

3.2.4. Virtual carbon credits

Waste Genie introduces the concept of virtual "carbon credits", which was initially issued by the United Nations [40], to measure the CO2 equivalent as the impact of users' waste classification efforts. Adopting the EPA's national overview on waste and recycling [41], we estimate the environmental benefits from correctly processing a unit of waste item through recycling, landfill, and composting. Furthermore, to enhance the user's understanding, alternative calculations for "carbon equivalents" are provided. By utilizing the green algorithm [42], the virtual carbon credits are converted into *tree-days* and *travel mileage* (Fig. 1d), which represents the number of days a tree can offset the equivalent carbon emis-

[1]https://platform.openai.com/

sions, and the emission produced by traveling certain miles in a car. The design mimics the complex carbon emission concept and paints a picture of the positive environmental effects of proper waste disposal.

3.3. User Study Design

A total of 54 students were recruited from two web usability classes at the author's university. Students were introduced to the study as part of the learning activity. Prior to using the Waste Genie platform, each participant was asked to complete a pre-survey that included basic sustainability awareness questions and waste categorization knowledge tests. There were 15 waste items being asked to classify into one of four categories: recycle, landfill, compost, or hazard. This pre-survey collected a baseline understanding of the participants' existing knowledge and attitudes toward environmental sustainability.

Participants were then asked to use Waste Genie over two days to complete a series of 30 pre-generated interactive waste sorting quizzes. Each participant received the same quizzes in the same order. During this period, participants' activities, including the time spent to complete each question, the steps used to complete a quiz, the mistakes they made, and the interactions with the AI agent, were recorded for the following data analysis.

A post-survey was distributed among the users after the completion of quiz sorting. The questionnaire included the same set of sustainability awareness questions, a reassessment of their waste classification knowledge, and a WG usability survey. The system usability scale (SUS) questions [43] were deployed to pertain to the systematic evaluation of the overall WG usability. We also gathered qualitative feedback on participants' opinions and suggestions about their experience with Waste Genie through open-ended questions in the survey.

3.4. Data Analysis

During the study, we collected 4,914 waste-sorting actions performed by the 54 participants on WG. We examined these participants' pre-knowledge distribution to analyze the differences and learning effects. A threshold of waste categorization accuracy of 0.8 was found based on the median accuracy in the pre-survey. We then distinguish these users into two groups, lower performing group (LPG, 33 users) and higher performing group (HPG, 21 users). This classification enabled us to analyze and compare the performance and responses of participants with varying levels of proficiency in waste classification. The participants with accuracy below or equal to the threshold were assigned to the LPG, while those with accuracy above the threshold were assigned to the HPG. To evaluate the effectiveness of the Waste Genie platform for improving waste management skills (RQ2), we measured the accuracy and time taken per quiz for each user and conducted an ANOVA analysis. The results from the waste sorting knowledge test and the sustainability awareness questions in our pre- and post-study surveys were also compared to evaluate the knowledge growth. System usability (RQ3) was assessed by calculating the SUS score from the post-study survey responses. The open-ended feedback also provided qualitative insights into our evaluations of users' experiences. As for the impact of AI-guided assistance (RQ1), we evaluated it by comparing users' accuracy in sorting waste items before and after receiving AI feedback. The number of AI guidance triggered per user was also tracked to gauge their engagement with the agent.

4. Evaluation Results

4.1. Efficient waste sorting practices: both LPG and HPG consistently increased waste sorting accuracy and efficiency

One of the underlying hypotheses in WG is that people can learn waste management from sorting virtual waste and reading organized waste management information. To collect the evidence of learning, we first measured each user's sorting accuracy throughout the 30 assigned quizzes. We then analyzed the trajectory of the waste sorting accuracy over time. We found that the overall sorting accuracy was 0.78, the average total sorting time spent per user was 294.8 seconds, and the average time spent on WG was 914.6 seconds (roughly about 15 minutes).

