# Designing Immersive, Narrative-Based Interfaces to Guide Outdoor Learning


Alan Y. Cheng
Stanford University
Stanford, CA, USA
ayc@stanford.edu

Jacob Ritchie
Stanford University
Stanford, CA, USA
jritchie@stanford.edu

Niki Agrawal
Stanford University
Stanford, CA, USA
nikhar@stanford.edu

Elizabeth Childs
Stanford University
Stanford, CA, USA
elchilds@stanford.edu

Cyan DeVeaux
Stanford University
Stanford, CA, USA
cyanjd@stanford.edu

Yubin Jee
Stanford University
Stanford, CA, USA
yubinjee@stanford.edu

Trevor Leon
Stanford University
Stanford, CA, USA
trevleon@stanford.edu

Bethanie Maples
Stanford University
Stanford, CA, USA
bethanie@stanford.edu

Andrea Cuadra
Stanford University
Stanford, CA, USA
apcuad@stanford.edu

James A. Landay
Stanford University
Stanford, CA, USA
landay@stanford.edu



**Figure 1: Our system, *Kiki's Adventure*, is a mobile app that uses narrative to engage fourth, fifth, and sixth graders in outdoor learning. In the leftmost image, a participant interacts with the outside world while using *Kiki's Adventure*. The leftmost screenshot portrays Kiki, the virtual koala, on a static background (narrative-only system variant). The middle screenshot depicts Kiki asking for help finding a eucalyptus tree in the real world, an interaction that happened in both system variants. In the screenshot on the right, augmented reality (AR) is used to put Kiki and Coco, another virtual koala, on the tree (augmented reality and computer vision variant)**

.

## ABSTRACT

Outdoor learning experiences, such as field trips, can improve children's science achievement and engagement, but these experiences are often difficult to deliver without extensive support. Narrative in educational experiences can provide needed structure, while also increasing engagement. We created a narrative-based, mobile application to investigate how to guide young learners in interacting with their local, outdoor environment. In a second variant, we added augmented reality and image classification to explore the value of these features. A study ($n = 44$) found that participants using our system demonstrated learning gains and found the experience engaging. Our findings identified several major themes, including participant excitement for hands-on interactions with nature, curiosity about the characters, and enthusiasm toward typing their thoughts and observations. We offer a set of design implications for supporting narrative-based, outdoor learning with immersive technology.

## CCS CONCEPTS

• **Human-centered computing** → *Interactive systems and tools*; **Mobile devices**; **Mixed / augmented reality**.

## KEYWORDS

augmented reality, outdoor learning, machine learning, immersive technology, mobile computing, ubiquitous computing, computer vision

**ACM Reference Format:**

Alan Y. Cheng, Jacob Ritchie, Niki Agrawal, Elizabeth Childs, Cyan DeVeaux, Yubin Jee, Trevor Leon, Bethanie Maples, Andrea Cuadra, and James A. Landay. 2023. Designing Immersive, Narrative-Based Interfaces to Guide Outdoor Learning. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 22 pages. https://doi.org/10.1145/3544548.3581365

## 1 INTRODUCTION

Prior research shows that most learning happens *outside* of formal learning environments, such as schools [8]. Even children in K-12 schools spend the majority of their time outside of the classroom, with some estimates putting the number at above 80% [8]. This presents a promising opportunity for developing learning tools to support learning outside the classroom. For example, we are surrounded by natural environments—excellent spaces for teaching subjects like biology and ecology. Trees, flowers, and wild animals, all readily available outdoors, can be used to make abstract scientific concepts in these subjects more concrete. Outdoor learning can also help increase students' interest in science [18] and feelings of connection to nature [32].

Despite the many benefits of outdoor learning, it is often difficult to implement in practice. Elementary school teachers are constrained by the heavy content demands of curriculum requirements and by limited time, making it difficult to carry out outdoor learning activities [7]. For outdoor learning, the widespread use of mobile devices, such as smartphones, can enable self-directed, readily-accessible educational experiences [43].

One way to provide structure to learning when outside the classroom is through narrative. Narrative-based learning has been shown to improve engagement and learning gains [53], and technology can be used to harness the power of narrative towards providing structure for outdoor learning. Moreover, recent technological advances make it possible to use augmented reality (AR) and context-aware technology like computer vision (CV) in mobile devices to interpret a user's surroundings and use this data to guide the learner in understanding the world around them. However, there is limited research about how to effectively integrate outdoor learning, narrative, and these immersive technologies. **The goal of the research described here is to address this gap by investigating how narrative-based mobile applications could guide young learners to learn about and engage with their local, outdoor environment. A secondary goal is to explore how integrating immersive technologies like AR and CV might support these activities.**

We created a narrative-based mobile application to teach fourth through sixth graders about a tree common in their local area: the eucalyptus tree, which has a rich history and is the subject of active debate. These grade levels were chosen because the ecological concepts relating to eucalyptus best matched the science standards for fourth through sixth grade. In our app, the learner meets a virtual koala named Kiki, who needs the learner's help to find a real eucalyptus tree. This sparks an adventure where the learner investigates the characteristics of eucalyptus trees, learns about their local history, and eventually weighs in on the ecological debate surrounding them.

To explore the potential benefits of immersive technologies that can understand and interact with the real world for outdoor learning, we also built a second variant of the app with CV (specifically, image classification) and AR. This variant had two aims. The primary aim was to enable a design exploration of newer features and tradeoffs. The secondary aim was to generate preliminary data for a formal comparison in the future. We compared the two app variants using the same criteria using pre- and post-intervention questionnaires in a between-subjects exploratory study.


*CHI '23, April 23–28, 2023, Hamburg, Germany*

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3581365

We investigate three main research questions:

**RQ1:** How do narrative, AR, and context-aware technology affect learning gains, engagement, and attitudes towards science, technology, and nature?

**RQ2:** How do fourth, fifth, and sixth graders interact with this technology (both the narrative-only and CV+AR variants) in an outdoor learning environment? What do they like and dislike about the experience?

**RQ3:** What are the opportunities and challenges of combining narrative, AR, and context-aware technology for outdoor learning?

We make three major contributions to the literature. First, we contribute the system with two variants used in our study for self-directed, narrative-based outdoor learning. Second, we provide quantitative and qualitative results from an exploratory study investigating how immersive educational technology is used by fourth, fifth, and sixth graders. We find that participants experienced learning gains and an upward shift in pro-environmental attitudes, and perceived the educational experience as engaging. Qualitative findings reveal that learners enjoyed interacting with the characters in the narrative, were surprised and excited by the hands-on and multisensory activities in our apps, and were enthusiastic about typing their thoughts and observations about nature. Third, we offer a set of design implications for supporting narrative-based, outdoor learning with immersive technology, such as the need to script learners’ movements when embedding such technology within a physical space. As a whole, our study’s artifacts and findings set a foundation for building self-directed and scalable outdoor lessons to supplement existing science education without overburdening schools and teachers.

In this paper, we first situate our study within the existing literature, describing the learning theories we build upon and the related work we contribute to. We then explain our design process—how we developed the learning content and narrative. We describe our system in detail, documenting the core features such as dialogue, in-app interactivity, and the CV and AR features. We detail our methods, share our findings, and discuss their implications.

# 2 RELATED WORK

Here we first describe the theories of learning that informed our artifact and study design. Next, we situate our work within existing research on narrative. We then survey existing work that measures the impact of outdoor learning, and technologies that have been used to guide outdoor learning, including AR. We highlight how our work is unique in that our system combines narrative-based and outdoor learning.

## 2.1 Relevant Theories of Learning and Education

Experiential learning is the process of learning through hands-on activities followed by reflective observation and abstract conceptualization. It has been applied in outdoor environments to increase students’ learning and engagement [11, 27, 29]. Place-based learning, which increases the relevancy of students’ learning by emphasizing the connection between scientific concepts and a learner’s local community and environment [57], can further increase engagement [58, 67]. Technology can support place-based learning by drawing a learner’s attention to how aspects of their location are relevant to specific fields of knowledge [66]. Our study design builds on the theories of experiential and place-based learning to deliver a locally-relevant, hands-on, interactive lesson.

## 2.2 Narrative-Based Learning

Our work extends existing lines of research on narrative-based learning materials, which embed educational content inside fictional or non-fictional stories. This creates a framing device where readers learn by progressing through a narrative. Research has found that including narrative elements in science education can help students retain facts [41]. Additionally narrative elements can help with 21st century skills (i.e., critical thinking, collaboration, creativity) [50] and provide cognitive support [13, 44]. Narrative can also produce an increased level of immersion, helping learners focus and stay engaged [35, 38, 39]. Recent work on educational escape rooms, for example, has explored combining digital technologies with narrative to engage learners and encourage collaborative problem solving [36]. However, other work has found that narrative-based learning may be ineffective in certain scenarios. For example, because the overall content is longer, learners may ignore parts of the narrative or educational content [19].

This points to a need for careful work to examine whether a particular narrative is appropriate for the learning domain in which it is deployed. Critical narrative elements include a clear storyline design [1, 65], and character design [13], all of which we incorporate into the narrative we developed for this study. Recent research in HCI has investigated how insights from narratology can be used to craft narratives that promote engagement with technology [40]. Like research on narrative-based learning in general, research on narrative-based educational technologies has shown mixed results, due in part to the cost and difficulty of creating educational technology systems [39, 53]. Some studies indicate that learning gains from narrative-based technologies are non-significant [1, 65], whereas others seem to indicate strong learning gains [12, 47, 53]. Our project extends this past research by investigating how a fictional narrative can support learners in the specific context of outdoor learning.

## 2.3 Outdoor Learning

Outdoor learning experiences enable children to learn while interacting with the outdoor environment. Prior work has demonstrated that outdoor learning experiences can positively impact children’s science achievement, engagement in learning, environmental attitudes, and connection to nature [14, 20, 29, 32, 51] with mixed results regarding their impact on self-efficacy [21, 60]. Moreover, several studies have shown that the benefits of outdoor learning may be more pronounced for students who are traditionally underrepresented in science [21, 60]. Our study relies on the widespread availability of mobile devices and the power of narrative to build a system that will eventually allow us to deliver self-directed, outdoor learning at scale.

### 2.4 Outdoor Learning Driven by Context-Aware Technology

Digital technologies are being increasingly used to augment and scaffold children's outdoor science learning experiences and have led to improved learning outcomes, including increased interpretation, reflection, and analysis among students [9, 11, 37, 52, 68]. Some prior work has relied on strategic placement of devices or scannable visual markers (such as QR codes) that trigger content and digital augmentations at specific locations to create a context-aware, outdoor learning experience. For instance, in the Ambient Wood project, students explored a woodland area while interacting with a variety of devices, including personal digital assistants, probes, and periscopes [52]. These devices enabled the students to read environmental data and receive images and audio information triggered at particular locations. Chien and colleagues [11] developed a mobile augmented reality system that used scannable codes to trigger 3D models of corresponding plant species.

While the context-aware technologies used in these outdoor learning experiences had a positive impact on students' learning outcomes, such experiences require others to equip the environment in advance with additional hardware devices or scannable visual markers. These dependencies would complicate attempts to enable self-directed, accessible learning and to replicate such experiences at scale.

Several smartphone applications leverage image classification to identify plants and animals in the photos taken by users and provide corresponding educational content [33, 45, 46]. These apps provide context-sensitive, educational content in outdoor settings without the need for strategically placed devices or markers or content triggered by geolocation. However, these apps are designed for a general audience rather than being catered towards children. Furthermore, their educational content does not go much deeper than providing factual information. Our study builds on and contributes to this prior work by developing an image classification pipeline to identify eucalyptus trees and accordingly trigger context-sensitive feedback while providing a rich learning experience within a standalone educational app designed for children.

### 2.5 Augmented Reality in Education

AR has been used in children's education in a variety of fields, including physics [17], mathematics [30], and ecology education [15]. Prior work demonstrates that AR can be used in outdoor science education to increase learning outcomes, motivation, engagement, positive emotions, and levels of immersion among students [11, 15, 22, 43, 61]. In addition, research has found that embedding educational AR activities within a larger narrative or supplementing the activities with guided commentary improves learning gains and engagement [23, 27, 30]. For example, Huang et al. [27] found that middle school students who had access to AR activities providing rich information about plants in a botanical garden, supplemented by guided commentary, achieved improved learning gains compared to students who had access only to the AR activities or only to the guided commentary. Georgiou and Kyza [23] found that when narrative-based, digital content was closely connected to the learner's physical space (e.g., via digital augmentations placed at specific points of interest), participants experienced higher learning gains and feelings of immersion compared to a loosely-coupled control. While some studies suggest that there is a positive relationship between more immersive experiences and conceptual understanding in science [22, 23], other studies have findings that are inconclusive or illustrate a negative relationship [10, 55]. In this work, we advance this research by building and evaluating a variant of our app that contains AR features.

