Engaging with Nature Sounds & Citizen Science by Designing for Creative & Contextual Audio Encounters

Jessica L. Oliver
Queensland University of Technology, jessie.l.oliver@gmail.com

Selen Turkay
Queensland University of Technology, selen.turkay@qut.edu.au

Margot Brereton
Queensland University of Technology, m.brereton@qut.edu.au

David M. Watson
Charles Sturt University, dwatson@csu.edu.au

Paul Roe
Queensland University of Technology, p.roe@qut.edu.au

Wildlife calls are the best witnesses to the health of ecosystems, if only we know how to listen to them. Efforts to understand and inform restoration of healthy ecosystems with environmental audio recordings languish from insufficient tools to learn and identify sounds in recordings. To address this problem, we designed and playtested the Bristle Whistle Challenge prototype with ten players. We explored how to design delightful interactions with audio for gaining awareness of nature sounds and supporting wildlife conservation through citizen science. We found that rather than presenting audio alone, it was necessary to connect sounds to other senses and experiences in creative ways to impart meaning and enhance engagement. We offer recommendations to design creative and contextual interactions with media to build awareness of nature's wonders. We call for greater efforts in interaction design to engage people with nature, which is the key to turning around our environmental crisis.

CCS CONCEPTS • Human-centered computing • Human-computer interaction (HCI)

Additional Keywords and Phrases: Agent, Audio, Birds, Citizen Science, Creativity, Emotion, Games, Games with a Purpose, Learning, Ludic Design, Nature, Participatory Design, Sound, Wildlife

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template: This is the subtitle of the paper, this document both explains and embodies the submission format for authors using Word. In Woodstock '18: ACM Symposium on Neural Gaze Detection, June 03–05, 2018, Woodstock, NY. ACM, New York, NY, USA, 10 pages. NOTE: This block will be automatically generated when manuscripts are processed after acceptance.

1 INTRODUCTION

The sounds of nature often go unnoticed in the everyday lives of most people. Interactions with audio are generally understudied in HCI, though there has been a recent surge of interest in designing interactions with calls of the wild [6, 41, 42, 60, 67]. Interactions with playfully designed devices that include environmental audio recordings have shown promise to spark interest in nature [4, 14, 57, 60]. Research has also explored how expert birdwatchers, or birders, identify birds in the wild by sight and sound, as well as when listening to audio recordings. Such studies have provided insights for developing systems to support environmental monitoring with audio recordings [10, 11, 66]. Studies have also highlighted difficulties with auditorily and visually interpreting audio recordings [7, 13, 40, 41]. For example, visualizations of sound, called spectrograms, can be difficult to interpret [13, 41]. However, with some training,

spectrograms can improve the interpretation of audio [61, 65]. We need to understand how to design enjoyable and rewarding experiences, so people can explore, learn about, and study the sounds in nature recordings.

With advances in technology, people can now capture short environmental audio recordings with mobile devices [55] or place acoustic sensors in environments for extended periods [30]. These recordings can potentially support finding and monitoring secretive and rare wildlife that may otherwise go undetected, such as little-known, endangered Eastern bristlebirds [63]. However, some species, like bristlebirds, have calls that are understudied and highly variable. This makes automatic detection of calls with computers difficult until a set of accurately identified calls is available to train algorithms [26, 42]. Having enticing designs that support learning calls may offer community members (i.e. as citizen scientists) and ecologists new opportunities to forge collaborative scientific investigations with audio that support saving species [15, 42].

For this exploratory study, we aimed to investigate how audio interactions can entice curiosity about nature and foster engagement with citizen science that supports wildlife conservation. Our two primary objectives were to 1) design a prototype that includes diverse interactions with audio recordings, and 2) observe how people interacted with this prototype. We were guided by two research questions (RQs):

1. Can technology support people to increase their interest in wildlife and their calls? If so, what forms might this technology take?
2. Can technology engender skills that increase engagement in the study of wildlife? If so, how can this inform the design of future citizen science technologies?

To explore these questions, we created the Bristle Whistle Challenge, a gameful prototype (hereafter called "the Challenge"; Figure 1). We focused the Challenge on Eastern bristlebirds because of our experience with the species [42] and the popularity of environmentally-focused citizen science [34, 46]. A narrative introduced bristlebirds and their conservation needs. Additionally, the narrative explained how people's experiences while playing the Challenge may inform the design of technologies that support learning and identifying bristlebird calls.