To dive deeper into the analysis, we found that the users showed an improvement in accuracy as they progressed through all the quizzes. As shown in Table 1 and Fig. 3, the average accuracy significantly ($p < 0.01$, Cohen's $d = 0.44$) increased from the first 6 quizzes ($M = 0.79$, $SD = 0.18$) to the last 6 quizzes ($M = 0.87$, $SD = 0.18$). The users also illustrated a steady improvement. To put it in perspective, we also used the amount of time spent per quiz to examine the waste sorting efficiency. We found that the users demonstrated a significant time reduction in waste classification ($p < 0.01, d = 0.71$) at the end of the sorting, from 15.34 ($SD = 12.99$) seconds to 7.5 ($SD = 8.48$) seconds. Such a finding is not trivial, it showed efficient waste sorting practices just within less than 5 minutes of overall use in WG. This trend was consistent across both the LPG and HPG. Despite variations in initial sorting proficiency, both groups showcased similar patterns of improvement in accuracy (0.79 to 0.88 for LPG and 0.80 to 0.87 for HPG) and decreasing sorting time (14.97s to 6.81s for LPG and 15.88s to 8.52s for HPG) as they engaged with WG. It is worth noting that the entire participant group illustrated a more coherent sorting efficiency after they had progressed 40% of the quizzes (the time spent was less than 9 seconds and the standard deviation was small, ranging from 5.04 to 8.48).

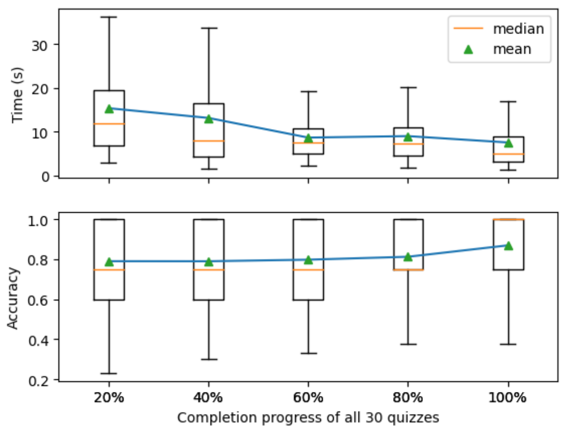


Figure 3: Accuracy and time used per quiz as users were making progress on the platform.

Table 1
Average accuracy and time (seconds) used per quiz as they were making progress on the platform.

Progress	All users		LPG		HPG	
(quizzes)	Accuracy	Time	Accuracy	Time	Accuracy	Time
20% (6)	0.79 ± 0.18	15.34 ± 12.99	0.79 ± 0.18	14.97 ± 13.18	0.80 ± 0.17	15.88 ± 12.61
40% (12)	0.79 ± 0.21	13.11 ± 14.08	0.78 ± 0.22	13.49 ± 14.82	0.82 ± 0.19	12.74 ± 13.16
60% (18)	0.80 ± 0.18	8.62 ± 5.04	0.80 ± 0.18	8.62 ± 4.98	0.80 ± 0.17	8.79 ± 5.18
80% (24)	0.81 ± 0.17	8.96 ± 7.13	0.82 ± 0.16	8.42 ± 5.25	0.81 ± 0.18	9.81 ± 9.27
100% (30)	0.87 ± 0.18	7.50 ± 8.48	0.88 ± 0.17	6.81 ± 7.00	0.87 ± 0.18	8.52 ± 10.35

Additionally, based on the knowledge tests in pre- and post-study surveys (see Table 2), the users showed overall knowledge growth to sort different types of waste. The overall average score improved from 0.79 in the pre-test to 0.91 in the post-test ($p < 0.01$). The compost type achieved the highest score in the post-test, nearing 100%. This suggests that the users comprehend well the concept of composting throughout the entire WG experience (sorting exercises and feedback). On the other hand, the *Landfill* category presented the greatest challenge among all types. The users exhibited a low average pre-knowledge score (67%) to begin with, and the LPG (57%) and HPG (83%) groups appeared to have a big accuracy gap. Although all users managed to increase their overall average accuracy to 81% in the post-test, HPG showed a slight knowledge score drop, which means the overall knowledge growth was predominately attributed to the LPG's growth. Such a result may be attributed to the unique definitions of compost, recycle, and hazardous wastes compared to the more ambiguous nature of landfill wastes. For instance, a piece of food-stained plastic consists of the key ingredients to dispose of in the compost and recycling categories. However, depending on the materials and the cleanliness, the item will be expected to be sorted into the landfill bin. Overall, these findings underscore the complexity and ambiguity of waste types, and the necessity of adaptive feedback, and suggest avenues for further research to enhance learning outcomes in waste management education.