## 3 DESIGN PROCESS

First, we identified a set of core learning objectives that we wanted our system to support. Then, we developed a feasibility prototype targeted at adults with the help of industry partners to test features and receive feedback, which we used to inform the design of the final system. Finally, we brainstormed and designed the narrative concurrently with the development of the final app.

### 3.1 Educational Content

For the educational content in our outdoor learning app, we targeted ecology and plant biology. We chose to focus on plants because their stationary nature makes it easier to design an interactive experience around them. Also, there is a documented tendency to pay less attention to plants despite their importance to the biosphere, a phenomenon that has been termed "plant awareness disparity" [48, 64], so engaging people to learn about plants is a worthwhile endeavor.

Our app centers around eucalyptus trees because they are common in our area, have a rich history, and are the subject of ongoing debate. This provided an effective foundation for delivering novel and enriching content that aligns with Next Generation Science Standards [59], and connects to local and current events that learners would recognize (i.e., wildfires).

To inform the design of our learning content, we used the cognitive domain of the revised Bloom's Taxonomy to articulate three levels of learning and evidence of that learning: Remembering, whereby the learner recalls verbatim facts they have heard; Understanding, whereby the learner constructs meaning by interpreting information; and Application, whereby the learner uses information in a new setting or applies the concept to new understandings [2]. From there, we decided on the following learning objectives for our app, which were selected to target different levels within Bloom's Taxonomy:

(1) Identify eucalyptus trees and distinguish them from other trees.
(2) Understand the ecological and historical context of a local plant.
(3) Give an example of how human activity affects natural ecosystems and vice versa.
(4) Create an argument based on evidence (by weighing in on the controversy surrounding the ecological role of eucalyptus trees in the local area).

### 3.2 Feasibility Prototype

To explore the potential of immersive technologies for learning, we first co-developed an app with industry partners to help adults learn about eucalyptus trees that was feasible, viable, and desirable.

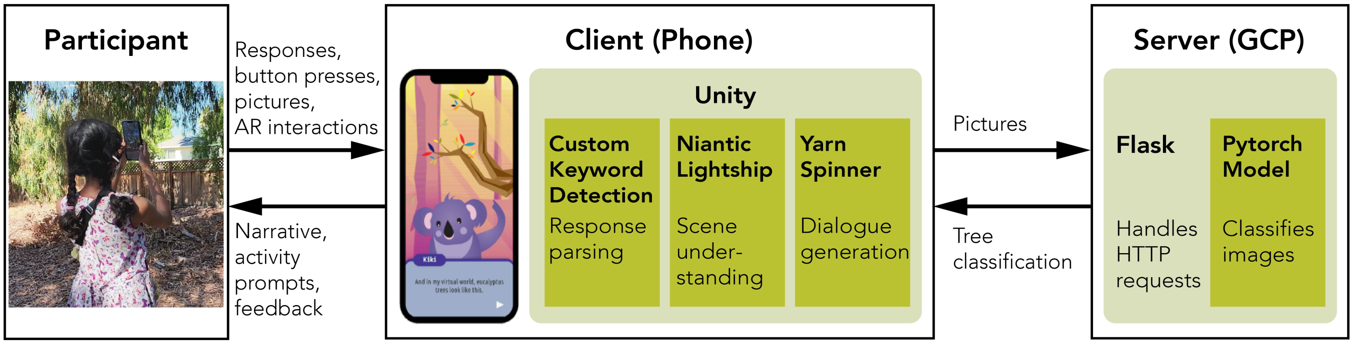


**Figure 2: Overview of our system design for the Narrative+CV+AR app variant. The participants interacts directly with the app on the phone, which handles all the narrative and AR features, while image classification is performed on a server. The Narrative-Only app shares the same infrastructure but does not use Niantic Lightship, nor does it upload photos to the server.**

We built a high-fidelity prototype to explore the feasibility of various features, including tree recognition, interactivity between the digital and physical worlds, and a chatbot to respond to questions the users might have. We tested this prototype with five adults affiliated with our institution, selected via convenience sampling. We found that the tree recognition, the interactivity between the digital and physical worlds, and interaction with the real environment were the most engaging features, so we decided to focus on these going forward.

## 3.3 Narrative Design

We used insights from our feasibility study, literature review, and consulting with children in the target age, allowed us to come up with the main requirements for the narrative. The two primary requirements were for a main character the learner could relate to and for the character's quests to be connected to the learning activities. With these requirements in mind, team members each wrote at least one narrative for the lesson. We then collected a total of nine possible narratives, and broke them apart by quest: reason for the main character to be at the study location, main character's motivation, tree identification, history of the trees, connection to wildfires, controversy surrounding the trees' role in the wildfires, and conclusion. Next, we discussed each quest and its relation to the learning activities to identify the strengths of the various ideas and then consolidated them into a cohesive narrative. For example, we selected the idea that Kiki, the main character, was a koala who lives in the virtual world and dreams of experiencing trees in the real world as a way to connect the quest to the activities involving augmented reality. Once we had a complete and cohesive story, we tested and refined it as we developed our system.

## 3.4 Final Narrative

Our app takes participants through a lesson that follows a story about a koala named Kiki who wants to learn about eucalyptus trees in the local area. In the narrative, the learner meets Kiki, who wishes to see what eucalyptus trees look like in the real world. Kiki asks the learner to find a eucalyptus tree in real life, and the learner interacts with the tree to learn more about its properties, such as its smooth, peeling bark and long, oily leaves. The pair then encounter another virtual koala, Coco, who has been living in the tree for a few years. Coco teaches the learner about the history of how eucalyptus trees arrived in the area. After the learner completes more educational activities, Kiki and Coco get into a disagreement about whether eucalyptus trees should be removed from the local area, since they contribute to wildfires and are an invasive species. The learner must choose one of the koalas to agree with or propose their own idea. Finally, they assist Coco and Kiki in returning to the virtual world.

In designing the narrative for *Kiki's Adventure*, we built off prior work that explored how theories from narratology might inform the design of framing narratives in an HCI context [40]. In particular, we found it important to include characters that learners could form an emotional attachment to, leading us to add the characters of Kiki and Coco. We include the storyboard for *Kiki's Adventure* in the Supplementary Material.

# 4 SYSTEM DESCRIPTION

In this section we describe the specific educational activities, capabilities, and technical implementation of the two variants of the *Kiki's Adventure* system. See Figure 2 for an overview.

## 4.1 Engine and Core Features

We developed our app in Unity, a popular cross-platform game engine, and we deployed it to a mobile device. In this subsection, we describe the components in the app common to both variants.

*4.1.1 Dialogue.* We implemented the narrative using a library called Yarn Spinner [34], which simplifies the writing and integration of dialogue into Unity. We also recorded voiceover audio for each line of Kiki's and Coco's dialogue. Our goal was to increase accessibility by offering both text and audio options and to encourage learners to look at the environment, rather than read on the phone. Each voiceover line had to finish playing before the learner could progress to the next line, to control for variable reading speeds among participants.

*4.1.2 Multiple-choice prompts and dialogue branching.* Drawing inspiration from video games, we designed the app to frequently prompt the learner for their input in the form of multiple-choice dialogue options. Our intention was to give the learner a more active role in the narrative to give them a greater sense of agency [4]. Depending on what option the learner chooses, the narrative will branch temporarily to react to that choice before resuming. This approach allows us to balance a linear narrative with the perception of meaningful choice.

*4.1.3 Free-response prompts.* Over the course of the narrative, Kiki and Coco ask the learner four free-response questions:

- What are some things you notice about the bark?
- What are some things you notice about the shape and smell of the leaf?
- Why would [reason] make wildfires worse? (Learner is asked to select a reason why eucalyptus trees contribute to wildfires before answering this question.)
- Why do you think eucalyptus trees should be removed? / Why do you think eucalyptus trees should stay? / What would you do about eucalyptus trees, and why? (Question depends on which stance the learner said they agree with.)

The former two questions encourage the learner to process what they see and verbalize it before the correct answer is revealed. This is based on the *generation effect*: Learners tend to remember information better when they generate it themselves, rather than simply read it [56]. These also correspond to the Understanding level of Bloom's taxonomy by having learners make sense of their observations.

The latter two questions ask the learner to take a stance and justify it with reasoning, which are core critical thinking skills. Furthermore, these questions align with the Application level of Bloom's taxonomy by asking learners to apply their newfound knowledge to the debate.

For the first three questions, the koalas reveal the correct answer after the learner responds. For the last question, however, the narrative positions the learner as the expert: Kiki and Coco disagree on what should be done about eucalyptus trees in the area and ask the learner for their suggestion.

*4.1.4 Keyword detection.* The app employed a very simple form of natural language understanding using keyword matching. Depending on keywords the app detected in the learner's answer to an open-ended question, the koalas' dialogue branches to customize their response. For example, if the learner's response to the bark observation question includes the word "flaky," Kiki will respond, "Yeah, exactly! Like you said, eucalyptus bark peels off in large strips." We opted for this approach as a lightweight way for the character in the narrative to interact meaningfully with the learner.

*4.1.5 Hardware.* Our target device was an iPhone 13 Pro Max, which we chose for its large screen size and its built-in Lidar sensor, which we wanted for more reliable depth sensing in the Narrative+CV+AR variant.

## 4.2 App Variants

We developed two variants of *Kiki's Adventure*: a narrative-only variant (Narrative-Only) and a variant with narrative, computer vision, and augmented reality (Narrative+CV+AR). Table 1 offers a comparison of the variants. The remaining subsections describe features exclusive to the Narrative+CV+AR variant.

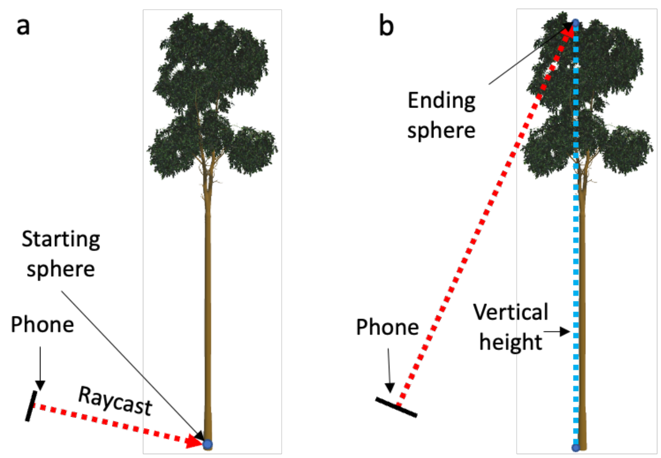


**Figure 3: Visualization of raycasting for height activity. *a)* The initial raycast to find the starting position. *b)* The final raycast against an invisible vertical plane to estimate the tree height.**

**Figure 4: Visualization of interaction for accurately putting the koala in the tree. *Crosshair:* location the user chose to place the koala. *Red arrow:* Trajectory of recursive raycast to find tree edge. Note: the trajectory of the raycast is not shown to the user.**

**Table 1: Story flow table showing the commonalities and differences between the two variants of our system. Left is the Narrative+CV+AR variant; right is the Narrative-Only variant.**

| | Narrative+CV+AR | Narrative-Only |
|---|---|---|
| 1 | Kiki, a virtual koala, wants to experience a real eucalyptus tree. | |
| 2 | Participant must find and identify a eucalyptus tree around them using a reference photo. | |
| 3 | Participant manually confirms whether the tree is a eucalyptus by comparing to the photo. | Image classification confirms whether the tree is a eucalyptus. |
| 4 | Participant places Kiki in the eucalyptus tree using AR. | Background in the app changes, but Kiki remains in the same spot on the screen. |
| 5 | Kiki asks participant to type their observations about the tree's bark before teaching them about the bark's properties. | |
| 6 | Coco, another virtual koala, appears in the tree using AR. | Coco, another virtual koala, appears on the screen. |
| 7 | Kiki and Coco ask participant to break a leaf off the real eucalyptus tree, smell the fragrant oil, and type their observations. | |
| 8 | Kiki and Coco explain the properties of eucalyptus leaves. | |
| 9 | Participant measures the height of the tree using AR. | Participant estimates the height of the tree through analogy to other objects. |
| 10 | Kiki and Coco ask participant to select a reason why eucalyptus trees contribute to wildfires. | |
| 11 | Kiki and Coco disagree on what to do about eucalyptus trees and ask participant to argue their own stance. | |
| 12 | Participant plants a virtual eucalyptus tree using AR. | *no equivalent activity* |
| 13 | Participant taps the koalas on the tree to bring them back to the virtual world. | *no equivalent activity* |

## 4.3 AR

The Narrative+CV+AR variant featured three augmented reality activities: measuring the tree height, placing the characters on the tree, and planting a tree. Below are in depth technical and motivational analyses of each AR aspect. For all of these activities, raycasting was generated using the depth buffer on an iPhone 13 Pro Max. To facilitate the AR activities, we used the Niantic Lightship's Unity plugin [42].