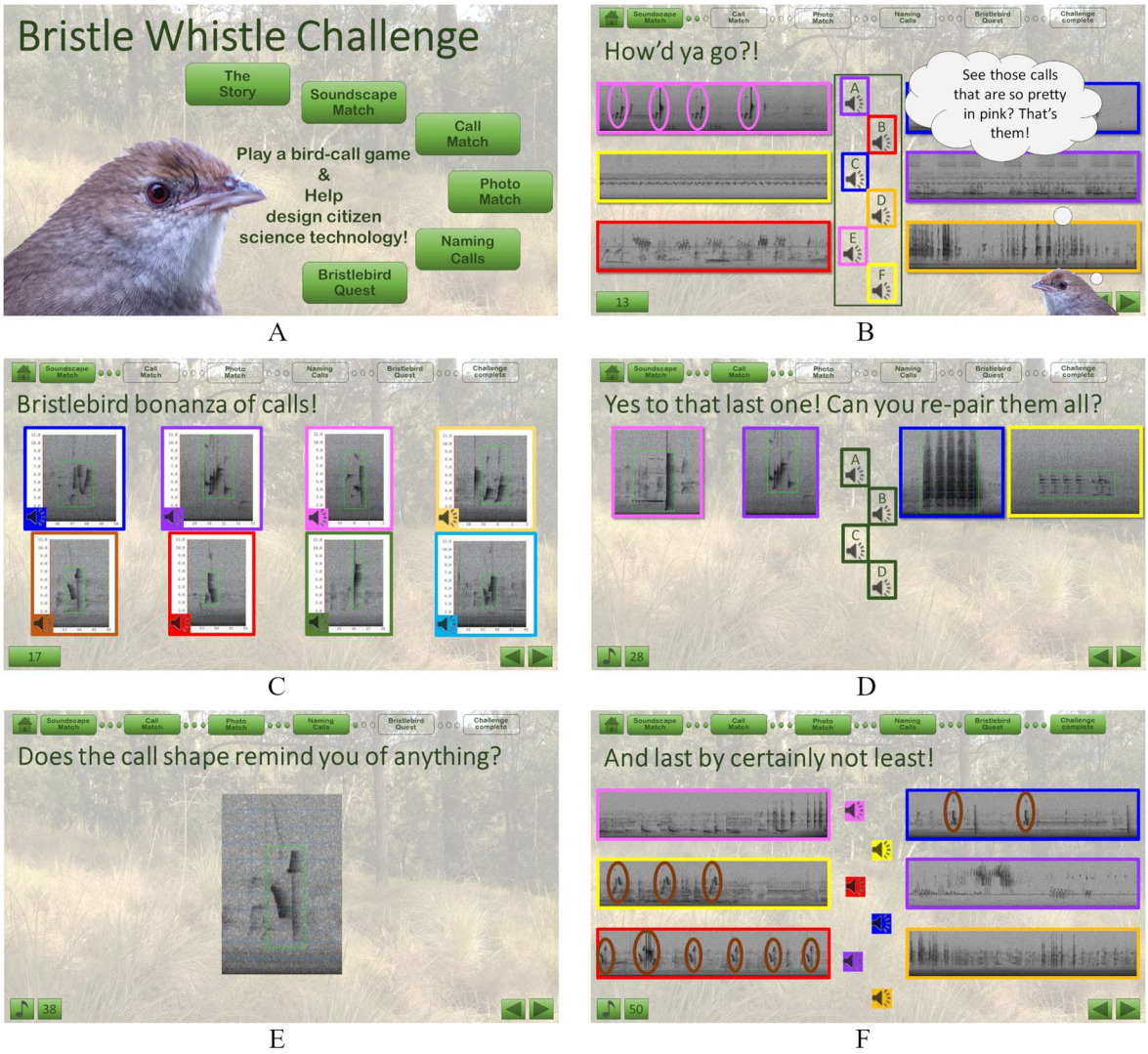

Figure 1: Screenshots of most Bristle Whistle Challenge sections. Players progressed, in sequence, to: A) view an activity overview and introductory narrative; B) gamefully match soundscape spectrograms with audio recordings; C) playful explore bristlebird calls; D) gamefully match bird call spectrograms with audio recordings, and then match this media with bird photos; E) playfully name calls by shape and then sound; F) gamefully identify bristlebird calls in soundscapes (see Supplementary Video).

We incorporated game elements within a non-game context (i.e. *gameful design* or *gamification* [17]) to explore the design of enjoyable experiences with audio recordings. Designing for enjoyment often involves *gameful*, goal-oriented tasks, or more free-form *playful* exploratory activities [17] . We designed the Challenge for people to receive guidance as they progressed through a mixture of gameful and playful activities. We set out to understand how each interaction may support enjoyment and learning (Table 1). Given the overall gamefulness of the Challenge prototype, we refer to participants as players and sessions as *playtests* [22].

Table 1: Research goals for interactions with the Challenge

Interactions (section)	To determine whether:
Receiving Guidance (3.1)	A gameful personified conversational agent provides useful support for audio interpretation
Matching Media (3.2)	Gameful tasks matching soundscapes*, calls, & photos support making sense of bird media
Exploring Media (3.2)	Playful interactions with birds by sight & sound or naming calls elicit interest & creativity
Questing for Calls (3.4)	Prior sections prove useful to gamefully identify bristlebird calls in soundscapes
Enjoying Experience (3.5)	Elements are engaging (e.g. narrative, agent, play, game, aesthetics, & feedback)

* Soundscape = 30-second environmental audio recordings capturing sounds of a landscape

Playtesting was conducted with 10 players to answer our RQs. We found that playful interactions with the conversational agent "Bob Bristlebird" supported players emotionally and logistically, while also making the game enjoyable (see Table 3). Each player readily applied their newly acquired experiences of interpreting audio to the final task of identifying bristlebird calls in soundscapes. Overall, players found the experience enjoyable, despite the cognitive demands of soundscape matching. Players engaged with multisensory experiences enthusiastically, which is promising for the design of future playful interactions with environmental audio.

This work contributes to the HCI and citizen science communities by providing a new understanding of how mixing supportive, creative, and structured interactions with multisensory media can afford sense-making of complex information, such as the interplay between how birds look, sound and live. Our research extends knowledge in understudied areas of HCI, including audio-visual interactions, playfulness in citizen science, gamefulness in training, and nature engagement. We offer recommendations to design creative and contextual interactions with media to build awareness of nature sounds. We call for greater efforts in interaction design to engage people with nature, to improve environmental awareness, ecological research with citizen science, and wildlife conservation.

2 LITERATURE REVIEW

2.1 Games & Citizen Science

Games are generally played for entertainment and include structured, goal-oriented tasks that often have competitive elements [17]. Games are also increasingly created for purposes beyond just entertainment, such as for use in education, health, and science [52, 68]. Researchers, for example, have explored games that engage people with sound (i.e. primarily music) while collecting data on their perceptions and annotations with the intent to inform audio searches [2, 19, 33, 36-38]. Similarly, citizen scientists are commonly recruited to complete a variety of gameful tasks, which provides researchers with data needed to answer scientific questions [43, 47, 48].

Tasks in most citizen science projects are designed to be easy, requiring little prior knowledge or training to participate [54]. However, identification of wildlife calls can vary in difficulty depending on the context of the species being sought after and the soundscape. For instance, the low frequency, consistent calls of a White-bellied heron are more distinct than the versatile calls of Eastern bristlebirds [13]. More people may explore wildlife calls when designs are available that improve call accessibility and offer enjoyable learning experiences.

2.2 Play to Understand Nature

> "*… Play is not just mindless entertainment, but an essential way of engaging with and learning about our world and ourselves — for adults as well as children*" [23, p. 2].

In design, play usually refers to creating freeform, exploratory, improvisational experiences [17]. Some view play and games on opposite ends of a spectrum, though people do regularly shift into playful modes of thinking while playing games [17]. Most technology is created with a utilitarian, task-oriented purpose, but there is value in designing playful explorations that tap into our inclinations to be curious, explore, and reflect [23, 25]. Gaver proposes that designing for activities to elicit curiosity, exploration, and reflection without a defined task "*can be a mechanism for developing new values and goals, for learning new things, and for achieving new understandings*" [25, p. 886].

Most citizen science is contributory, being led by researchers (typically scientists), with the public contributing data [50, 59]. Additionally, many design studies on citizen science are reflexive, exploring factors like participant motivations, engagement rates, gameful design, and data quality, to maximize scientific outcomes [12, 21, 48, 64]. Relatively few citizen science technologies are designed with a participatory design approach, gearing towards empowering communities, rather than controlling engagement [50]. We take inspiration from those who have designed to foster holistic experiences with nature that evoke playful curiosity, learning, sharing, and desires to help nature [18, 58]. Since learning animal calls is daunting, we are exploring how to design engaging experiences with audio that make animal calls more accessible to learn through guided playful and gameful activities.

3 DESIGNING THE CHALLENGE

The Challenge design is a culmination of our data corpus exploring how to design enticing nature engagement and citizen science with audio recordings to inform wildlife conservation[6, 40-42, 60]. We created the Challenge using Microsoft PowerPoint because the software was likely familiar to players. The Challenge consisted of 61 slides, embedded with diverse combinations of media, animations, and clickable buttons in arrangements aimed to be intuitive and aesthetically cheerful. Players read text, considered agent comments, played audio recordings, interpreted spectrograms, interrogated photos, and received feedback on answers. Playtests were designed to last 1.5 hours, assuming players would engage with all information presented, describe impressions and decisions, and reflect on their overall experience upon session completion. In this section, design goals, rationale, and mechanics for key interactions in the Challenge are explained (see Table 1 for goals). Interactions are grouped by categories of player actions and experiences, including *Receiving Guidance*, *Gamefully Matching Media*, *Playfully Exploring Media*, *Gamefully Questing for Calls*, and *Enjoying Experience (*Supplementary Video).

3.1 Receiving Guidance

We designed light-hearted Bob Bristlebird to explore whether the use of a playful conversational agent supports players to make sense of spectrogram variables, soundscapes, sound events (e.g. bird calls), and bird photographs (Figure 1B). Conversational agents can support people with online learning [44]. However, agents need to be designed to appear without interrupting activity flow, or being presumptuous of human intention and feelings about interventions [56] Given this, we designed Bob to pop-up and provide guidance and encouragement, usually between activity transitions. Bob also prompted players to reflect on media in the Challenge, to reveal insights to researchers, such as experiences players may have with bird species featured in the Challenge.