Table 2
Before and after the experiment, users' capability of sorting different types of wastes increased.

	All Users		LPG		HPG	
Waste Type	Pre	Post	Pre	Post	Pre	Post
Compost	0.86	0.98	0.81	0.92	0.92	0.98
Recycle	0.82	0.93	0.74	0.95	0.95	0.99
Landfill	0.67	0.81	0.57	0.83	0.83	0.75
Hazardous	0.78	0.92	0.68	0.94	0.94	0.98

4.2. Positive effects of AI-generated feedback: increased first attempt success

With the improved waste sorting accuracy and efficiency over time, we were motivated to find out how much of the improvement was attributed to the AI-generated feedback. We looked into the statistics of tips and feedback usage to analyze what happened before and after the feedback was provided. We found that on average, each user received approximately 30.96 times AI-generated content. This number means that on average a user will interact with the AI agent at least once per quiz, either by making the mistake and receiving the feedback or explicitly clicking on the tips to review the sorting summary. It showcased their active engagement with the AI agent. Notably, users exhibited an increase in accuracy after receiving AI guidance, with an average first-attempt success rate of 0.76 compared to 0.69 before the assistance (Table 3). Such an effect was found in both LPG and HPG groups. Specifically, HPG showed significant accuracy improvement after the AI guidance ($p < 0.0167$), from 0.68 to 0.79 (Cohen's $d = 1.16$). The p-value was adjusted with Bonferroni correction with $\alpha = 0.05$. It is reassuring to learn that AI-generated feedback has a positive influence in helping users achieve a higher first-attempt success rate. The finding is also encouraging to learn that HPG which has already had higher pre-knowledge can capitalize on the AI guidance and improve itself further.

Furthermore, based on the post-survey, users also provided high ratings on the helpfulness of the AI-generated content ($M = 3.64, SD = 1.11$) and the AI agent's contribution to their understanding ($M = 3.68, SD = 1.04$). These subjective assessments along with the observed improvements in the data logs suggested that the AI agent indeed enhanced the waste sorting effectiveness.

4.3. Positive impacts on sustainability awareness

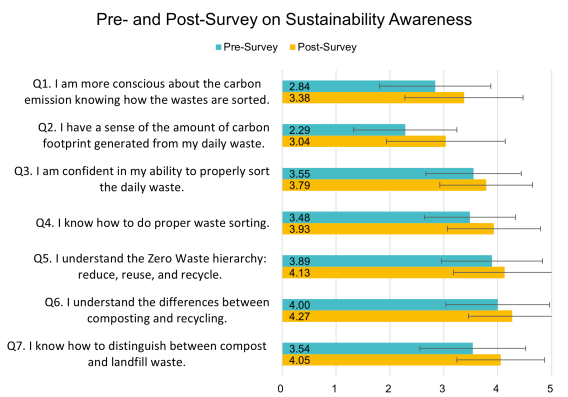


Figure 4: Survey of sustainability awareness before and after the usage of the platform.

The pre- and post-study sustainability awareness questions from the surveys revealed that our platform had a positive impact on increasing users' knowledge and attitudes to sustainable waste management practices (Fig. 4).

Table 3
The number of AI-guidance generated per user and users' performance before and after they saw the tips from AI.

	All users	LPG	HPG
Avg. AI-guidance received per user	30.96 ± 11.67	31.18 ± 10.40	30.48 ± 9.59
First-attempt success before AI-guidance	0.69 ± 0.08	0.70 ± 0.07	0.68 ± 0.09
First-attempt success after AI-guidance	0.76 ± 0.11	0.74 ± 0.10	0.79 ± 0.10

Users' core concepts such as *"distinguishing between compostables and landfills"* increased from 3.54 (standard deviation $SD = 0.98$) to 4.05 ($SD = 0.81$), *"understand the differences between composting and recycling"* increased from 4.00 ($SD = 0.96$) to 4.27 ($SD = 0.81$), ratings on *"know how to do proper waste sorting"* increased from 3.48 ($SD = 0.84$) to 3.93 ($SD = 0.86$), and ratings on *"confidence in my ability to properly sort the daily waste"* increased from 3.55 ($SD = 0.88$) to 3.79 ($SD = 0.86$).