*4.3.1 Measuring tree height.* Part of the learning content referenced how tall and how quickly eucalyptus trees grow, and this measurement activity allowed the users to gather data to support their understanding of this content.

To measure the height of a tree, the user first placed a marker at the base of the tree using a raycast. Then, an invisible vertical plane was instantiated at the position of the marker. A second raycast was performed, and as the user angled their phone upwards, a vertical measurement line was shown between the tree base and the vertical position of the point of intersection of the raycast and the plane. Once the measurement line matched the tree's height, the user placed another marker at the top of the tree. The vertical distance between the two markers was calculated to estimate the height of the tree (see Figure 3).

*4.3.2 Appropriate positions: Koalas in the tree.* Bringing virtual characters into the real world was a key part of the narrative framing. Therefore, users were instructed to place a marker anywhere on the tree, similar to the height measurement activity. The system then conducted a recursive raycast to the right of the marker position to the border of the tree to place Kiki on the vertical tree edge (see Figure 4). Coco, the other koala, was similarly placed using a recursive raycast to the left of the marker and instantiated at that point in the narrative.

At the end of the experience, both koalas return into the phone. A crosshair marker (provided via raycasting) allowed the users to target each koala. By tapping either the crosshair or the koala they would pull the koalas out of the real world and back into the phone.

*4.3.3 Planting a sapling.* In an attempt to increase engagement and underscore the socially-relevant message about eucalyptus trees' fast speed of growth, the participants planted a virtual eucalyptus tree sapling in their environment by placing a marker on the ground at a minimum distance of six feet away to ensure the full content would be visible. We then showed a sequence of eucalyptus tree models, which grow progressively in size.

## 4.4 Image Classification Pipeline

Many activities in Kiki's Adventure involve observing and interacting with a eucalyptus tree. Thus, early in our narrative, Kiki guides the user to find a eucalyptus tree. In the Narrative+CV+AR variant, the user photographs the tree. If the photo is of a eucalyptus tree, Kiki congratulates the user and the lesson proceeds. Otherwise, the user is prompted to try again. Thus, the app is aware of the user's physical context and accordingly provides feedback to ensure the user is near a eucalyptus tree for the remainder of the lesson.

To enable this context-aware feedback, we created an image classification pipeline that takes as input a photo taken by the user and classifies the photo as being class "eucalyptus tree" or class "other tree." We created our image classification model by finetuning a ResNet-18 model [26] that had been pre-trained on the ImageNet dataset [54].

Our training data consists of a mix of images from iNaturalist, a citizen science initiative where the public can share observations of biodiversity [63], and pictures we took ourselves. We collected images from iNaturalist by randomly sampling research-grade observations. We constrain observations of the other tree class to lie within the rectangle of latitude/longitude lines that bounds our state's border. Our training data also includes photos of trees that we took at our user study site and in the surrounding area and labeled ourselves. In total, our training set consists of 4100 photos. Additional details on model training are provided in the Supplementary Material.

We tested the model's performance on two datasets. The first test set (53 eucalyptus trees, 86 other trees) was collected at our user study site, using our system's in-app camera feature, to validate end-to-end performance ahead of the user study. The second (211 eucalyptus trees, 452 other trees) used observations gathered by multiple users and devices at our user study site and in the surrounding area. We achieved 96% and 97% accuracy respectively on the two test sets (see Table 2).

We deployed the model using the Flask web framework and hosted on a Google Cloud Platform virtual machine. The mobile phone queried the server using mobile data, which resulted in latency ranging from one to thirty seconds, due to the large file size of the images and poor cellular signal strength at the user study site.

## 5 PILOT STUDY

We conducted a pilot study ($n$ = 14) to identify bugs, assess usability, gain insights into how learners interacted with our system, and refine the procedure described in detail below in Section 6.1.

We found that some participants, especially younger ones, were taking longer than expected to complete the open-ended questions in the pre- and post-lesson questionnaires because of the effort required to type their responses. In response, we changed the procedure to have all participants dictate the answers to those questions instead, which were recorded and later transcribed.

Upon analyzing the photos taken by pilot study participants when they were asked to identify a eucalyptus tree, we found that some participants took photos at sub-optimal angles and distances from the tree that would increase the likelihood of false negatives occurring. Through qualitative analysis of our model's performance on two test sets (as detailed in Section 4.4), we had determined that most false negatives were either photos where the ground, rather than the tree itself, occupied a significant portion of the photo or photo that were a close up of a eucalyptus tree's leaves. Thus, we added the following instructions before users took their photos: "Make sure you can see the trunk and branches of the tree. Try not to capture too much of the ground in your picture."

Also, we realized that the ordering of activities in the narrative needed to be changed to facilitate smoother movement within the physical space. We elaborate on this finding in depth in Section 8.5.

Finally, we made various minor improvements to increase the robustness of our app. For example, we updated the keywords in our keyword detection algorithm based on the responses from our pilot participants. We also implemented more ways to recover from simple errors, such as adding confirmation screens after the participant takes photos.

**Table 2: Precision and recall metrics of our image classification model for both test sets.**

| | Test Set 1 | | Test Set 2 | |
|---|---|---|---|---|
| | Precision | Recall | Precision | Recall |
| Eucalyptus Tree | 96% | 92% | 97% | 91% |
| Other Tree | 95% | 98% | 96% | 99% |

## 6 EVALUATION

This section details the procedure of our study, our measures, and how the data was gathered and analyzed.

### 6.1 Procedure

We conducted a study of our app ($n$ = 44) where participants were invited to come onsite to an outdoor location near our institution. Sessions were conducted with a single participant at a time. Each session took approximately one hour, and participants were compensated with a $20 gift card for their time.

Before each study session, we emailed participants' parents or guardians a copy of the study consent form. Parents or guardians were asked for consent and for participant demographics upon arrival at the entrance of a trail, and the learners were asked to assent to participate. All but one learner (P56) assented. With participants' assent, we asked them to fill out a pre-lesson questionnaire on a touchscreen device we provided.

Subsequently, we walked approximately 100 yards to a location with a eucalyptus tree nearby, and handed the participant a mobile phone with the study app opened. We asked participants to complete the lesson in a self-directed manner and to think aloud. Experimenters recorded written and video observations of participants engaging in the lesson and intervened only if participants became stuck. The lesson itself lasted 10-20 minutes.

After the lesson, we walked back to the entrance of the trail and asked participants to fill out the post-lesson questionnaire. Following this step, we interviewed participants to ask about their experience in the study. Finally, we emailed participants a delayed post-lesson questionnaire four days (96 hours) after their study session.

### 6.2 Condition Assignment

Once we felt that all major issues with the app and the study procedure had been addressed, we began our exploratory evaluation. Participants were assigned to one of two experimental conditions, corresponding to the two variants of our system (Narrative-Only or Narrative+CV+AR). To ensure a roughly even distribution across grade and gender, we assigned participants using systematic stratified randomization. Initial assignments were randomized by strata (where each stratum was a pair of gender and grade level), and assignments within each stratum alternated between the Narrative-Only and Narrative+CV+AR conditions.

## 6.3 Participants and Recruitment

A total of 44 children in grades 4, 5, and 6 (F:24, M:19, Other:1) participated in the study. Children attending schools near our study location in these grades are typically literate, able to navigate real-world spaces with some autonomy, and new to exploring ecological concepts—allowing our subject area to be novel but not intellectually overtaxing. Participants were recruited through university mailing lists, social media ads, physical flyers, calling and emailing local schools, and word of mouth. They attended 33 different schools. Of the participants, 45% reported their race as Asian, 23% white, 4.5% Hispanic or Latino. A further 25% reported two or more races, and 2% reported other. 11 children were in 4th grade, 13 in 5th grade, and 20 in 6th grade. The majority of participants (29 out of 44) had used AR at least once before in the past, primarily through Pokémon Go or Google Search.

## 6.4 Measures

Here, we describe how we set up our evaluation to answer the three research questions posed in Section 1. To address RQ1, we deployed questionnaires to measure participants' learning gains, engagement, and attitudes toward science, technology, and nature to understand how our system affects them. We also compare these measures between the two variants to identify any differences. For RQ2, we took observational notes while participants used our system, video recorded these sessions, and conducted post-completion interviews to learn how they interacted with the technology and what they liked and disliked about it. For RQ3, we reflected on our design process and the findings from RQ1 and RQ2 to identify opportunities and challenges of combining narrative, AR, and CV for outdoor learning.

*6.4.1 Pre-lesson and post-lesson questionnaires.* All participants completed a pre-test questionnaire as well as two post-test questionnaires (one immediately following completion of the task, and another after a delay of four days). The questionnaires measured learning gains, attitudes, and intervention desirability and effectiveness. Learning gains were measured using six content-related questions that were designed to target multiple levels of Bloom's Taxonomy. These questions ranged from fact-based Remembering questions ("List at least two reasons why eucalyptus trees can catch fire easily.") to more complex Application questions that asked participants to use what they learned in a new context ("Why do you think Australia has a wildfire problem?"). We asked an external researcher at our institution's school of education to review and provide feedback on our questionnaires before deploying them. All questionnaires employed are available in the Supplementary Materials.

To analyze learning gains, two researchers first coded the pilot study participants' responses to the learning content questions and developed a rubric, which was reviewed by an external expert. The rubric used a four-point ordinal scale (*no credit, weak, medium, strong*) for the short-answer questions, while the multiple choice questions were graded by correctness. The researchers independently coded the responses from the formal study using the rubric (blinded to condition), jointly compared the codes, and created a final set of codes for each response. Two participants were excluded from some questions due to audio data loss.

Shifts in attitudes toward science, technology, and nature were assessed using pre-post attitude items with five-point Likert scales. We used six items from the Students' Motivation Toward Science Learning (SMTSL) questionnaire [62] (e.g., *"I am not sure that I can understand difficult science ideas."*), and we adapted six more items from the SMTSL questionnaire to measure attitude toward learning science with technology (e.g., *"When I do not understand something in science, I use technology to find things that will help me"*). For attitude toward nature, we used ten items from the 2-MEV environmental attitudes scale [28], which has items that measure the affinity toward preservation of nature (e.g., *"I like to go on trips to places like forests away from cities."*), as well as the utilization (i.e., human exploitation) of nature (e.g., *"Building new roads is so important that trees should be cut down"*.)

Finally, we adapted fifteen items from the web-based learning tools (WBLT) evaluation scale [31] to measure the engagement and perceived effectiveness of our intervention (e.g., *"The app made learning fun."* and *"I felt like Kiki and Coco understand what I said."*). For these items, we used a five-point Smiley Face Likert Scale [25], and they were included only in the post-lesson questionnaire.

*6.4.2 Interview.* Following completion of the post-test questionnaire, we conducted a brief semi-structured interview about the participant's experience with the application. For example, we asked, "Overall how did that go?" and "How did this experience compare to ways you learned about plants in the past, like in school?" The interview script is provided in the Supplementary Material.

Interviews were recorded and transcribed using an automated transcription service. Transcripts were edited to fix transcription errors and to remove filler words. We created a codebook using interview transcripts from the pilot studies, and four researchers divided the transcripts from the formal studies and coded them using inductive coding. Lastly, one researcher went through each coded transcript and fixed inconsistencies.

# 7 FINDINGS

Our data revealed that our system was predominantly perceived as engaging and resulted in learning gains (RQ1). We also share major themes identified in our qualitative analysis, such as participants being surprised and excited about the hands-on and multisensory activities, and motivated to express their opinions in the app (RQ2).