3.2 Gamefully Matching Media

Matching tasks were designed to elucidate *how* players make sense of soundscapes, specific sounds, and bird photos (Figure 1B, D, & F). We were interested in whether gameful aspects of memory recall, scaffolding, repetition, multisensory perception, and feedback were useful and enjoyable. Upon completing the *Soundscape Match* section, players progressed to a playful section to explore bird media (i.e. Bioacoustics Bonanza, see below). Then, players progressed to gamefully *Call Match* and *Photo Match* sections (Supplementary Video).

3.2.1Soundscape Match

We designed the *Soundscape Match* section to engage players with gamefully matching 30-second soundscape audio recordings with spectrograms (Figure 1B). During preliminary prototyping, we observed that matching pairs of soundscape spectrograms and audio recordings were most difficult for people when multiple soundscapes had equivalent complexity or spacing of prominent sound events, such as bird calls. Given this, we designed three *Soundscape Match* levels to become progressively more difficult, by increasing:

1) The number of new spectrogram-audio pairs offered (i.e. 2, 4, and then 6 new pairs)
2) Complexity similarity across soundscapes; and
3) The similarity of call spacing within and across soundscapes.

In Level 1, players listened to a soundscape audio recording and then were asked to choose which one of two spectrograms matched their memory of the audio. This introduced listening, recalling memories of sounds, and interpreting spectrograms. Then, players were tasked with matching two audio recordings to the spectrograms (Figure 2). After players verbalized decisions, answers were revealed with borders of audio icons changing to colors that match corresponding spectrogram borders (e.g. compare Figure 1B & D audio icons). Level 2 tasked players with matching four new soundscapes, and the two original ones again (Figures 1B & 2). Soundscapes varied in complexity, with unevenly spaced sounds, such as cricket chirps, car engines, dogs barking, and birds calling (Figure 1B). In Level 3, six new soundscapes with similar complexity were presented. We strategically selected soundscapes that featured relatively evenly spaced Eastern bristlebird calls, with the intent to maximize pattern recognition difficulty.

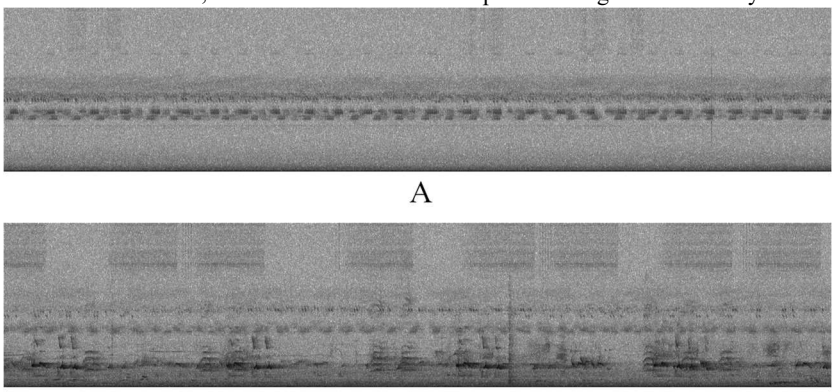

Figure 2: The first two soundscape spectrograms in the Challenge. The middle band is crickets calling at night. B) The top band is likely cicada insects, the middle is crickets, and the bottom half is birds calling (x = 30 seconds; y = 0 to 11kHz frequency).

3.2.2Call & Photo Matches

We designed the *Call Match* section to explore whether evoking player prior knowledge, memory recall, and multisensory perception supports noticing differences in how bird calls look and sound. Four calls from four species were used. First, two spectrograms were presented, with Bob encouraging players to interpret calls vocally (e.g. by whistling). Then, four audio recordings were provided, and players were tasked to determine, from memory, which audio recordings matched the two previous spectrograms. Players were then tasked to match additional call spectrograms and audio, with opportunities to confirm their answers (Figure 1D).

In the Photo Match, we looked to assess if players knew four species by sight or sound. Four photographs were provided to discuss. The species in photos were also featured in the Call Match. We were interested if players would make this connection. To provide them with opportunities to make this connection, the call spectrograms were provided again, followed by audio, which allowed us to assess prior knowledge and reinforce experiences from previous sections.

3.3 Playfully Exploring Media

We designed media exploration activities to consider whether playful interactions with audio that aren't goal-oriented would be engaging for players and elicit curiosity that may help them make sense of audio recordings (Figure 1C & E; Supplementary Video).

3.3.1Bioacoustics Bonanza

With the *Bioacoustics Bonanza* section, we set out to explore whether players enjoyed engaging in a freeform exploration of bird media between gameful tasks. Additionally, we were interested to know if such opportunities proved useful in later activities (Figure 1). First, eight different bristlebird call spectrograms and audio recordings were

presented to players, giving them opportunities to hear and see how much bristlebird calls can vary (Figure 1C). Then, players were provided information about four additional bird species (i.e. one photo, call spectrogram, and call audio per species). Players received prompts to notice a few common and nuanced features that can help distinguish bristlebird calls from other species with similar calls.

3.3.2Naming Calls

In a prior study, one participant commented that the call shape of the Sulphur-crested cockatoo looked like "*... a highway, or car tracks*" [40, p. 1690]. This led us to consider how people might creatively name the shapes of calls in spectrograms. We were also inspired by the *Rorschach test* [53], a psychological test that examines the shapes people perceive in the random or ambiguous visual pattern of inkblot plates (an example of *pareidolia* [39]).

To explore this further, we posted a 30-second soundscape spectrogram containing several identical bristlebird calls on social media and invited participants to name the call based on its shape "*a little like a Rorschach test*". One participant downloaded the spectrogram, cropped a single call (Figure 3A), drew a shape over the top, and shared the sketch indicating it was an "emu" (Figure 3B). Another participant provided two suggestions, drawing "a swimmer" (Figure 3C) and superimposing an "Enterprise" spaceship image on a cropped call spectrogram. With this positive outcome, we decided to include freeform *pareidolia* exploration in the forthcoming Challenge.

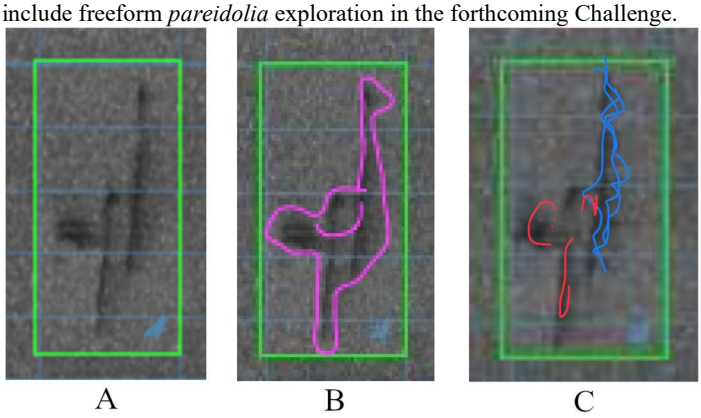

Figure 3: Calls creatively named by their shape in a spectrogram. An annotation of a bristlebird call that is: A) unaltered, B) drawn on and tagged "the emu", and C) drawn on and tagged "a swimmer".

While *pareidolia* typically refers to visual perception, we were also interested in extending this to consider whether there was an auditory equivalent of perceiving a specific sound in a random or ambiguous auditory pattern, such as a bird call. To ensure clarity between the two forms of perception, we will hereafter specify naming calls in spectrograms as "*visual pareidolia*" and naming calls by sound as "*auditory pareidolia*".