Interestingly, the users found an increasing understanding of carbon emission on waste sorting and their daily waste carbon footprint (Q1 & Q2 - carbon emissions with the sorted waste ($M = 3.38, SD = 1.09$) and daily waste carbon footprint ($M = 3.04, SD = 1.10$), but the ratings of these two questions were still considered relatively low. After the exposure to sequences of waste sorting quizzes and AI feedback in WG, the users indicated their awareness and consciousness level increased, but the concept of the association between waste sorting and carbon emission is still arguably insufficient (just a bit higher than the neutral reference point in the Likert scale).

4.4. Waste Genie demonstrated high system usability

According to the System Usability Scale [43] in the post-study survey, WG was found to achieve good usability across the board. Specifically, users thought the system was easy to use ($M = 3.93, SD = 0.88$) and they felt confident while using it ($M = 3.88, SD = 0.87$). Users also expressed that *"I would imagine that most people would learn to use Waste Genie very quickly"* (M= 4.04, $SD = 0.91$). Such an outcome is consistent with our previous findings [44].

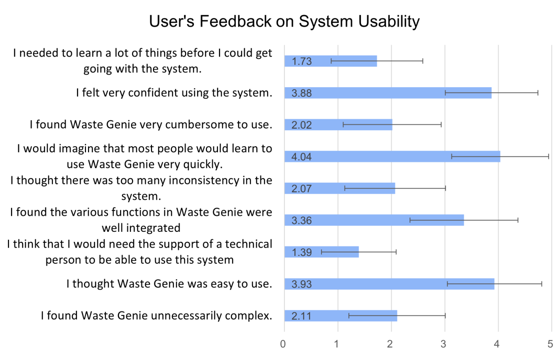


Figure 5: Results from the system usability survey.

5. Conclusions

5.1. Summary

In this work, we proposed to leverage the power of off-the-shelf GPT to generate adaptive feedback to facilitate waste sorting. We implemented the AI agent and integrated it into the web-based application Waste Genie (WG). We designed and conducted a user study to evaluate the effectiveness of learning and the impact of AI-generated feedback. The results demonstrated the feasibility of leveraging such AI agents to enhance educational technologies and effectively support waste management education.

Participants exhibited increased waste sorting accuracy and efficiency as well as improved sustainability awareness after engaging with WG. The AI-generated feedback had a positive impact on users' first-attempt success rates for waste sorting (RQ2), highlighting the potential of using off-the-shelf GPT models to provide concise guidance in the complex domain of waste classification. Furthermore, users demonstrated positive knowledge growth and heightened sustainability awareness (RQ1), underscoring the educational implications of WG, particularly in the context of informal lifelong learning and the global concern surrounding waste management and sustainability education. Waste Genie's usability was again validated by the users with overall high ratings (RQ3), it showcased the viability of the integration of AI feedback in WG.

5.2. Limitations and future work

Several limitations and avenues for future research have been identified during our study. For instance, the AI explanations were sometimes too brief and could have been better if there had been more conversational interactions. One user suggested that a smarter arrangement of quizzes and pace of explanations, such that sorting similar wastes together may aid in the reinforcement of knowledge.

Future iterations of our work could address these issues by further tuning the AI agent for dynamic and deeper dialogue; optimizing the quiz generation algorithms to add adaptive features for personalized sequences and paces; and providing more types of challenges to boost the learning process.

References

[1] W. Lin, Yolo-green: A real-time classification and object detection model optimized for waste management, in: 2021 IEEE International Conference on Big Data (Big Data), IEEE, 2021, pp. 51–57.

[2] P. F. Proença, P. Simões, Taco: Trash annotations in context for litter detection, arXiv preprint arXiv:2003.06975 (2020).

[3] A. B. Wahyutama, M. Hwang, YOLO-based object detection for separate collection of recyclables and capacity monitoring of trash bins, Electronics (Basel) 11 (2022) 1323.