## 7.1 Quantitative Analysis

First, we share our analysis surrounding learning gains and changes in attitudes towards science, technology, and nature. Within each exploratory condition, we compare pre- and post-intervention changes, and we also compare those change between the two conditions. In addition, we report engagement measures, participants' responses to the usability of our system, and their perceptions of our system's role in their learning.

*7.1.1 Learning gains.* We found strong evidence for learning gains. For this analysis, we had a sample size of $n = 42$ because the audio files for two participants were lost. Using Wilcoxon signed-rank tests to measure improvement between the pre- and post-intervention questionnaires, we observed a significant improvement across all six learning content questions ($p < 0.001$ for each, see Table 4 in the Appendix for details). Similarly, we observed significant learning gains for those same questions between the pre-intervention questionnaire and the delayed questionnaire that was sent out four days later ($p < 0.001$ for each, see Table 5 in the Appendix for details), indicating retention of learned content.

Additionally, we found a significant difference in learning gains between the two variants for Q1, "Why do you think Australia has a wildfire problem?" (Mann-Whitney U Test,$\text{Median}_{\text{Narrative-Only}}$ = 2, $\text{Median}_{\text{Narrative+CV+AR}}$ = 1, W = 304, p = 0.01795). These values indicate the difference between their post-test score and their pre-test score on a four-point scale (*no credit, weak, medium, strong*). For example, a value of 2 indicates an improvement from "no credit" to "medium" or from "weak" to "strong." We found no significant differences for the remaining questions and no significant differences in delayed learning gains (i.e., retention) between the two exploratory conditions.

*7.1.2 Motivation toward science learning.* In the following analyses we used Mann-Whitney U tests for between-group comparisons and Wilcoxon-Signed Rank Tests for within-group comparisons. We found that participants' responses related to motivation toward science learning improved from pre-test to post-test, when considering the summed score of the entire scale. Moreover, there was a significant difference between the median levels of improvement observed with the two variants of the system. While the Narrative-Only variant shows a significant improvement in participants' science attitudes, the Narrative+CV+AR variant does not. The pre-test science scores were significantly higher in the Narrative+CV+AR condition, which may have contributed to the lower observed level of improvement. See Table 6 and Figure 5 for additional details.

*7.1.3 Attitudes towards technology.* We found no significant difference in attitudes towards technology from pre- to post-test, or in the amount of change observed between variants. See Table 7 and Figure 5 for additional details.

*7.1.4 Environmental attitudes.* For both the preservation of nature and utilization of nature subscales, we found significant differences from pre-test to post-test. Overall, attitudes related to preservation of nature increased while attitudes related to utilization of nature decreased. For preservation attitudes, this difference was significant when the Narrative-Only variant was considered in isolation but not for Narrative+CV+AR variant. For utilization, it was significant for both Narrative-Only and Narrative+CV+AR. There was no significant difference between the median pre-test answers in the two conditions, for either the preservation or utilization subscales. See Table 8, Table 9, and Figure 5 for additional details.

*7.1.5 Engagement.* We found high levels of self-reported engagement (e.g., 84% of participants agreed or strongly agreed with the statement "I found the app engaging."). Moreover, participants were also engaged with the narrative itself. Most participants reported wanting to learn more about the two koalas (80%), and wanting to go on another adventure with Kiki (73%). Only one participant responded negatively toward the engagement questions. There was no significant difference in self-reported engagement between the two variants (Mann-Whitney U Test, $\text{Median}_{\text{Narrative-Only}}$ = 12.5, $\text{Median}_{\text{Narrative+CV+AR}}$ = 12 , W = 272.5, z= -0.71 , p= 0.48) or grade levels (Mann-Whitney U Test, $\text{Median}_{\text{grade6}}$ = 12, $\text{Median}_{\text{other}}$ = 12, W = 225 , z= -0.11, p= 0.91).

*7.1.6 Usability and perceptions of the app's role in learning.* Most participants thought the app was easy to use and helped them learn (see Table 10 in the Appendix for details). We did not observe any large differences in these categories between the two variants.

*7.1.7 Task completion.* Participants were able to complete the educational activities with minimal assistance. Almost all participants (42 out of 44) successfully identified the eucalyptus tree without the researcher intervening to help, and only one participant in each variant did not. Five participants encountered a false negative (i.e., the image classifier could not successfully identify the eucalyptus tree), but four out of those five were able to correct for it by taking another picture of the eucalyptus from a different position or angle. In all false negatives, significant portions of the photo were occupied by the ground rather than the eucalyptus tree. In some cases, the lighting conditions varied significantly from images in the model's training set. No participants encountered a false positive (i.e., the image classifier mistook another tree for a eucalyptus). Most participants in the Narrative+CV+AR condition (20 out of 22) were able to understand the AR features without external help.

In the activity where participants were asked to pick and sniff a eucalyptus leaf, participants occasionally made errors that required researcher intervention. Three out of 44 participants accidentally took a leaf from a small tree that was growing at the base of the eucalyptus. One participant misunderstood the instructions and picked up a piece of bark instead of a leaf. In these cases, the researcher stepped in to correct the misunderstanding.

## 7.2 Qualitative Findings

In this section we share our qualitative findings, including comparisons between the variants; reactions to the narrative, in particular the characters in the story; and participants' favorite parts of the experience.

*7.2.1 Participants found hands-on, multi-sensory interaction with nature to be surprising and enjoyable.* Many participants (14 out of 44) expressed surprise or requested confirmation when the app prompted them to break a leaf off of the eucalyptus tree or to take a picture of the tree, e.g., *"Am I supposed to find one?"* (P51) and *"Should I get the leaf or...?"* (P43). However, in the follow-up interviews, no participants reported actual confusion regarding these tasks. While children may be initially surprised when a mobile app asks them to perform a real-world task, they are able to make sense of it and execute accordingly.

As anticipated, we found that participants enjoyed the hands-on interactions with nature, with many participants mentioning that they liked some aspect of the interactions with the real world. The lesson also included an activity to take a leaf from the tree, break it, and sniff it to experience the scent of eucalyptus oil. Many participants (16 out of 44) brought up this activity as one of their

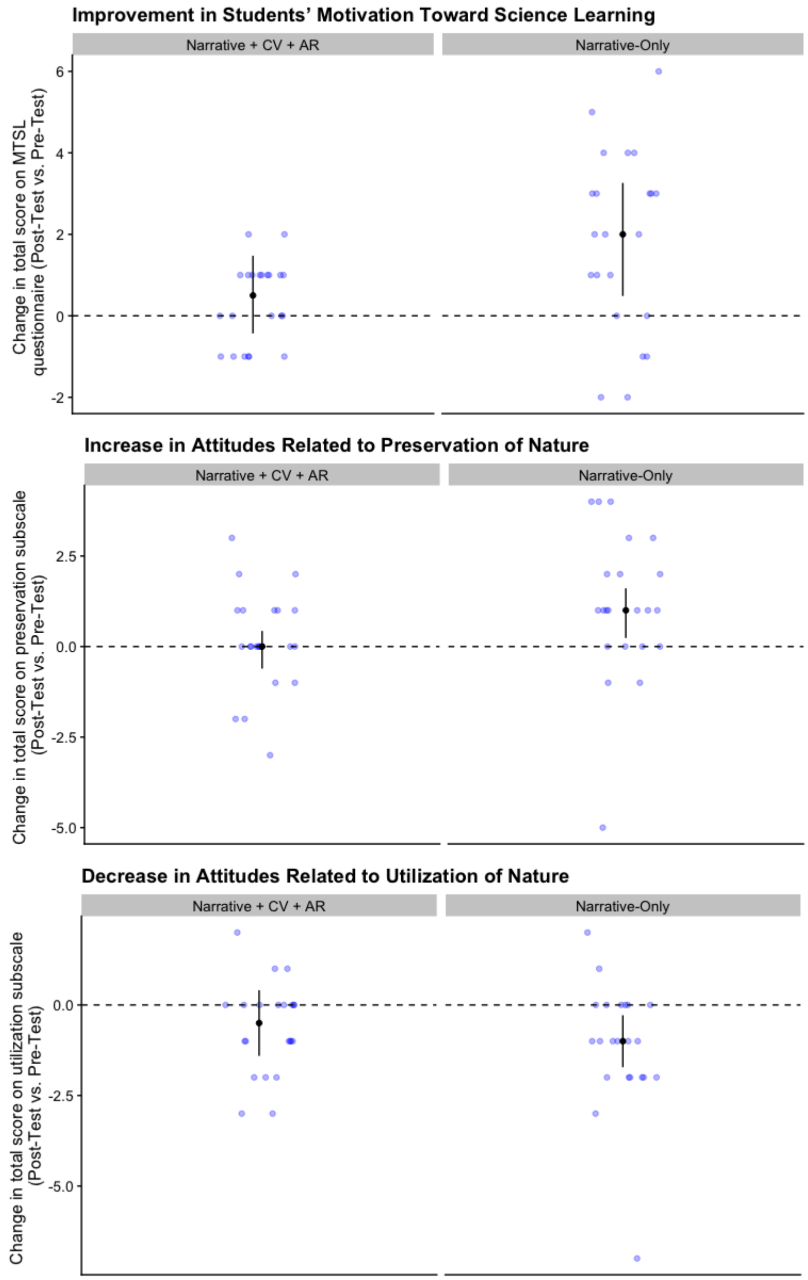


**Figure 5: Changes in motivation toward science learning and environmental attitudes following interaction with our system. Each blue point is the difference in scores from pre-to-post for a single participant, jittered along the horizontal axis to reduce overlap. Black dot shows median change, and black bars are 95% bootstrapped confidence intervals (CIs) for the median.**

favorite parts of the lesson (see Table 3). One participant (P58) compared it to their prior experiences of learning with technology, which had been sedentary and hands-off: *"My experiences with online learning, you don't really do anything, you just sit in front of a computer and answer your questions online. So it was cool to actually see the eucalyptus tree and learn about it."* Several participants (10 out of 44) even went up to the tree and rubbed the bark with their hands, unprompted. One such participant (P22) recalled, *"I liked the smell, it's cool. And like the bark, the feel of the bark, was awesome."* (P22).

*7.2.2 Participants found the locally-relevant content and activities compelling.* Many participants (12 out of 44) enjoyed learning about the history of eucalyptus trees in the local area, either because they

**Table 3: A table showing how many participants reported enjoying each activity in the interview. Some activities were only available in the Narrative+CV+AR app variant, so the counts are grouped by condition.**

| Activity | Narrative+CV+AR | Narrative-Only |
|---|---|---|
| Finding the eucalyptus tree | 2 | 5 |
| Smelling the leaf | 8 | 8 |
| Answering questions | 6 | 6 |
| Taking a picture of the tree (CV) | 4 | |
| Putting Kiki on the tree (AR) | 2 | |
| Planting the virtual eucalyptus tree (AR) | 5 | |
| Measuring the height of the tree (AR) | 6 | |

thought eucalyptus trees only existed in Australia or because they were unaware of the trees' origin: *"Because these eucalyptus trees, like I do see them often but sometimes, like you don't know where the trees come from. Or, like, do people bring them over here? [...] I've seen that type of tree before but I'd never knew it was called the eucalyptus tree."* (P26). Integrating place-based learning into our app helped engage our participants by connecting concepts and filling gaps in their knowledge. Our findings add to a large body of existing research that emphasizes the need for local outdoor learning experiences [5] and show that it is possible to engage children in outdoor learning using technology.

*7.2.3 Participants' first-hand interactions with the characters in the narrative positively impacted how they engaged with the educational content.* Participants had generally positive perceptions of the main characters of the narrative and talked about how the characters *"came to life"* (P49), *"[had] their own personalities"* (P24), and *"were really nice"* (P30). Furthermore, some participants described the educational content of the app in relation to the main characters of the narrative. For example, P48 said "*I like the way they make me to try it by myself. [...] Like, they told me to smell the eucalyptus tree leaves and I did. And I never tried that before. And they didn't just explain it, they told me to try it.*" In this excerpt, P48 describes the lesson by explaining that the koala made her try something by herself. This is distinctly different than following a prompt devoid of narrative in a textbook, something that participants appreciated about this activity.