For the Challenge, we designed the *Naming Calls* section to explore whether playfully naming calls by sight and sound elicits creativity. Additionally, we sought to know whether named calls are then more recognizable and memorable. After completing the gameful *Photo Match*, Bob briefly introduced players to the idea of naming calls, relating the activity to people seeing shapes in Rorschach inkblots. Players were offered spectrograms of three different bristlebird calls and asked if the call shapes reminded them of anything (Figure 1E). Once finished interpreting calls by shapes, players advanced to naming each of the same calls again, this time metaphorically by the audible sound.

3.4 Gamefully Questing for Calls

The *Call Quest* was designed to explore how players applied experiences from previous sections (Figure 1F). Players were given two 30-second soundscape spectrograms and corresponding audio recordings. They were asked to identify familiar calls. Next, photos of three species were presented, and their calls in spectrograms were revealed, to reinforce awareness of calls from multiple species , including bristlebird. To explore if players recognized calls from previous

sections, they were then tasked to identify bristlebird calls from six pairs of soundscape spectrograms and audio recordings. Four pairs were new, whereas two were from the previous level. Upon completing the section, players were given the option to complete two bonus levels of the Challenge or begin a reflective interview.

3.5 Enjoying Experience

All Challenge interactions were created to provide players with an enjoyable experience engaging with audio visually and auditorily. We reflected on our playtest observations to identify how people felt while completing activities. Also, we considered which experiences were most memorable for players, positive or negative, based on what they discussed in reflexive interviews afterwards.

4 PLAYTESTING THE CHALLENGE

As design researchers, we were interested in how players experienced gameful and playful interactions with audio, rather than evaluating, in a hypothesis-driven sense, whether players engaged and completed activities "successfully". Instead, we adopted an exploratory approach to assessment, which focuses, *"… on capturing a rich description of the encounters people have with the prototypes, addressing their activities, experiences, and interpretations from a variety of perspectives…*". [24, p. 128]. We aimed to develop a depth of understanding, and while not generalizable, readers can ascertain whether our findings apply to their respective contexts.

To do this, we recruited a total of ten participants by *purposive sampling*, with the intent to identify unique and rich experiences of value [20]. All participants were recruited for interests in design, citizen science, or wildlife. We included five people with design expertise given their experience in navigating *wicked problems* [5], by skillfully articulating, ideating, and evaluating technology interactions. The remaining five participants were members of the public, to gain diverse perspectives and experiences.

As sessions began, we asked players about their experiences with design, citizen science, birds, and audio. Additionally, we noted relevant experiences revealed during playtests (Table 2). To gain insights into player cognitive processes and experiences during playtesting, we asked players to verbalize actions, decisions, and feelings following *concurrent think-aloud protocols* [1] . Upon completing the Challenge, we asked players reflexive questions following *retrospective think-aloud protocols* [1], to 1) triangulate observations; 2) understand what was most memorable (i.e. positive or negative); and 3) invite additional ideas for overcoming barriers, increasing fun, improving learning, or facilitating cooperation. Session observational notes and audio transcripts were reviewed to identify patterns in player experiences with Challenge interactions and to understand overall experiences.

Table 2: Participant backgrounds, interests, and experiences revealed during sessions

Identifier*	Relevant Experience & Interests
DR1	HCI senior lecturer & researcher focused on user interface design and evaluation
DR2	HCI professor, instructor, & researcher with diverse areas of study (e.g. digital interactions with nature)
DR3	HCI lecturer and researcher focused on games and learning
DR4	Designer & software developer who completed a PhD in HCI on audio interactions & birdwatchers. Has tried citizen science projects involving classification of galaxy photos online & submitting nature photos.
DR5	Industry project manager overseeing design & development of software & websites (e.g. for citizen science initiatives). Has contributed wildlife sightings as a citizen scientist and played in a hobbyist band.
PR1	Citizen science enthusiast who moderates social media for a national organization. Has tried a diverse range of projects too. Has played saxophone in a twenty-piece big band for more than 25 years.
PR2	Interested in learning about birds (e.g. recruitment from visiting bird conservation booth at a music festival)
PR3	Citizen science practitioner for 5 years, developing, supporting, and participating in projects and global working groups. Has aided in development of interfaces and platforms. Played trombone for nearly 10 years, being "a sound nerd ever since" via sound design and data sonification. Interested in soundscape ecology.

PRC1 Interested in citizen science from learning of it during University studies in ecology (e.g. a symposium led by the lead author). Aspires to learn more about birds (i.e. owns an identification field guide).

PRC2 Broadly aware of the term citizen science and became interested in this research from brother

***Identifier acronyms**: **DR** = design recruit; **PR** = public recruit; **PRC** = public recruit cooperative. DRs and PRs completed playtests individually; whereas PRC1 brought his sister to his playtest session, who was also interested in wildlife, so we welcomed the serendipitous opportunity to observe them cooperatively playtest the Challenge.

5 FINDINGS

Overall, players enjoyed all Challenge interactions, and the skills gained were useful to identify bird calls (Table 3). This section discusses how players engaged with each activity.

Table 3: Summary of findings from player experiences with the Challenge

Interactions (section)	Interactions led to players:
Receiving Guidance (5.1)	Being supported emotionally & logistically by Bob Bristlebird, a conversational agent
Matching Media (5.2)	Perceiving soundscapes in varied ways, homing in on details, & applying prior experiences
Exploring Media (5.3)	Playing with media for insights & imagining creative call names, often with personal meaning
Questing for Calls (5.4)	Applying experiences with Bob, gameful, & playful interactions to find bristlebird calls
Enjoying Experience (5.5)	Enjoying soundscapes despite being taxing & engaging enthusiastically with all sections

5.1 Receiving Guidance

5.1.1 Being Supported Emotionally & Logistically by Bob

Bob Bristlebird, the agent, provided players with emotional and logistical support. Players enjoyed seeing Bob, regularly referring to him and his dialog in endearing terms, such as being "*cute*". For instance, after completing the most difficult *Soundscape Match* level deciphering different bristlebird calls, Bob popped up and said, "*That was tricky huh?! I would make a wicked DJ the way I also remix all of these songs!*", which both reassured and amused most players after a taxing section. Additionally, players typically followed Bob's recommendations. When starting the *Call Match*, for example, Bob said, "*You might realize what the calls are if you sound them out aloud* ☺", as two call spectrograms appeared without audio, and players who hadn't tried this on their own took the advice. PR3 elaborated he thought including the emoji was a good idea, as it might help more self-conscious players be more comfortable with the idea of interpreting a spectrogram verbally. Players also engaged when Bob prompted them to notice details. When exploring *Bioacoustic Bonanza*, for example, Bob alerted players that bristlebird and Australian King Parrot calls "*can be pretty easily mixed up*", which led DR2 to spend time enthusiastically comparing those calls. Such observations highlight how an agent can playfully provide entertainment, encouragement, and guidance.

5.2 Gamefully Matching Media

5.2.1 Employing Diverse Approaches to Matching Soundscapes

Each player approached deciphering patterns in spectrograms and audio recordings based on soundscape features and personal preferences, with no two players using identical strategies. Common actions included:

- Comparing overall soundscape differences in complexity & regularity
- Comparing overall soundscape patterns in musical terms (beats, rhythms, tempo, etc.)
- Comparing temporally when sounds began or ended in soundscapes
- Comparing locations of specific familiar sounds within a soundscape
- Comparing features of specific calls to musical terms (pitch, crescendo, etc.)
- Counting the number of calls
- Counting seconds between calls

- Whistling or onomatopoetically verbalizing calls

When deciphering the first set of six soundscapes that include diverse environmental sounds (i.e. level 2), players often began by identifying broad differences, and then gradually homed in on more nuanced details of difficult pairs (Figure 1B). When matching the next set of six soundscapes that included evenly spaced bristlebird calls (i.e. level 3), players often zeroed in on features and differences of specific calls. Illuminating exemplar strategies for interpreting soundscape spectrograms and audio recordings, as well as processes for deciphering between pairs, may make audio recordings less abstract and more appealing to explore.