[4] E. Longo, F. A. Sahin, A. E. C. Redondi, P. Bolzan, M. Bianchini, S. Maffei, A 5g-enabled smart waste management system for university campus, Sensors (Basel) 21 (2021) 8278.

[5] I. Reif, F. Alt, J. D. Hincapié Ramos, K. Poteriaykina, J. Wagner, Cleanly: trashducation urban system, in: CHI'10 Extended Abstracts on Human Factors in Computing Systems, Association for Computing Machinery, 2010, pp. 3511–3516.

[6] R. M. Jacobsen, P. S. Johansen, L. B. L. Bysted, M. B. Skov, Waste wizard: Exploring waste sorting using ai in public spaces, in: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society, 2020, pp. 1–11.

[7] A. Thieme, R. Comber, J. Miebach, J. Weeden, N. Kraemer, S. Lawson, P. Olivier, " we've bin watching you" designing for reflection and social persuasion to promote sustainable lifestyles, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2012, pp. 2337–2346.

[8] H. Kumar, D. M. Rothschild, D. G. Goldstein, J. M. Hofman, Math education with large language models: Peril or promise?, Available at SSRN 4641653 (2023).

[9] S. Sarsa, P. Denny, A. Hellas, J. Leinonen, Automatic generation of programming exercises and code explanations using large language models, in: Proceedings of the 2022 ACM Conference on International Computing Education Research-Volume 1, 2022, pp. 27–43.

[10] C. Xiao, S. X. Xu, K. Zhang, Y. Wang, L. Xia, Evaluating reading comprehension exercises generated by llms: A showcase of chatgpt in education applications, in: Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023), 2023, pp. 610–625.

[11] G. Christensen, G. Corpron, J. Engel, P. Hoila, The llm-based mini-lecture generation for enhanced learning project (2023).

[12] P. Sridhar, A. Doyle, A. Agarwal, C. Bogart, J. Savelka, M. Sakr, Harnessing llms in curricular design: Using gpt-4 to support authoring of learning objectives, arXiv preprint arXiv:2306.17459 (2023).

[13] S. Asthana, T. Arif, K. C. Thompson, Field experiences and reflections on using llms to generate comprehensive lecture metadata (????).

[14] H. A. Nguyen, H. Stec, X. Hou, S. Di, B. M. McLaren, Evaluating chatgpt's decimal skills and feedback generation in a digital learning game, in: European Conference on Technology Enhanced Learning, Springer, 2023, pp. 278–293.

[15] M. Pankiewicz, R. S. Baker, Large language models (gpt) for automating feedback on programming assignments, arXiv preprint arXiv:2307.00150 (2023).

[16] Z. A. Pardos, S. Bhandari, Learning gain differences between chatgpt and human tutor generated algebra hints, arXiv preprint arXiv:2302.06871 (2023).

[17] J. V. Dempsey, Interactive instruction and feedback, Educational Technology, 1993.

[18] P. Brusilovsky, et al., Adaptive and intelligent technologies for web-based eduction, Ki 13 (1999) 19–25.

[19] S. Gombert, A. Fink, T. Giorgashvili, I. Jivet, D. Di Mitri, J. Yau, A. Frey, H. Drachsler, From the automated assessment of student essay content to highly informative feedback: a case study, International Journal of Artificial Intelligence in Education (2024) 1–39.

[20] N. Yannier, S. E. Hudson, H. Chang, K. R. Koedinger, Ai adaptivity in a mixed-reality system improves learning, International Journal of Artificial Intelligence in Education (2024) 1–18.

[21] I.-H. Hsiao, P. Brusilovsky, The role of community feedback in the student example authoring process: An evaluation of annotex, British Journal of Educational Technology 42 (2011) 482–499.

[22] S. Wang, C. Christensen, W. Cui, R. Tong, L. Yarnall, L. Shear, M. Feng, When adaptive learning is effective learning: comparison of an adaptive learning system to teacher-led instruction, Interactive Learning Environments 31 (2023) 793–803.

[23] A. Buragohain, B. Mali, S. Saha, P. K. Singh, A deep transfer learning based approach to detect COVID-19 waste, Internet Technology Letters 5 (2022).

[24] A. S. Padalkar, An Object Detection and Scaling Model for Plastic Waste Sorting (Doctoral dissertation), Ph.D. thesis, Dublin, National College of Ireland, 2021.