*7.2.4 Many participants were impressed by the characters' AI.* Despite our simple keyword detection-based approach to natural language understanding, a majority of participants felt like the koalas understood what they said in the app, with 30 out of 44 indicating "Agree" or "Strongly agree" with the statement, "I felt like Kiki and Coco understood what I said." P31 exclaimed while using the app, *"The AI is pretty good!"* The AI helped make the interactions with the characters feel more personal. One participant (P33) remarked in the interview, "*They actually seemed like they were real people talking back to you, it was like a real conversation.*" Another (P32) said, *"They actually sounded like they had like, feelings about what I said. And they sometimes changed their mind."*

*7.2.5 Participants frequently mentioned responding to free-response prompts as their favorite part of the activity.* Interestingly, many participants (14 out of 44) identified free-response questions as one of their favorite features of the app. We found several reasons for this. First, they enjoyed being able to voice their own opinion. A few participants highlighted the debate specifically, such as P31, who said, "*It was kind of nice that the app, like you could agree or disagree with it on like an opinion with nature. [...] I found that really cool, actually.*" Another reason participants enjoyed the free-response feature is because it helped them to actively engage with the educational content. P40 said, "*I just, like don't like memorizing stuff, because it's hard, but like, talking to someone it makes it like, it makes me like understand and like memorize more? Yeah. Helps my memory, I guess.*" (P40). The interactivity helped participants think through the prompts, transforming knowledge into understanding. Finally, some participants enjoyed being able to record their observations about the natural world. For instance, one participant described how recording their observations made them feel as though they were conducting a science experiment: "*"It was kind of like an experiment. With technology. It was like, yeah, it was like an experiment that you could do. And it like, told you what to do. And you could take a picture, and then write down in the app, what you could, like, realize from the plan."* (P28) However, two participants (P22 and P34) said that they did not like typing and would have preferred to dictate their answers instead.

*7.2.6 Participants generally enjoyed the AR activities.* Most participants (12 out of 22) in the Narrative+CV+AR condition explicitly mentioned one of the AR activities (putting the koala on the tree, measuring the height of the tree, or planting the virtual eucalyptus) as a feature of the app that they enjoyed. P17, who had no prior experience with AR, explained their excitement about AR: *"I find it really cool how you could just put, take your phone out and just virtually just place anything there. Like you could place a giant chicken right there and it'd work. Anything anywhere. I find that really cool."* Another participant, P51, explicitly mentioned enjoying this blend between the virtual and the physical that AR affords, saying, *"[The app] kind of put like virtual reality and real life together. That was fun."*

## 8 DISCUSSION

Through the process of designing, building, piloting, iterating, and testing Kiki's Adventure, we found that our system successfully guides outdoor learning, though there were few differences between the two variants of our system (RQ1). We synthesize our findings on how children interact with our technology to uncover important design implications for creating narrative-based outdoor learning activities (RQ2 & RQ3).

## 8.1 Narrative-based technology effectively guides local, self-directed outdoor learning

We developed and tested an app that successfully engaged young learners to learn by interacting with their local environment using narrative. The majority of participants (39 out of 44) were able to complete the lesson independently, using only the narrative-based system as a guide. Although our system is grounded in a fictional narrative, its activities engage users with their real-world environments. In our design, we draw heavily from experiential learning (by having learners experience the eucalyptus tree firsthand) and from place-based learning (by emphasizing the local history and ecology of eucalyptus), which resulted in learning gains from using our system. Despite some participants initially being surprised that a mobile app was telling them to interact with the physical world, many cited the hands-on activities and learnings relevant to their local area as highlights of their experience.

Since narrative-based educational technology requires skills in storytelling, application development, and teaching, it can be more difficult to implement a narrative-based educational app compared to a purely educational one. However, *Kiki's Adventure* provides an example of how to incorporate fictional narrative into educational technology that helps users learn about their real-world environment. By showing the feasibility of using narrative-based technology to guide informal outdoor learning, we hope to encourage more designers of educational technology to explore this space.

## 8.2 Interactive characters spark curiosity

The vast majority of participants indicated they want to learn more about the story's main characters, Kiki and Coco, and go on another adventure with Kiki. This suggests that the curiosity the characters sparked engaged participants in the narrative, something that could be built on to continue delivering more educational content.

Our participants seemed to highly enjoy and benefit from the interactivity with the characters. This opens up design opportunities for creating character designs in future work that feel personally relevant for learners with different preferences, a way to make it more individualized [3], such as offering options to select characters personalities. For example, learners could choose to learn with characters that were more chatty or more active based on their preferences, or even mood for that day. Different characters could be carefully tailored to a learner's age, background, interests, and existing knowledge. Furthermore, characters could have narrative arcs that resemble the challenges a specific learner may be encountering, providing emotional support and companionship while learning.

## 8.3 Chat-based interactions motivate young learners to write

Writing is a critical skill that is generally hard to motivate children to practice [6]. Many children are behind in spelling [16], which affects their writing skills, whether typing or handwriting. Furthermore, the generation effect shows that writing can support learning in general—having learners generate their own responses helps them learn information better than simply reading it [56].

We suspected that typing in responses would be considered more work (and therefore less fun) than just tapping an option. Thus, we were surprised that so many of our participants indicated that answering the free-response questions within the app was a highlight of their experience. Based on participant feedback and behavior during user studies, we hypothesize that these factors in the design of our app were responsible:

(1) The questions were embedded within a narrative with characters that participants liked.
(2) The questions encouraged participants to engage with the real world and share their own observations. Participants frequently paused to observe the tree when typing their answers.
(3) The questions asked participants for their opinions, rather than always seeking a right or wrong answer. Instead of being either right or wrong (as can be the case with closed-ended questions), open-ended questions could be right in different ways. This could improve learners' self-efficacy.
(4) The characters gave customized feedback to the participants' responses, and the feedback was designed to be encouraging and positive. Participants may have felt that their thoughts and opinions were appreciated by the characters.

We encourage future designers of educational technology to leverage open-ended, free-response questions not just as a tool for learning, but also as a tool for engagement. We offer the above points as suggestions on how to design such questions to be motivating.

A few participants disliked the act of typing into the phone, and in general, many children struggle with writing. We suggest exploring the option of using dictation as scaffolding for writing, allowing young learners to build confidence expressing themselves and ease into writing.

## 8.4 The Narrative+CV+AR variant did not outperform the Narrative-Only one

No statistically significant advantages were observed for the Narrative+CV+AR variant over the Narrative-Only one. Moreover, science attitudes were significantly less impacted by the Narrative+CV+AR variant. However, in the pretest, the Narrative-Only group had significantly lower positive attitudes towards science. As a result, there were more opportunities for them to improve. This may have contributed to the larger improvement in science attitudes.

## 8.5 A learner's location in space must be carefully choreographed

Designers hoping to combine narrative, AR, or context-aware features with location-based technology must consider the "choreography" of the experience. When narrative is added to location-based technology, the designer needs to consider where the user should be at each point in the narrative and come up with a progression that feels seamless within the physical space. This is made even more challenging with the addition of technology that interacts with the physical world, such as camera-based image recognition and augmented reality. For example, the height measuring activity was highly inaccurate if the user was too close to the tree, due to the steep angle needed to see the top of the tree. Users needed to stand sufficiently far away, so we added a prompt in the narrative to facilitate this ("First, make sure you're standing far enough away from the tree so you can comfortably see the entire tree.").

During our pilot studies, we discovered that participants using the Narrative+CV+AR variant were struggling to find the AR koalas in the tree at the end of the lesson, even though they had placed the koalas there earlier. We realized that when participants were standing far away when the place the AR koalas on the tree. However, because the narrative required the user to go under the tree and grab a leaf, participants would often walk past the spot where they had placed the koala without realizing it.

To fix this, we reordered the activities in the narrative by moving the height measuring activity right before the end. This encouraged the user to step away from the tree, which then allows them to easily find the koala they had placed earlier. We suggest that designers and developers of similar apps test the flow of their activities at their intended locations regularly—with unfamiliar users whenever possible—to check whether it makes sense within the space. We also recommend including activities (like our height measuring activity) that can act as forcing functions to reliably constrain the user's position.

## 9 LIMITATIONS AND FUTURE WORK

Our study and system design have several limitations that can be addressed through additional research. First, novelty effects and demand characteristics may have increased the apparent effectiveness of the intervention apps. Longer-term, do-it-yourself deployments are needed to mitigate these potentially confounding factors.

Second, our participant pool was skewed toward locals who had the availability and resources to drive to our study location, and whose attitudes towards science going into the study were already enthusiastic. A worthwhile direction for future research would be to repeat the study in a different community, potentially with participants whose attitudes towards science are not enthusiastic. It would be particularly beneficial for future work to examine how this type of learning activity may benefit students who may feel discouraged or disengaged in school. Our intervention apps are different from traditional methods of teaching science, and many participants mentioned preferring our method. Therefore, our work creates a promising new option to capture the attention of those for whom traditional methods of learning science are less effective.

Third, even though there is extensive research supporting the use of narrative to increase learning outcomes [24, 44, 50], we did not explicitly test a variant without narrative. Thus, more research is needed to determine how effective the narrative itself was relative to a control.

Fourth, the AR features, while providing a level of immersion and engagement, did not contribute directly to the learning goals of the experience. In this study, we demonstrate that it is possible to make use of AR features (and also CV technology) to create engaging activities within an educational narrative, but these technologies could more directly support the educational content by taking better advantage of semantic understanding of the world. For example, learners could use AR to investigate the eucalyptus leaf at a magnified perspective and see the oil glands, or even learn geometric math by measuring shadows with AR and computing the tree height. Due to the large amount of new content being introduced, we focused on using AR to keep the narrative fun and engaging. We would like refine the AR activities and run a formal controlled study to better understand the effects of AR on outdoor learning.

Fifth, researcher intervention was required when our users behaved in ways that we did not expect. The confusion between the leaves of the eucalyptus and the small tree growing at the base of the eucalyptus is one example of a way in which our system was unable to correct users' misunderstandings. Ultimately, we wish to support self-directed learning situations where there may not be an instructor or experimenter to step in. Achieving this goal would require mechanisms to allow either the learner or the system to detect and recover from these types of identification errors. There is a large range of spatial layouts of different plant species in typical outdoor locations, in addition to a large number of actions that learners can take during an outdoor experience. If we were able to capture and model the common errors that learners make, this model could be the basis for context-aware, educational systems that are capable of diagnosing misunderstandings, which can aid learning by highlighting teachable moments.

Sixth, our system was built with a hyperlocal narrative, for a specific location, using an overfitted image classifier for real-world understanding. However, the classifier can be trained to recognize more trees in more locations. Moreover, the structure of the lesson (e.g., combining outdoor learning and narrative) can be used to teach other topics meeting other educational goals. In the future, we would like to test our system with other eucalyptus trees at other locations, as well as implement lessons for other topics in outdoor learning. We also see an opportunity to extend our work to enable computer supported collaborative inquiry learning [49] by bringing multiple learners into an outdoor space and leveraging augmented reality's ability to create persistent virtual objects to establish joint attention. These changes would allow our system to achieve opportunities of scale and reach more learners.

Seventh, we discovered that the design of our system motivated learners to write, but we could not precisely identify the exact causes that led to this. We described some hypotheses in Section 8.3, but future work could look deeper into what factors are most effective to inform the design of educational technologies that promote writing.

Finally, creating a single cohesive, narrative-based, outdoor learning experience with immersive technologies required substantial effort in terms of both design of the curriculum and content and technical implementation. We would like to explore how to enable teachers, or other instructors, to adapt our technology to various contexts around the world with different geographical features, flora, and fauna, increasing our potential for positive impact. An interesting area for future work would be to explore the development of specialized authoring tools that allow educators to easily incorporate AR and computer vision into their learning activities.

## 10 CONCLUSION

In this paper, we described a design exploration of a narrative-based app to guide outdoor learning. We evaluated two variants of our app with fourth, fifth, and sixth grade participants learning about eucalyptus trees in a local, outdoor setting. Through these variants, we explored the trade-offs of adding immersive features that made use of AR and computer vision technologies. We found that our app was effective at engaging learners and helping them learn, but we did not find large differences between the app variants. Participants demonstrated excitement for hands-on interactions with nature, curiosity about the characters, and enthusiasm toward typing their thoughts and observations. This paper offers a set of design implications for employing advances in contextual understanding and AR technology to support the design of narrative learning systems that connect educational content to the real world.

## ACKNOWLEDGMENTS

We are indebted to the Brown Institute for Media Innovation for supporting this research through a Magic Grant and providing input during development. Part of this research was funded by a grant from Meta, in collaboration with Anoop Sinha, Justin Haaheim, James Yang, Pranav Deshpande, Roman Rädle, and Somayan Chakrabarti. We gratefully acknowledge our collaborators who contributed to the design (Laura Leventhal, Yousef AbuHashem, Thanawan Atchariyachanvanit, Rowena Chen), graphics (Alexa Lempel), and voiceovers (Will Yu, Jordan Troutman) of *Kiki's Adventure*. We thank Rachel Wolf and Roy Pea for their guidance on questionnaire design and analysis. Finally, we are grateful to Tatiana Chamorro and Carlos Sebastian Rivers for their feedback on an early prototype, as well as all of the wonderful children and parents who participated in our user studies.