5.2.2Homing in on Finer Details of Specific Birds & Their Calls

Once finished exploring *Bioacoustics Bonanza* freely, players progressed to the gameful *Call Match* section, by first enthusiastically attempting to whistle or verbalize the first two call spectrograms provided (i.e. with pink and blue borders in Figure 1D). Remembering call matches proved less difficult than remembering soundscape matches, though people often replayed audio for both when answers were revealed to confirm their recollection of their answers. When provided all four call spectrograms and audio recordings, most players were able to correctly match them, regardless of whether they knew any of the bird species calling (Figure 1D).

Once transitioning to the *Photo Match* section, only the loud, ubiquitous cockatoo was regularly recognized. Some of the players thought they recognized other bird species in photos, but then described conflicting attributes that revealed players were thinking of a different species. Some players playfully personified species they did not recognize, such as DR3 and PR1 independently referring to the Eastern whipbird as, "*the angry-looking bird*". Players without prior knowledge, could not match photos with previous call spectrograms and audio recordings when given again. Those with prior knowledge who had correctly matched call audio recordings and spectrograms, however, were sometimes then able to match the correct photo. For example, DR5 did this while also telling stories about hearing whipbirds camping. Being new to Brisbane, DR3 was unfamiliar with the birds. She described that matching call audio recordings and spectrograms as easier than matching calls to bird photos. DR3 elaborated that patterns in the calls could be recognized and linked by sight and sound without experience with birds, but a bird's appearance couldn't be intuitively associated to a call without prior knowledge.

5.2.3Applying Prior Experiences to Sensemaking

Players were recruited based on having experiences and interests in either technology design, nature, or citizen science. Initial interview questions and playtest sessions, however, revealed that players were often interested in more than one of these topics, and had broader relevant experiences, such as sound designing, computer science with audio, and playing music (Table 1). Several players, for example, related visuals, soundscape audio, and bird calls to experiences with music. For instance, DR2 shared that when previously trying to distinguish calls of two bird species, her musician son introduced her to thinking in musical terms, saying *"[…] he described it in quite an interesting way, and it helped me to learn the difference*". Additionally, PR1 attempted to distinguish differences in spectrograms, saying, "*I will put this in music language because that's how I understand it for myself*". When deciphering soundscapes, PR1 described, *"counting it out like beats to a bar".* She likened spectrograms to being analogous to a "treble clef" type of "stave" if lines were added. Additional music terminology she applied included, "*emphasis mark*", "*gliss*", types of "*notes*" ("*grace*", "*high*", "*low*", "*quaver*", "*triple forte*"), "*strokes*" (up and down), "*timbre*", and "*upswing*". Similarly, DR5 also described observed patterns in musical terms. After completing the most difficult level of the *Soundscape Match*, containing audio with only evenly spaced bristlebird calls, for example, DR5 justified mismatches made, saying "*to my credit, there's some pretty even beats and it's the same bird*". Later, while matching call spectrograms and audio, he also commented, "*[…] the whipbird was a bit easier too because it kind of has got that crescendo […]*". His description referred to the male bird's characteristic call that begins as a gradually increasing whistle that reaches a final, loud cracking whip sound. Those with prior experience interacting with audio, music or otherwise, appeared to have benefited from having an established language that perhaps made thinking about and discussing sounds of nature more accessible and engaging.

5.3 Playfully Exploring Media

5.3.1 Broadening Contextual Knowledge to Decipher Calls

In the *Bioacoustics Bonanza* section, most players openly explored the photo and call media provided. However, as with the *Soundscape Match*, the amount of time and effort each player spent interrogating nuanced audible and visible differences between calls varied (Figure 1C). For instance, DR2 took time to play each call at least twice. She attempted to verbalize and whistle every call, switching between pages multiple times. DR2 carefully considered call attributes, after reading comments and "dialog" between Bob and a bird in a photograph. Others questioned why this information was provided and anticipated being tested. All players broadly commented on some aspect of the section being interesting or useful to better understand birds and calls.

5.3.2 Imagining Visual Pareidolia that Often Reflected Personal Connection

Players readily provided names for call shape, which reflected visual pareidolia. Ideas were often personalized, reflecting specific knowledge, interests, and past experiences. Some players had exhibited a propensity for visual pareidolia during the *Spectrogram Match*, likening call shapes to objects, such as chairs (DR4), broken bones (DR3), and mushrooms (DR3 for two different calls). During the *Naming Calls* section, after Bob primed players with a Rorschach inkblot image, responses varied. Most players readily offered name ideas based on shapes in spectrograms for each of the three calls provided. Two exceptions included one player naming only two of the calls, and another player not imagining names based on shape for any of the three call spectrograms provided.

Collectively, some ideas were similar, while others were distinct. Multiple players likened, for instance, the first call to a mushroom (PR1, PR3, PRC2 and previously DR3) or a lamp (DR4, DR1; Figure 4A); the second to a jellyfish (PRC1, DR4, DR2; Figure 4B); and the third to steps (DR4, PRC1, PR1) or stairs (PR1, DR1; Figure 4C). Some ideas were of a similar vein, such as the second call resembling plant-like forms, including a kelp forest and grasses (DR2); a pine plantation and bamboo (PR1); and a forest of thin eucalypts (PRC1 elaborated from PRC2's suggestion of trees; Figure 4B). Other ideas were also unique, such as the shape of the second call also evoking ideas of a window with rods (DR1); a sideways squeegee for cleaning windows (DR3); windshield wipers (DR5); blinds on the side of a house (PR1); electric cables with an electricity charge (DR2); the mythological creature "*Slenderman*" and a barcode (PRC2); and the Star Wars Imperial droids specifically in The Empire Strikes Back movie (PR3; Figure 4B).

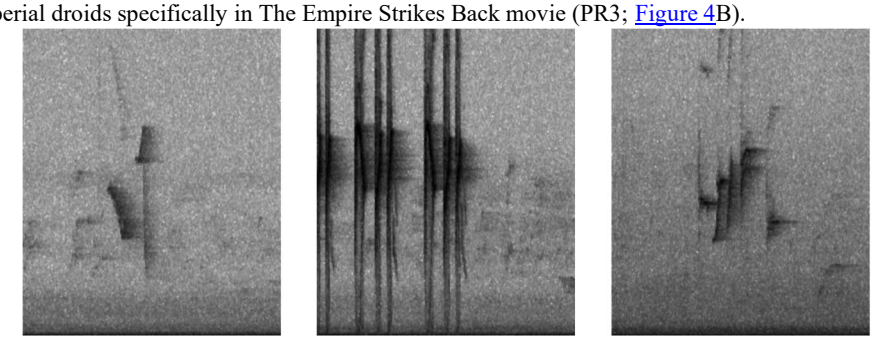

Figure 4: Spectrograms of bristlebird calls that players were invited to name creatively by shape and then by sound.

Ideas by players often included reflections of knowledge and personal stories. For instance, DR3 described the first call as looking like specific rock formations in Cappadocia, Turkey, known as fairy chimneys, which she knew of with Turkey being her country of origin (Figure 4A). Similarly, the same call evoked DR2's memories of living abroad in Japan, and she likened the shape to her recollection of the Chinese character for water, which she then confirmed via her nearby book of Chinese logograms (Figure 4A).