[25] S. Majchrowska, A. Mikołajczyk, M. Ferlin, Z. Klawikowska, M. A. Plantykow, A. Kwasigroch, K. Majek, Deep learning-based waste detection in natural and urban environments, Waste Management 138 (2022) 274–284.

[26] A. Assor, A. Prouzeau, P. Dragicevic, M. Hachet, Exploring augmented reality waste data representations for eco feedback., in: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1–4.

[27] Q. Sun, I.-H. Hsiao, S.-Y. Chien, Immersive educational technology for waste management learning: A study of waste detection and feedback delivery in augmented reality, in: International Conference on Human-Computer Interaction, Springer, 2023, pp. 509–515.

[28] Q. Sun, I.-H. Hsiao, S.-Y. Chien, Immersive educational recycling assistant (era): Learning waste sorting in augmented reality, in: International Conference on Immersive Learning, Springer, 2023, pp. 469–483.

[29] N. Jiménez Barriga, B. Hernández Villalba, Robote: Interactive educational tool to teach basic education children to classify and collect waste in their school environment, in: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1–8.

[30] K. Wang, Z. D. Tekler, L. Cheah, D. Herremans, L. Blessing, Evaluating the effectiveness of an augmented reality game promoting environmental action, Sustainability 13 (2021) 13912.

[31] K. N. Jespersen, R. Odgaard, K. Julsgaard, J. L. Madsbøll, M. H. Lundbak, M. Niebuhr, M. M. Skovfoged, M. Löchtefeld, Foodfighters-improving memory retention of food items through a mobile serious game, in: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, pp. 1–7.

[32] B. Robelia, C. Greenhow, L. Burton, Adopting environmentally responsible behaviors: How learning within a social networking application motivated students to act for the environment, Environmental Education Research 17 (2011) 553–575.

[33] F. Yalvaç, V. Lim, J. Hu, M. Funk, M. Rauterberg, Social recipe recommendation to reduce food waste, in:

CHI'14 Extended Abstracts on Human Factors in Computing Systems, 2014, pp. 2431–2436.

[34] D. Casado-Mansilla, D. Foster, S. Lawson, P. Garaizar, D. López-de Ipiña, 'close the loop' an ibeacon app to foster recycling through just-in-time feedback, in: Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems, 2015, pp. 1495–1500.

[35] Q. Sun, I.-H. Hsiao, Waste genie: A web-based educational technology for sustainable waste management, in: Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology, 2023, pp. 1–3.

[36] California Environmental Protection Agency, Calrecycle, https://calrecycle.ca.gov/, 2023. Accessed: 2024-05-12.

[37] Environmental Protection Agency, State universal waste programs in the united states, https://www.epa.gov/hw/state-universal-waste-programs-united-states, 2023. Accessed: 2024-05-12.

[38] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, M. Chen, Hierarchical text-conditional image generation with clip latents, arXiv preprint arXiv:2204.06125 (2022).

[39] N. Ruiz, Y. Li, V. Jampani, Y. Pritch, M. Rubinstein, K. Aberman, Dreambooth: Fine tuning text-to-image diffusion models for subject-driven generation, arXiv preprint arXiv:2208.12242 (2022).

[40] United Nations, United nations carbon offset platform, https://unfccc.int/climate-action/united-nations-carbon-offset-platform, 2015. Accessed: 2024-05-12.

[41] Environmental Protection Agency, National overview: Facts and figures on materials, wastes and recycling, https://www.epa.gov/facts-and-figures-about-materials-waste-and-recycling/national-overview-facts-and-figures-materials, 2022. Accessed: 2024-05-12.

[42] L. Lannelongue, J. Grealey, M. Inouye, Green algorithms: quantifying the carbon footprint of computation, Advanced science 8 (2021) 2100707.

[43] J. Brooke, et al., Sus-a quick and dirty usability scale, Usability evaluation in industry 189 (1996) 4–7.

[44] Q. Sun, I.-H. Hsiao, Waste genie: Learning environmental sustainability from waste sorting and interactive feedback, in: 2023 IEEE Global Humanitarian Technology Conference (GHTC), IEEE, 2023, pp. 310–317.