## REFERENCES

[1] Deanne M. Adams, Richard E. Mayer, Andrew MacNamara, Alan Koenig, and Richard Wainess. 2012. Narrative games for learning: Testing the discovery and narrative hypotheses. *Journal of Educational Psychology* 104 (2012), 235–249. https://doi.org/10.1037/a0025595

[2] Lorin W. Anderson and David R. Krathwohl (Eds.). 2001. *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives* (2 ed.). Allyn & Bacon, New York.

[3] Ronald J Anson. 1994. *Systemic Reform: Perspectives on Personalizing Education.* ERIC, Washington, D.C.

[4] Albert Bandura. 2001. Social cognitive theory: An agentic perspective. *Annual review of psychology* 52, 1 (2001), 1–26.

[5] Simon Beames, Pete Higgins, and Robbie Nicol. 2012. *Learning outside the classroom: Theory and guidelines for practice.* Routledge, New York.

[6] Roger Bruning and Christy Horn. 2000. Developing motivation to write. *Educational psychologist* 35, 1 (2000), 25–37.

[7] Sarah J Carrier, Linda P Tugurian, and Margareta M Thomson. 2013. Elementary science indoors and out: Teachers, time, and testing. *Research in Science Education* 43, 5 (2013), 2059–2083.

[8] The LIFE Center. 2005. The LIFE Center Lifelong and Lifewide Learning Diagram. http://life-slc.org/about/about.html Accessed: 2022-09-15.

[9] Yuh-Shyan Chen, Tai-Chien Kao, and Jay-Ping Sheu. 2003. A mobile learning system for scaffolding bird watching learning. *Journal of computer assisted learning* 19, 3 (2003), 347–359.

[10] M-T Cheng, H-C She, and Leonard A Annetta. 2015. Game immersion experience: its hierarchical structure and impact on game-based science learning. *Journal of computer assisted learning* 31, 3 (2015), 232–253.

[11] Yu-Cheng Chien, Yen-Ning Su, Ting-Ting Wu, and Yueh-Min Huang. 2019. Enhancing students' botanical learning by using augmented reality. *Universal Access in the Information Society* 18, 2 (2019), 231–241.

[12] Diana I. Cordova and Mark R. Lepper. 1996. Intrinsic motivation and the process of learning: Beneficial effects of contextualization, personalization, and choice. *Journal of Educational Psychology* 88 (1996), 715–730. https://doi.org/10.1037/0022-0663.88.4.715

[13] Michele D Dickey. 2006. Game design narrative for learning: Appropriating adventure game design narrative devices and techniques for the design of interactive learning environments. *Educational technology research and development* 54, 3 (2006), 245–263.

[14] Jürgen Drissner, Hans-Martin Haase, and Katrin Hille. 2010. Short-term environmental education-does it work?-An evaluation of the 'green classroom'. *Journal of Biological Education* 44, 4 (2010), 149–155.

[15] Julie Ducasse. 2020. Augmented reality for outdoor environmental education. In *Augmented Reality in Education: A New Technology for Teaching and Learning.* Springer, Switzerland, 329–352.

[16] Per Engzell, Arun Frey, and Mark D Verhagen. 2021. Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences* 118, 17 (2021), e2022376118.

[17] Noel Enyedy, Joshua A Danish, Girlie Delacruz, and Melissa Kumar. 2012. Learning physics through play in an augmented reality environment. *International journal of computer-supported collaborative learning* 7, 3 (2012), 347–378.

[18] John H Falk, Nancy Staus, Lynn D Dierking, William Penuel, Jennifer Wyld, and Deborah Bailey. 2016. Understanding youth STEM interest pathways within a single community: The Synergies project. *International Journal of Science Education, Part B* 6, 4 (2016), 369–384.

[19] Joey R. Fanfarelli. 2020. Impact of narrative and badging on learning and engagement in a psychology learning game. *British Journal of Educational Technology* 51, 2 (2020), 387–419. https://doi.org/10.1111/bjet.12838 arXiv:https://bera-journals.onlinelibrary.wiley.com/doi/pdf/10.1111/bjet.12838

[20] James Farmer, Doug Knapp, and Gregory M Benton. 2007. An elementary school environmental education field trip: Long-term effects on ecological and environmental knowledge and attitude development. *The journal of environmental education* 38, 3 (2007), 33–42.

[21] Carol Fuller, Daisy Powell, and Simon Fox. 2017. Making gains: the impact of outdoor residential experiences on students' examination grades and self-efficacy. *Educational Review* 69, 2 (2017), 232–247.

[22] Yiannis Georgiou and Eleni A Kyza. 2018. Relations between student motivation, immersion and learning outcomes in location-based augmented reality settings. *Computers in Human Behavior* 89 (2018), 173–181.

[23] Yiannis Georgiou and Eleni A Kyza. 2021. Bridging narrative and locality in mobile-based augmented reality educational activities: Effects of semantic coupling on students' immersion and learning gains. *International Journal of Human-Computer Studies* 145 (2021), 102546.

[24] Manuela Glaser, Bärbel Garsoffky, and Stephan Schwan. 2009. Narrative-based learning: Possible benefits and problems. *Communications: The European Journal of Communication Research* 34, 4 (2009), 429–447. https://doi.org/10.1515/COMM.2009.026

[25] Lynne Hall, Colette Hume, and Sarah Tazzyman. 2016. Five degrees of happiness: Effective smiley face likert scales for evaluating with children. In *Proceedings of the the 15th international conference on interaction design and children.* Association for Computing Machinery, Manchester, U.K., 311–321.

[26] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition.* IEEE, US, 770–778.

[27] Tien-Chi Huang, Chia-Chen Chen, and Yu-Wen Chou. 2016. Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. *Computers & Education* 96 (2016), 72–82.

[28] Bruce Johnson and Constantinos C Manoli. 2010. The 2-MEV scale in the United States: a measure of children's environmental attitudes based on the theory of ecological attitude. *The Journal of Environmental Education* 42, 2 (2010), 84–97.

[29] Sara Jose, Patricia G Patrick, and Christine Moseley. 2017. Experiential learning theory: the importance of outdoor classrooms in environmental education. *International Journal of Science Education, Part B* 7, 3 (2017), 269–284.

[30] Seokbin Kang, Ekta Shokeen, Virginia L Byrne, Leyla Norooz, Elizabeth Bonsignore, Caro Williams-Pierce, and Jon E Froehlich. 2020. ARMath: augmenting everyday life with math learning. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* Association for Computing Machinery, US, 1–15.

[31] Robin Kay. 2011. Evaluating learning, design, and engagement in web-based learning tools (WBLTs): The WBLT Evaluation Scale. *Computers in Human Behavior* 27, 5 (2011), 1849–1856.

[32] Alida Kossack and Franz X Bogner. 2012. How does a one-day environmental education programme support individual connectedness with nature? *Journal of Biological Education* 46, 3 (2012), 180–187.

[33] Neeraj Kumar, Peter N Belhumeur, Arijit Biswas, David W Jacobs, W John Kress, Ida C Lopez, and João VB Soares. 2012. Leafsnap: A computer vision system for automatic plant species identification. In *European conference on computer vision.* Springer, Springer, Firenze, Italy, 502–516.

[34] Secret Lab. 2022. Yarn Spinner. https://yarnspinner.dev/ Accessed: 2022-09-13.

[35] Kwan Min Lee, Namkee Park, and Seung-A Jin. 2006. Narrative and interactivity in computer games. In *Playing video games.* Routledge, US, 304–322.

[36] Agoritsa Makri, Dimitrios Vlachopoulos, and Richard A Martina. 2021. Digital escape rooms as innovative pedagogical tools in education: a systematic literature review. *Sustainability* 13, 8 (2021), 4587.
[37] Lucy R McClain and Heather Toomey Zimmerman. 2016. Technology-mediated engagement with nature: sensory and social engagement with the outdoors supported through an e-Trailguide. *International Journal of Science Education, Part B* 6, 4 (2016), 385–399.
[38] Scott W McQuiggan, Jennifer L Robison, and James C Lester. 2010. Affective transitions in narrative-centered learning environments. *Journal of Educational Technology & Society* 13, 1 (2010), 40–53.
[39] Scott W. McQuiggan, Jonathan P. Rowe, Sunyoung Lee, and James C. Lester. 2008. Story-Based Learning: The Impact of Narrative on Learning Experiences and Outcomes. In *Intelligent Tutoring Systems*, Beverley P. Woolf, Esma Aïmeur, Roger Nkambou, and Susanne Lajoie (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 530–539.
[40] Elizabeth L Murnane, Xin Jiang, Anna Kong, Michelle Park, Weili Shi, Connor Soohoo, Luke Vink, Iris Xia, Xin Yu, John Yang-Sammataro, et al. 2020. Designing ambient narrative-based interfaces to reflect and motivate physical activity. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, US, 1–14.
[41] Aquiles Negrete. 2005. *Fact via Fiction: Stories that Communicate Science*. The Pantaneto Press, UK, 95–102. https://doi.org/10.13140/RG.2.1.5110.1207
[42] Niantic. 2022. NianticLightship. https://lightship.dev/ Accessed: 2022-09-15.
[43] Danakorn Nincarean, Mohamad Bilal Alia, Noor Dayana Abdul Halim, and Mohd Hishamuddin Abdul Rahman. 2013. Mobile augmented reality: The potential for education. *Procedia-social and behavioral sciences* 103 (2013), 657–664.
[44] Elena Novak. 2015. A critical review of digital storyline-enhanced learning. *Educational Technology Research and Development* 63, 3 (2015), 431–453.
[45] Cornell Lab of Ornithology. 2014. Merlin Bird ID. http://merlin.allaboutbirds.org/ Accessed: 2022-08-26.
[46] California Academy of Sciences and National Geographic Society. 2008. iNaturalist. http://www.inaturalist.org/ Accessed: 2022-08-26.
[47] Louise E. Parker and Mark R. Lepper. 1992. Effects of fantasy contexts on children's learning and motivation: Making learning more fun. *Journal of Personality and Social Psychology* 62 (1992), 625–633. https://doi.org/10.1037/0022-3514.62.4.625
[48] Kathryn M Parsley. 2020. Plant awareness disparity: A case for renaming plant blindness. *Plants, People, Planet* 2, 6 (2020), 598–601.
[49] Roy Pea, Marcelo Milrad, Heidy Maldonado, Bahtijar Vogel, Arianit Kurti, and Daniel Spikol. 2011. Learning and technological designs for mobile science inquiry collaboratories. In *Orchestrating Inquiry Learning*. Routledge, US, 114–136.
[50] Meihua Qian and Karen R Clark. 2016. Game-based Learning and 21st century skills: A review of recent research. *Computers in human behavior* 63 (2016), 50–58.
[51] José M Rios and Jessica Brewer. 2014. Outdoor education and science achievement. *Applied Environmental Education & Communication* 13, 4 (2014), 234–240.
[52] Yvonne Rogers, Sara Price, Geraldine Fitzpatrick, Rowanne Fleck, Eric Harris, Hilary Smith, Cliff Randell, Henk Muller, Claire O'Malley, Danae Stanton, et al. 2004. Ambient wood: designing new forms of digital augmentation for learning outdoors. In *Proceedings of the 2004 conference on Interaction design and children: building a community*. Association for Computing Machinery, US, 3–10.
[53] Sherry Ruan, Jiayu He, Rui Ying, Jonathan Burkle, Dunia Hakim, Anna Wang, Yufeng Yin, Lily Zhou, Qianyao Xu, Abdallah AbuHashem, et al. 2020. Supporting children's math learning with feedback-augmented narrative technology. In *Proceedings of the Interaction Design and Children Conference*. Association for Computing Machinery, US, 567–580.
[54] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. 2015. Imagenet large scale visual recognition challenge. *International journal of computer vision* 115, 3 (2015), 211–252.
[55] Claudia Schrader and Theo J Bastiaens. 2012. The influence of virtual presence: Effects on experienced cognitive load and learning outcomes in educational computer games. *Computers in Human Behavior* 28, 2 (2012), 648–658.
[56] Daniel L Schwartz, Jessica M Tsang, and Kristen P Blair. 2016. *The ABCs of how we learn: 26 scientifically proven approaches, how they work, and when to use them*. WW Norton & Company, US.
[57] David Sobel. 2004. Place-based education: Connecting classroom and community. *Nature and listening* 4, 1 (2004), 1–7.
[58] Kurt D Squire and Mingfong Jan. 2007. Mad city mystery: Developing scientific argumentation skills with a place-based augmented reality game on handheld computers. *Journal of science education and technology* 16, 1 (2007), 5–29.
[59] Next Generation Science Standards. 2022. Next generation science standards. https://www.nextgenscience.org/
[60] Kathryn T Stevenson, Rachel E Szczytko, Sarah J Carrier, and M Nils Peterson. 2021. How outdoor science education can help girls stay engaged with science. *International Journal of Science Education* 43, 7 (2021), 1090–1111.
[61] Wernhuar Tarng, Kuo-Liang Ou, Chuan-Sheng Yu, Fong-Lu Liou, and Hsin-Hun Liou. 2015. Development of a virtual butterfly ecological system based on augmented reality and mobile learning technologies. *Virtual Reality* 19, 3 (2015), 253–266.
[62] Hsiao-Lin Tuan*, Chi-Chin Chin, and Shyang-Horng Shieh. 2005. The development of a questionnaire to measure students' motivation towards science learning. *International journal of science education* 27, 6 (2005), 639–654.
[63] Grant Van Horn, Oisin Mac Aodha, Yang Song, Yin Cui, Chen Sun, Alex Shepard, Hartwig Adam, Pietro Perona, and Serge Belongie. 2018. The inaturalist species classification and detection dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, US, 8769–8778.
[64] James H Wandersee and Elisabeth E Schussler. 1999. Preventing plant blindness. *The American biology teacher* 61, 2 (1999), 82–86.
[65] Pieter Wouters, Christof van Nimwegen, Herre van Oostendorp, and Erik D. van der Spek. 2013. A meta-analysis of the cognitive and motivational effects of serious games. *Journal of Educational Psychology* 105 (2013), 249–265. https://doi.org/10.1037/a0031311
[66] Heather Toomey Zimmerman and Susan M Land. 2014. Facilitating place-based learning in outdoor informal environments with mobile computers. *TechTrends* 58, 1 (2014), 77–83.
[67] Heather Toomey Zimmerman, Susan M Land, Chrystal Maggiore, Robert W Ashley, and Chris Millet. 2016. Designing outdoor learning spaces with iBeacons: combining place-based learning with the Internet of learning things. In *Proceedings of the 12th International Conference of the Learning Sciences*. International Society of the Learning Sciences, Singapore.
[68] Heather Toomey Zimmerman, Susan M Land, Chrystal Maggiore, and Chris Millet. 2019. Supporting children's outdoor science learning with mobile computers: integrating learning on-the-move strategies with context-sensitive computing. *Learning, Media and Technology* 44, 4 (2019), 457–472.