Some players asked us months later, using their call names, whether we found any bristlebird calls in audio recordings from the wild. For instance, DR5 asked whether we found any "*Washington Monuments*", which was his name for the first call (Figure 4A). This demonstrates that players applying personal meanings of experiences and knowledge to calls through creating names holds promise to strengthen call relatability, recognizability, and memorability.

5.3.3 Imagining Auditory Pareidolia that Often Reflected Personal Connection

Players readily created auditory pareidolia manifestations relating to calls via onomatopoeias, phrases, emotional tones, and everyday sounds. Such ideas sometimes also reflected personal experiences, interests, and auditory awareness. A few players had a propensity to mimic sounds encountered in audio by whistling and creating verbal onomatopoeias before beginning the *Naming Call* activity (DR2, DR4). Most attempted creating verbalizations of calls only after being prompted by Bob to do so (DR1, PR3). Most players found the idea of naming calls based on a likeness to another sound ambiguous and requested examples, which the interviewer provided. Players regularly shared sentiments that they only heard bird calls, lacked ideas, or were not creative (PR3, PRC2, ST, DR1, PR2). Regardless of initial apprehensions, all players interpreted at least one call auditorily, often offering multiple ideas per call.

Some players related calls to existing onomatopoeias and human language. For instance, DR2 verbalized the first call to be akin to an "achoo!" sneezing sound (Figure 4A). Similarly, DR3 described the second call resembling, "*psh, psh, psh, psh, psh*", as a shooting sound effect in a video game (Figure 4B). Players sometimes likened calls to English words, such as the first call being akin to, "*Here I am! Here I am*" (PR1), "*I like chips*" (DR4), "*What's up?*" (PRC2), and "*far out*" (PRC1; Figure 4A).

Some players translated calls by tone, anthropomorphically relating calls to a human emotional state. For instance, the second call conjured ideas of the bird having a "*real pissed off*" (PRC2), aggressive, or persistent attitude, due to the mixed registers (PR3) and short, sharp staccato notes (DR5; Figure 4B). Words applied to the second call likewise, often tonally matched being aggressive or persistent. Examples include PRC1 saying, "*danger here! danger here!*" after being startled by the intense call, "*Julie Shut up!*" (PRC2), "*Go away, get out of here*" (DR5), or akin to seagulls in the animated Disney movie *Finding Nemo* incessantly saying "mine" (PR1; Figure 4B). By contrast, DR2 related the same call to have a secretive tone, and sounded reminiscent of "*Pst, come here*" (Figure 4B).

Occasionally, players related what they heard to familiar, everyday sounds or memories. To PR2, the second call was unpleasant and like someone erratically scratching a polystyrene box (Figure 4B). The third call was reminiscent of '*I want more tea please*' for DR2 (Figure 4C). She then shared memories of her and her mum learning together years earlier that a local bird of the United Kingdom named a yellowhammer has a call reminiscent of the phrase, "*a little bit of bread and no cheese*". This story exemplifies that becoming familiar with bird calls by similar-sounding human speech resonates as an engaging activity that can help with remembering calls. Auditory pareidolia shows similar promise as visual pareidolia for bird vocalizations, with players applying meaning to calls that have the potential to support them becoming auditorily more relatable, recognizable, and memorable.

5.4 Gamefully Questing for Calls

5.4.1 Finding Bristlebird Calls by Sensemaking with Previous Challenge Experiences

All prior sections provided players playful and gameful experiences that were then applicable to the gameful task of identifying bristlebird and other calls in soundscapes. Bob's guidance through each of the previous sections assisted players in developing a process for identifying vocalizations. Players applied strategies gained through experiences gamefully matching spectrograms, calls, and photos; playfully exploring bird media; and creatively naming calls. Often players began to search for bristlebird calls by applying a broad process of elimination that was commonly adopted when matching soundscape audio recordings and spectrograms previously. Once soundscapes that seemed to lack bristlebirds were ruled out, players would begin homing in on specific calls. This meant that players then attended to determining what features make a call unique, as a common discrimination strategy. Those players who commonly verbalized calls onomatopoeically or via whistling continued to do so during *Call Questing*. Additionally, players often explored how calls in the soundscapes related to their personalized names created. Players related their visual pareidolia names more often than auditory ones. The playtest session highlighted the value in providing players with a diverse mixture of agent

guided playful and gameful interactions that engaged them with learning a suite of processes to identify specific sounds from environmental audio recordings.

5.5 Enjoying Experience

5.5.1 Gaining Interest in Bristlebirds

The narrative of the Challenge prototype focused on understanding design needs to support the use of audio for bristlebird conservation. The focus on bristlebirds piqued player interest in learning more about the secretive species. Playtesting confirmed that no player knew of Eastern bristlebirds by sight or call. For instance, when Bob inquired if any bristlebird calls were heard or seen in the second level of the Soundscape Match, no players identified the planted bristlebird calls, beyond DR3 randomly guessing correctly (i.e. calls within pink circles in Figure 1B). Most players became curious to know more about the species than what was revealed explicitly in the Challenge sections. For example, some inquired about the bristlebird's eponymous feather bristles at the top of its beak (PR3); how audio recordings of the bird were collected (DR1); the meaning of calls, and whether players may encounter the birds in the wild (PR2). The reactions of players to the narrative of the bristlebird story show promise for interest to be sparked to learn more and potentially take action to help the species.

5.5.2 Matching Soundscapes was Taxing, Albeit Engaging

Overall, the Soundscape Match was enjoyable, despite being the most cognitively demanding section of the Challenge. Players enjoyed how they were introduced to audio interpretation and other game mechanics, though the lack of explicit direction led to some confusion. For example, not telling players that the initial habitat photo was just for broad context caused some to search the image expecting to find a bird (DR1). Players also indicated wanting advanced warning that they would need to remember the audio they heard to match it by memory to a spectrogram later in the Challenge. While players readily matched the first two soundscapes (Figure 2), some did not recognize that Level 2 repeated soundscapes from Level 1 (Figure 1B). Additionally, some players found the jump from two soundscapes to six a bit overwhelming (DR3).

Given our use of PowerPoint, players were required to remember simultaneously which audio recordings had been played, which spectrograms had been matched, and which media needed further interrogation. Having to retain such information proved an arduous and cognitively demanding task that caused temporary frustration for some players. This may, in part, explain why a few took a rather haphazard approach to matching tasks. Bob's sassy comments helped to lighten the mood and caused most players to laugh, helping to alleviate tension from cognitively demanding interactions. Some players, such as DR2 and PR1, however, took delight in the difficulty of tasks, happily replaying calls several times when matches were forgotten and to resolve initial ambiguity. PR1 shared this natural inclination, saying, "*I'm gonna play that again and see if I'm right because as a musician, I don't play one thing once. I play it like 20 or 30 different times*".

Overall, the gestures and comments of players during gameplay indicated they were engaged, irrespective of any temporary frustration. Upon completing matches, all players sought to be reminded of their decided pairs, so they could determine if the feedback received matched their responses. All players were eager to know how they did with each gameful activity level, indicating overall enjoyment and that the task enticed them to compete with themselves. During reflections, players regularly stated finding that the task was fun. Additionally, they readily suggested useful ideas to reduce the memory burden, such as to drag and drop media, draw lines, or tick boxes. They also suggested other ways to elucidate associations between soundscape spectrograms and audio, which further demonstrated their interest in the experience. The *Soundscape Matching* section was likely most difficult in part because it was also the first section introducing players to game mechanisms, as well as crossmodal perception of audio by sight and sound simultaneously. Playtests showed promise that orienting people to the sounds of the wild with supportive guidance may help to abate feelings of frustration while learning a difficult task.

5.5.3 Feeling Enthusiastic About the Mixture of Gameful & Playful Multisensory Activities

As with the *Soundscape Match*, players exhibited and expressed enjoyment for other activities, often offering ideas to improve and extend interactions. Designers liked the gameful design of matching spectrograms and audio recordings in a

scaffolded way, and then searching for bristlebirds, with feedback received at every level (Figure 1B, D, & F). While all broadly enjoyed exploring materials in the *Bioacoustics Bonanza* section, some players (DR2; PR3) were enthralled with exploring and critically thinking about "dialog" between Bob and other birds describing call features. Several players also enthusiastically opted to complete optional bonus *Soundscape Match* and *Questing for Calls* sections.

While matching spectrograms with audio recordings could easily be done without prior experience, as DR3 pointed out, matching the birds by what they looked like to their calls was impossible without prior knowledge or training. We consciously designed the section this way to assess prior knowledge, but since players were not told this explicitly, the feedback was useful to validate our assumption. DR3 elaborated that despite being unable to accurately match photos to calls, she enjoyed seeing the birds and imagining their personalities based on appearance. She also suggested design ideas for restructuring the Challenge to create more opportunities for learning about how the appearances of birds may be related to their calls.

Collectively, players enjoyed the opportunity to be creative by imagining names for calls (Figure 4). Though preferences varied for naming calls by sight or sound. Most players were confused when first asked "*What does it sound like this bristlebird is saying?"*. To clarify, the researcher described a range of example mnemonic devices commonly used by birders, including creating *birdie talk* [9] (i.e. verbally mimicking calls as some players had already naturally been doing), relating the calls to human words, and likening calls to everyday sounds. Some players suggested that prompts would be helpful, with PR2 suggesting Bob Bristlebird could provide this information. Once given guidance, most players enjoyed naming each call by sound and expressed a curiosity for how others named calls. When naming calls together, PRC1 and PRC2, sometimes reached an agreement. Other times, they teasingly joked about one having a better idea than the other, hinting at a light-hearted competitive nature. PRC1 later reflected, "*I am quite competitive as a person*". He then described enjoying being able to compare his answers and name interpretations with PRC2. Then recalling a particular moment in completing the *Questing for Calls section*, PRC2 said to his sister, who did not identify as a birder, "*I was jealous of your ability to be like, 'oh, that's a noisy miner', and you knew exactly what it was, it's impressive*". Our playtests show promise that combining playful, gameful, and social interactions may create an enjoyable experience exploring environmental audio recordings. Additionally, the study shows how people's feelings about certain tasks can be wide-ranging, albeit providing a variety of tasks can provide an overall engaging experience.

6 DISCUSSION

In recent years, Preece [48, 49] highlighted opportunities for HCI researchers to advance the design for nature engagement and saving species with citizen science. This study builds upon this call to action by exploring design with environmental audio recordings to inspire more people to take notice of nature's sounds and support wildlife conservation. Our study revealed that delightful interactions with environmental audio recordings hold promise to pique people's interest in learning more about wildlife (RQ1 & Table 3). Engaging our players with environmental information through both creative and goal-oriented audio encounters enticed players to be curious about a little-known bird species and nature. Additionally, our findings show promise that such designs can engender skills likely to support people in the study of wildlife (RQ2; Table 3). While we did not quantitatively evaluate skill increase, our detailed qualitative investigation of players' interactions with the Challenge suggests activities were fun and useful. This was particularly evident when reflecting on how players applied experiences earlier in the Challenge playtest to the final task of identifying bristlebird calls in soundscapes.

However, this is just one piece of a very large design puzzle. We still need to understand how design can support people with meaningful and engaging ways to explore, study, and save the wild. Looking forward, we reflect on how players engaged with each type of interaction. We provide recommendations for designing to engage people with nature and species conservation using audio recordings (Table 4). Additionally, we frame the remainder of the discussion with key questions worth further investigation.

Table 4: Recommendations from interactions with the Challenge

Interactions (section)	To support fun experiences learning sounds of the wild, consider including:

Receiving Guidance (5.1)	Fun guidance for mechanics, sensemaking, & emotional support (e.g. with a playful agent)
Matching Media (5.2)	Gameful, scaffolded media with diverse examples to gradually home in attention on calls
Exploring Media (5.3)	Playful sections that offer freeform exploration of context and elicit multisensory creativity
Questing for Calls (5.4)	A mixture of guidance, gameful, & playful interactions for useful context to identify birds
Enjoying Experience (5.5)	A range of diverse, changing interactions to account for sensory & animal preferences

6.1 Receiving Guidance: What Form of Logistical & Emotional Support Advances Learning?

The need for guidance to support the learning of animal calls and interpretation of environmental audio recordings is evident [40-42]. Bob Bristlebird, the conversational agent, was central to providing logistical instruction on game mechanics, audio interpretation, and nature information. The emotional support he provided was also a critical ingredient. Bob kept a lighthearted tone, even as players became self-conscious or taxed completing tasks. Agents built with ill-timed or presumptuous interventions are known to be disliked [56], so understanding how people feel about agents like Bob with long-term learning about nature deserves attention. Also, other forms of emotional support would be worth exploring.

6.2 Matching Media: How Can Gameful Interactions Lead to Stronger Contextualization?

Previous studies demonstrate that spectrograms can help people interpret environmental sound [13, 41]. Understanding what factors influence the shapes of sounds in spectrograms is difficult without training [40]. This study deepens our understanding of *how* players related auditory cues embedded in audio to visual representations. What makes this possible is the human brain's ability to receive and interpret multiple cues that share space and time (i.e. *crossmodal perception* [35]). While our spectrograms were a static visual on a screen, our players perceived how sounds in audio related to the visual over a common time. Whether relating audio and spectrograms of soundscapes or short calls, players crossmodally perceived differences between recordings. Matching calls was likely easier than matching soundscapes for players because of the difference in the amount of audio being presented at once. This study highlights the processes players used to crossmodally interpret spectrograms and audio recordings. Aspects such as how much audio to provide in activities, how many calls to feature, and the influence of environmental context, are worthwhile to investigate, for those seeking to support nature engagement and learning wildlife calls.

Players without prior knowledge were unable to match bird photos with calls, though they still had fun playfully personifying birds. Similarly, such serendipitous engagements personifying animals were also noted for some citizen scientists who were enticed to take more notice of local bees visiting their garden [58]. Personifying animals is common, whether pets or wildlife, and this may well be useful to help people connect with animals. One study has shown that zoo visitors who observe orangutans using interactive technologies "*responded with cognitive, affective and motor empathy for the animals*" [69, p. 6075]. Given our interest in helping wildlife, we want to ensure that creatively relating animals to how humans engage with the world is balanced with relating people to how animals interact with the world and their survival needs. Studies have begun to explore the design of playful and gameful encounters with information about nature [40, 60]. Though more research is needed to understand if such interactions result in greater long-term environmental awareness and meaningful actions.

Our players demonstrated that they possessed a wide range of prior experiences that applied to making sense of bird media. Several had experiences with music, local birds, or visualizations of audio that aided with media interpretation. Perhaps designing with groups that have specific skills could create citizen science synergies. A group of *ecological sound artists*, for example, have the potential to be one such ideal group given their mastery of sound and passion for raising environmental awareness on a variety of topics, from elephant conservation to preservation of river ecosystems [3, 27, 28].