# A APPENDIX

**Table 4: Pre-Test to Post-Test Learning Gains**

| Item | Pre-Test Distribution | Post-Test Distribution | Gains | p |
|---|---|---|---|---|
| Why do you think Australia has a wildfire problem? | 1: 14, 2: 28, 3: 0, 4: 0 | 1: 0, 2: 3, 3: 20, 4: 19 | Median: 2, Std.Dev:0.71 | 0.000000014 |
| What would you tell Lulu about this tree to help her identify its species? | 1: 4, 2: 8, 3: 19, 4: 11 | 1: 3, 2: 5, 3: 13, 4: 21 | Median: 0, Std.Dev:0.58 | 0.000643560 |
| Suppose your friend is going camping in a forest with eucalyptus trees in it. What would you want them to know? | 1: 27, 2: 11, 3: 2, 4: 2 | 1: 1, 2: 6, 3: 23, 4: 12 | Median: 2, Std.Dev:0.89 | 0.000000066 |
| List two reasons why eucalyptus trees catch fire easily. | 1: 35, 2: 7, 3: 0, 4: 0 | 1: 0, 2: 5, 3: 18, 4: 19 | Median: 2, Std.Dev:0.70 | 0.000000009 |
| Click any picture with a eucalyptus tree. | Median: 0, Std.Dev:1.19 | Median: 2, Std.Dev:0.59 | Median: 1, Std.Dev:1.27 | 0.000003902 |
| How did eucalyptus trees come to be in California? | Correct: 8,Incorrect:36 | Correct:41,Incorrect: 3 | Median: 1, Std.Dev:0.44 | 9.79e-09 |

**Table 5: Pre-Test to Delayed Post-Test Learning Gains**

| Item | Pre-Test Distribution | Delayed Post-Test | Gains | p |
|---|---|---|---|---|
| Why do you think Australia has a wildfire problem? | 1: 14, 2: 28, 3: 0, 4: 0 | 1: 0, 2: 7, 3: 19, 4: 9 | Median: 1, Std.Dev:0.88 | 0.000001042 |
| What would you tell your friend about this tree to help her identify its species? | 1: 4, 2: 8, 3: 19, 4: 11 | 1: 0, 2: 4, 3: 14, 4: 17 | Median: 0, Std.Dev:0.66 | 0.000274529 |
| Suppose your friend is going camping in a forest with eucalyptus trees in it. What would you want them to know? | 1: 27, 2: 11, 3: 2, 4: 2 | 1: 2, 2: 3, 3: 22, 4: 8 | Median: 2, Std.Dev:1.03 | 0.000001741 |
| List two reasons why eucalyptus trees catch fire easily. | 1: 35, 2: 7, 3: 0, 4: 0 | 1: 0, 2: 6, 3: 8, 4: 21 | Median: 2, Std.Dev:0.77 | 0.000000162 |
| Click any picture with a eucalyptus tree. | Median: 0, Std.Dev:1.19 | Median: 2, Std.Dev:0.73 | Median: 1, Std.Dev:1.40 | 0.000122680 |
| How did eucalyptus trees come to be in California? | Median: 0, Std.Dev:0.41 | Correct:20,Incorrect: 5 | Median: 1, Std.Dev:0.50 | 0.0001227 |

**Table 6: Results for Student's Motivation Towards Science Learning. Range of possible values: [6,30]**

| Pre-test vs. post-test scores for motivation towards science learning | | | | | |
|---|---|---|---|---|---|
| | Median (pre-test) | Median (post-test) | V | Z | p |
| Overall | 22 | 22 | 571 | -3.37 | **0.00075** |
| Narrative-only | 20.5 | 22 | 188 | -3.10 | **0.002** |
| Narrative+CV+AR | 22 | 22 | 105 | -1.44 | 0.15 |
| Amount of change in motivation towards science learning between the two variants | | | | | |
| | Median (Narrative-only) | Median (Narrative+CV+AR) | W | Z | p |
| | 2 | 0.5 | 131.5 | -2.63 | **0.0086** |
| Do pre-test scores for motivation towards science learning differ by variant? | | | | | |
| | Median (Narrative-only) | Median (Narrative+CV+AR) | W | Z | p |
| | 20.5 | 22 | 335.5 | -2.21 | **0.027** |

**Table 7: Results for Attitudes Towards Technology**

| Pre-test vs. post-test scores for attitudes towards technology. Range of possible values: [6,30] | | | | | |
|---|---|---|---|---|---|
| | Median (pre-test) | Median (post-test) | V | Z | p |
| Overall | 23 | 23.5 | 370.5 | -0.91 | 0.369 |
| Amount of change in attitudes towards technology between the two variants | | | | | |
| | Median (Narrative-only) | Median (Narrative+CV+AR) | W | Z | p |
| | 1 | 0 | 199.5 | -1.0 | 0.32 |

**Table 8: Results for Preservation of Nature Subscale. Range of possible values: [6,30]**

| Pre-test vs. post-test scores for preservation of nature | | | | | |
|---|---|---|---|---|---|
| | Median (pre-test) | Median (post-test) | V | Z | p |
| Overall | 22 | 22 | 365.5 | -2.33 | **0.02** |
| Narrative-only | 21 | 22.5 | 143 | -2.52 | **0.012** |
| Narrative+CV+AR | 22 | 22 | 51.5 | -0.4 | 0.69 |
| Do pre-test scores for preservation of nature differ by variant? | | | | | |
| | Median (Narrative-only) | Median (Narrative+CV+AR) | W | Z | p |
| | 21 | 22 | 284 | -.98 | 0.33 |

**Table 9: Results for Utilization of Nature Subscale. Range of possible values: [4,20]**

| Pre-test vs. post-test scores for utilization of nature | | | | | |
|---|---|---|---|---|---|
| | Median (pre-test) | Median (post-test) | V | Z | p |
| Overall | 10 | 9 | 66 | -3.5 | **0.00049** |
| Narrative-only | 9 | 7.5 | 15 | -2.8 | **0.0057** |
| Narrative+CV+AR | 10.5 | 9 | 19.5 | -2.1 | **0.037** |
| Do pre-test scores for utilization differ by variant? | | | | | |
| | Median (Narrative-only) | Median (Narrative+CV+AR) | W | Z | p |
| | 9 | 10.5 | 287.5 | -1.1 | 0.29 |

**Table 10: Post-Test Only Questions**

| Item | Narrative-Only | Narrative+CV+AR |
|---|---|---|
| I found the app engaging. | SA: 5, A:15, NO: 2, D: 0, SD: 0 | SA: 7, A:10, NO: 4, D: 1, SD: 0 |
| The app made learning fun. | SA: 9, A:10, NO: 3, D: 0, SD: 0 | SA: 8, A: 8, NO: 5, D: 1, SD: 0 |
| I would like to use the app again. | SA: 6, A:11, NO: 5, D: 0, SD: 0 | SA: 6, A: 8, NO: 7, D: 0, SD: 1 |
| I want to learn more about the koalas (Kiki and Coco). | SA: 4, A:14, NO: 4, D: 0, SD: 0 | SA: 4, A:13, NO: 4, D: 0, SD: 1 |
| I felt like Kiki and Coco understand what I said. | SA: 5, A:11, NO: 4, D: 2, SD: 0 | SA: 6, A: 8, NO: 6, D: 1, SD: 1 |
| I want to go on another adventure with Kiki. | SA: 5, A:13, NO: 4, D: 0, SD: 0 | SA: 7, A: 7, NO: 6, D: 1, SD: 1 |
| Working with the app helped me learn. | SA:10, A:11, NO: 1, D: 0, SD: 0 | SA:13, A: 8, NO: 1, D: 0, SD: 0 |
| The koalas' responses to my answers in the app helped me learn. | SA: 6, A:11, NO: 5, D: 0, SD: 0 | SA: 6, A:12, NO: 3, D: 1, SD: 0 |
| The graphics from the app helped me learn. | SA: 4, A: 6, NO: 6, D: 5, SD: 1 | SA: 6, A:10, NO: 2, D: 3, SD: 1 |
| The app helped teach me a new concept. | SA:11, A: 9, NO: 2, D: 0, SD: 0 | SA: 8, A:13, NO: 0, D: 1, SD: 0 |
| Overall the app helped me learn. | SA:15, A: 7, NO: 0, D: 0, SD: 0 | SA:12, A: 9, NO: 0, D: 1, SD: 0 |
| The instructions in the app were easy to follow. | SA:12, A:10, NO: 0, D: 0, SD: 0 | SA: 8, A:14, NO: 0, D: 0, SD: 0 |
| The app was easy to use. | SA:11, A:11, NO: 0, D: 0, SD: 0 | SA:11, A: 9, NO: 2, D: 0, SD: 0 |
| The app was well organized. | SA: 8, A:14, NO: 0, D: 0, SD: 0 | SA: 9, A:11, NO: 2, D: 0, SD: 0 |
| I liked the overall theme of the app. | SA: 7, A:10, NO: 4, D: 1, SD: 0 | SA: 7, A: 8, NO: 6, D: 1, SD: 0 |

**Table 11: Pre-Test Attitude Questions**

| Item | Narrative-Only | Narrative+CV+AR |
|---|---|---|
| I am sure that I can do well on science tests | SA: 0, A: 8, NO:10, D: 4, SD: 0 | SA: 0, A:13, NO: 4, D: 5, SD: 0 |
| When I do not understand things in science I still try to learn them | SA: 2, A:14, NO: 6, D: 0, SD: 0 | SA: 4, A:15, NO: 3, D: 0, SD: 0 |
| I think that learning science is important because I can use it in my daily life. | SA: 4, A:13, NO: 3, D: 1, SD: 1 | SA: 7, A: 9, NO: 6, D: 0, SD: 0 |
| I am not sure that I can understand difficult science ideas (*reverse scored) | SA: 0, A: 3, NO:10, D: 6, SD: 3 | SA: 1, A: 7, NO: 9, D: 4, SD: 1 |
| When I make a mistake I try to understand why. | SA: 3, A:12, NO: 4, D: 2, SD: 1 | SA: 3, A:17, NO: 2, D: 0, SD: 0 |
| I think that learning science is important because it makes me think more. | SA: 5, A: 7, NO:10, D: 0, SD: 0 | SA: 5, A:10, NO: 7, D: 0, SD: 0 |
| When I do not understand something in science I use technology to find things that will help me. | SA: 2, A:14, NO: 3, D: 3, SD: 0 | SA: 4, A:12, NO: 5, D: 0, SD: 1 |
| I'm not good at learning science even with the help of technology. (*reverse scored) | SA: 4, A:12, NO: 6, D: 0, SD: 0 | SA: 0, A:13, NO: 5, D: 4, SD: 0 |
| I prefer to use technology to learn science rather than textbooks. | SA: 4, A: 6, NO: 9, D: 3, SD: 0 | SA: 2, A: 6, NO:11, D: 2, SD: 1 |
| I don't think technology is useful for learning science (*reverse scored) | SA:10, A: 8, NO: 3, D: 1, SD: 0 | SA: 8, A: 9, NO: 3, D: 1, SD: 1 |
| I use technology to explore and learn new science ideas. | SA: 1, A:11, NO: 9, D: 1, SD: 0 | SA: 4, A:11, NO: 5, D: 2, SD: 0 |
| I feel comfortable using technology to help me learn difficult science ideas. | SA: 6, A:11, NO: 5, D: 0, SD: 0 | SA: 6, A:14, NO: 1, D: 0, SD: 1 |
| I would help raise money to protect nature. | SA: 4, A:12, NO: 5, D: 1, SD: 0 | SA: 6, A: 8, NO: 6, D: 2, SD: 0 |
| I always turn off the light when I do not need it any more. | SA: 3, A:10, NO: 5, D: 4, SD: 0 | SA: 4, A:13, NO: 3, D: 2, SD: 0 |
| I would like to sit by a pond and watch dragonflies. | SA: 2, A: 7, NO: 6, D: 6, SD: 1 | SA: 3, A:10, NO: 5, D: 2, SD: 2 |
| Weeds should be killed because they take up space from plants we need. | SA: 4, A: 4, NO: 7, D: 5, SD: 2 | SA: 1, A:10, NO: 9, D: 2, SD: 0 |
| Building new roads is so important that trees should be cut down. | SA: 1, A: 0, NO: 3, D: 7, SD:11 | SA: 0, A: 0, NO: 0, D:13, SD: 9 |
| Because mosquitoes live in swamps we should drain the swamps and use the land for farming. | SA: 0, A: 1, NO: 3, D:12, SD: 6 | SA: 1, A: 1, NO: 6, D: 8, SD: 6 |
| I try to save water by taking shorter showers or by turning off the water when I brush my teeth. | SA: 7, A: 9, NO: 5, D: 1, SD: 0 | SA: 4, A:15, NO: 2, D: 1, SD: 0 |
| I like a grass lawn more than a place where flowers grow on their own. | SA: 1, A: 2, NO: 7, D: 9, SD: 3 | SA: 2, A: 2, NO: 7, D:10, SD: 1 |
| I like to go on trips to places like forests away from cities. | SA: 4, A: 6, NO: 4, D: 6, SD: 2 | SA: 5, A:10, NO: 4, D: 3, SD: 0 |
| I try to tell others that nature is important. | SA: 3, A: 9, NO: 9, D: 1, SD: 0 | SA: 3, A: 8, NO: 9, D: 2, SD: 0 |

**Table 12: Post-Test Attitude Questions**

| Item | Narrative-Only | Narrative+CV+AR |
|---|---|---|
| I am sure that I can do well on science tests | SA: 1, A:11, NO: 9, D: 1, SD: 0 | SA: 1, A:13, NO: 5, D: 3, SD: 0 |
| When I do not understand things in science I still try to learn them | SA: 4, A:13, NO: 5, D: 0, SD: 0 | SA: 6, A:12, NO: 4, D: 0, SD: 0 |
| I think that learning science is important because I can use it in my daily life. | SA: 6, A:13, NO: 3, D: 0, SD: 0 | SA: 7, A: 9, NO: 6, D: 0, SD: 0 |
| I am not sure that I can understand difficult science ideas (*reverse scored) | SA: 1, A: 8, NO:11, D: 2, SD: 0 | SA: 1, A: 8, NO: 6, D: 7, SD: 0 |
| When I make a mistake I try to understand why. | SA: 4, A:14, NO: 2, D: 2, SD: 0 | SA: 6, A:14, NO: 2, D: 0, SD: 0 |
| I think that learning science is important because it makes me think more. | SA: 5, A: 8, NO: 8, D: 1, SD: 0 | SA: 4, A:11, NO: 7, D: 0, SD: 0 |
| When I do not understand something in science I use technology to find things that will help me. | SA: 3, A:16, NO: 2, D: 1, SD: 0 | SA: 3, A:13, NO: 4, D: 1, SD: 1 |
| I'm not good at learning science even with the help of technology. (*reverse scored) | SA: 5, A:11, NO: 4, D: 2, SD: 0 | SA: 1, A:13, NO: 3, D: 4, SD: 1 |
| I prefer to use technology to learn science rather than textbooks. | SA: 4, A:10, NO: 5, D: 3, SD: 0 | SA: 4, A: 8, NO: 7, D: 2, SD: 1 |
| I don't think technology is useful for learning science (*reverse scored) | SA: 6, A:11, NO: 4, D: 1, SD: 0 | SA: 5, A:11, NO: 3, D: 2, SD: 1 |
| I use technology to explore and learn new science ideas. | SA: 4, A:13, NO: 4, D: 1, SD: 0 | SA: 4, A:15, NO: 1, D: 2, SD: 0 |
| I feel comfortable using technology to help me learn difficult science ideas. | SA: 5, A:17, NO: 0, D: 0, SD: 0 | SA: 5, A:13, NO: 2, D: 2, SD: 0 |
| I would help raise money to protect nature. | SA: 6, A:11, NO: 5, D: 0, SD: 0 | SA: 6, A: 7, NO: 7, D: 2, SD: 0 |
| I always turn off the light when I do not need it any more. | SA: 5, A:13, NO: 4, D: 0, SD: 0 | SA: 6, A:10, NO: 4, D: 2, SD: 0 |
| I would like to sit by a pond and watch dragonflies. | SA: 5, A: 4, NO: 8, D: 4, SD: 1 | SA: 3, A:10, NO: 4, D: 5, SD: 0 |
| Weeds should be killed because they take up space from plants we need. | SA: 1, A: 5, NO: 5, D: 5, SD: 6 | SA: 1, A: 7, NO:10, D: 3, SD: 1 |
| Building new roads is so important that trees should be cut down. | SA: 1, A: 0, NO: 1, D: 9, SD:11 | SA: 0, A: 0, NO: 1, D: 8, SD:13 |
| Because mosquitoes live in swamps we should drain the swamps and use the land for farming. | SA: 1, A: 0, NO: 3, D: 7, SD:11 | SA: 1, A: 0, NO: 2, D:13, SD: 6 |
| I try to save water by taking shorter showers or by turning off the water when I brush my teeth. | SA: 6, A:12, NO: 3, D: 1, SD: 0 | SA: 7, A:11, NO: 3, D: 1, SD: 0 |
| I like a grass lawn more than a place where flowers grow on their own. | SA: 0, A: 3, NO: 4, D:10, SD: 5 | SA: 3, A: 0, NO: 9, D: 9, SD: 1 |
| I like to go on trips to places like forests away from cities. | SA: 5, A: 6, NO: 5, D: 2, SD: 4 | SA: 4, A:10, NO: 7, D: 1, SD: 0 |
| I try to tell others that nature is important. | SA: 4, A: 9, NO: 8, D: 0, SD: 1 | SA: 3, A: 8, NO: 9, D: 2, SD: 0 |

**Table 13: Pre-Test to Post-Test Changes for Attitude Questions**

| Item | Narrative-Only | Narrative+CV+AR |
|---|---|---|
| I am sure that I can do well on science tests | Median: 0, Mean:0.36, Std.Dev:0.79 | Median: 0, Mean:0.18, Std.Dev:0.79 |
| When I do not understand things in science I still try to learn them | Median: 0, Mean:0.14, Std.Dev:0.56 | Median: 0, Mean:0.05, Std.Dev:0.56 |
| I think that learning science is important because I can use it in my daily life. | Median: 0, Mean:0.32, Std.Dev:0.95 | Median: 0, Mean:0.00, Std.Dev:0.95 |
| I am not sure that I can understand difficult science ideas (*reverse scored) | Median:-0, Mean:-0.77, Std.Dev:1.02 | Median: 0, Mean:0.00, Std.Dev:1.02 |
| When I make a mistake I try to understand why. | Median: 0, Mean:0.27, Std.Dev:0.77 | Median: 0, Mean:0.14, Std.Dev:0.77 |
| I think that learning science is important because it makes me think more. | Median: 0, Mean:0.00, Std.Dev:0.93 | Median: 0, Mean:-0.05, Std.Dev:0.93 |
| When I do not understand something in science I use technology to find things that will help me. | Median: 0, Mean:0.27, Std.Dev:0.77 | Median: 0, Mean:-0.09, Std.Dev:0.77 |
| I'm not good at learning science even with the help of technology. (*reverse scored) | Median: 0, Mean:0.05, Std.Dev:1.00 | Median: 0, Mean:0.00, Std.Dev:1.00 |
| I prefer to use technology to learn science rather than textbooks. | Median: 0, Mean:0.18, Std.Dev:0.85 | Median: 0, Mean:0.27, Std.Dev:0.85 |
| I don't think technology is useful for learning science (*reverse scored) | Median: 0, Mean:0.23, Std.Dev:0.81 | Median: 0, Mean:0.23, Std.Dev:0.81 |
| I use technology to explore and learn new science ideas. | Median: 0, Mean:0.36, Std.Dev:0.58 | Median: 0, Mean:0.18, Std.Dev:0.58 |
| I feel comfortable using technology to help me learn difficult science ideas. | Median: 0, Mean:0.18, Std.Dev:0.73 | Median: 0, Mean:-0.14, Std.Dev:0.73 |
| I would help raise money to protect nature. | Median: 0, Mean:0.18, Std.Dev:0.50 | Median: 0, Mean:-0.05, Std.Dev:0.50 |
| I always turn off the light when I do not need it any more. | Median: 0, Mean:0.50, Std.Dev:0.80 | Median: 0, Mean:0.05, Std.Dev:0.80 |
| I would like to sit by a pond and watch dragonflies. | Median: 0, Mean:0.23, Std.Dev:0.53 | Median: 0, Mean:0.05, Std.Dev:0.53 |
| Weeds should be killed because they take up space from plants we need. | Median: 0, Mean:-0.59, Std.Dev:1.01 | Median: 0, Mean:-0.27, Std.Dev:1.01 |
| Building new roads is so important that trees should be cut down. | Median: 0, Mean:-0.09, Std.Dev:0.53 | Median: 0, Mean:-0.14, Std.Dev:0.53 |
| Because mosquitoes live in swamps we should drain the swamps and use the land for farming. | Median: 0, Mean:-0.18, Std.Dev:1.10 | Median: 0, Mean:-0.27, Std.Dev:1.10 |
| I try to save water by taking shorter showers or by turning off the water when I brush my teeth. | Median: 0, Mean:0.05, Std.Dev:0.79 | Median: 0, Mean:0.09, Std.Dev:0.79 |
| I like a grass lawn more than a place where flowers grow on their own. | Median: 0, Mean:-0.27, Std.Dev:0.63 | Median: 0, Mean:0.05, Std.Dev:0.63 |
| I like to go on trips to places like forests away from cities. | Median: 0, Mean:0.09, Std.Dev:0.61 | Median: 0, Mean:0.00, Std.Dev:0.61 |
| I try to tell others that nature is important. | Median: 0, Mean:0.05, Std.Dev:0.49 | Median: 0, Mean:0.00, Std.Dev:0.49 |