6.3 Exploring Media: How Can Playful Exploration & Creativity be Harnessed?

Playfulness has received little attention in relation to environmental awareness or citizen science [47, 60]. Several players invested substantial time in playfully interacting with bristlebird calls and media for four additional species. People commonly switch into playful modes of thinking momentarily in gameplay [16], which was also observed during our playtests. During the taxing task of matching soundscapes, some imagined how calls related by shape and sound without prompting. Additionally, some creatively personified birds during gameful tasks, which appeared to keep them engaged, even when having little knowledge of the birds to relate to sounds. Later, players also readily engaged with creating names for the three bristlebird calls by sight and sound in the *Naming Calls* section. Both planned and unplanned moments of play illustrate the engaging nature of play. This study demonstrated the value in designing to foster exploration and creativity in nature-focused and citizen science gameplay activities. Other citizen science studies similarly found affording moments of such open exploration and creativity to be beneficial as well [47, 58].

HCI studies often highlight how interactions of verbalizing, sketching, photograph viewing, and audio listening can elicit emotive memories, aid in developing or regaining language, and support music composition or listening [31, 45, 51, 62]. Contrarily, our players often attached memories to calls when engaged with naming them by shapes and sounds. Being able to create a personalized language for bird calls based on shape and auditory characteristics may entice people who would traditionally feel overwhelmed with trying to directly learn the complexities of nature. Assigning these types of meaning may make calls more relatable, memorable, and recognizable. Further investigation of shape naming, sound naming, and memory-making activities for relating to nature is worthwhile.

6.4 Questing for Calls: How can Playful, Gameful, & Multisensory Experiences be Synergistic?

Players readily applied their previous playtest multimodal game experiences to the final task of identifying calls of Eastern bristlebirds. Given the exploratory nature of this study, we were not concerned with whether players accurately identified the bristlebird calls. However, if ecologists and citizen scientists want to identify calls from audio recordings to study wildlife, then there is a need to ensure that identifications are accurate and accessible [26]. This would help increase the scalability of using audio so that we can all learn about our wildlife. A variety of questions could be asked, such as which gameful and playful interactions most strongly influence call identification accuracy?

Our study found players were interested in other players' naming attempts as well as enjoying playful freeform activities themselves. To share call names remotely, players would need enticing, easy to use support to examine, tag, co-create, share, and discuss creative interpretations of bird calls by sight and sound. Sharing of call names has the potential to lead to new forms of folksonomies, which arise "*from data about how people associate terms with content that they generate, share, or consume*" [29, p. 1]. One text-based, freehand folksonomy exploration for audio with wildlife showed that usefulness for small-scale audio libraries may be limited because of being sensitive to differences in variables, such as spellings of a common name or onomatopoeic interpretations [66]. Nevertheless, folksonomies can help to coordinate citizen science efforts if systems support citizen scientists to curate the tags as well [32]. The visual and auditory pareidolia interpretations provide the basis for a new type of multisensory folksonomy. Along with textual tags, people may provide sketches and audio recordings to elucidate ideas. They may also provide context, offer musical interpretations, and share personal stories, to make learning what birds look and sound like more playful and memorable. Perhaps inspiration can come from the audio aspects of *Games with a Purpose* for music tagging [19, 33, 36]. Further investigation is needed to understand how player's mnemonic aides influence their ability to recognize calls in soundscapes. Additionally, exploring, whether sharing mnemonics is helpful, and how folksonomies can support both visual and auditory interpretations of wildlife calls is worthwhile.

6.5 Enjoying Experience: How Can Contextual Audio Encounters Build Care for Nature?

Our study highlights a variety of ways that players enjoyed interacting with environmental sounds when provided with guidance performing gameful tasks and exploring playfully. When searching for bristlebirds in the final section, players applied the strategies they developed in previous sections, as well as concurrently applying relevant prior knowledge. Citizen science projects commonly engage citizen scientists to complete single tasks, such as submit or classify photographs or audio recordings to eventually inform artificial intelligence algorithms [8]. However, people exploring bird media craved access to contextual information about the species and habitats [40, 41]. Our players gained satisfaction from personifying animals that caught their attention, which may signify the development of empathy whether interactions were in situ or digital explorations. It's worthwhile to remain vigilant to design towards community interests and empowerment. Design can provide ways that lead communities towards greater holistic engagement with nature and actions that support wildlife. This is contrast to most contributory citizen science, designed with communities being guided by scientific objectives with limited opportunities to foster true partnership.

Exploration of the Challenge evoked ideas for interaction strategies and revealed players gaining satisfaction at different levels based on sensory preferences. While gameful tasks in the Challenge were designed to compete against oneself, players mused about what completing it in pairs or teams would be like, whether through competitive or cooperative approaches. More broadly, each player had preferences with sight or sound. Deeply understanding the influences of such tendencies, such as differences in learning styles, physical ability, or prior experiences, may inform designs to reduce hurdles of engagement.

7 CONCLUSION

Our study explores the design and playtesting of the Bristle Whistle Challenge. This prototype was created to investigate how gameful and playful interactions with audio and visual media can support engaging with nature and learning how to identify bird calls. The Challenge is focused on endangered Eastern bristlebirds since the species stands to gain if the public becomes awareness of its plight and the citizen science community can support its conservation by identifying calls in audio recordings. Our exploratory Challenge playtests revealed that, with guidance from an agent, players were engaged, curious, creative, and gameful while completing the multisensory experience. This study highlights the value in designing creative multisensory interactions that provide environmental context when designing audio encounters. Such interactions also show promise to engender people with skills to identify wildlife. We contribute a detailed understanding of how players engaged with different media. We also offer future design research recommendations. Our studies point to the gulf between existing interactions with audio data and what is possible through attentive interaction design that explores multisensory and creative engagement. Our natural environment is in desperate need of human beings to hear its calls. Engaging people with nature is more so an interaction design problem than any other kind of problem. We, therefore, call upon the CHI community to step up efforts in designing to engage with nature.

ACKNOWLEDGEMENTS

Sincerest thanks to anonymous reviewers, Justin Cappadonna, Anne Bowser, Jessica Wardlaw, and Mark Cottman-Fields for their invaluable manuscript feedback. Additional thanks to Alex Mayor and Joee Kelk for permission to include their social media sketches. We are appreciative to numerous artists for permission to feature their gorgeous bird photos in the Challenge and music in the Supplementary Video (see video for details). The Australian Research Council (DP140102325) supported this work.

REFERENCES

[1] Obead Alhadreti and Pam Mayhew. 2018. Rethinking Thinking Aloud: A Comparison of Three Think-Aloud Protocols. In *Proceedings of the 2018 Chi Conference on Human Factors in Computing Systems*, Association for Computing Machinery, Paper 44. http://dx.doi.org/10.1145/3173574.3173618

[2] Anna Aljanaki, Dimitrios Bountouridis, John Ashley Burgoyne, Jan Van Balen, Frans Wiering, Henkjan Honing and Remco Veltkamp. 2014. Designing Games with a Purpose for Data Collection in Music

Research. Emotify and Hooked: Two Case Studies. In *Proceedings of Games and Learning Alliance* Springer International Publishing, 29-40. http://dx.doi.org/10.1007/978-3-319-12157-4_3
[3] Leah Barclay, Toby Gifford and Simon Linke. 2020. Interdisciplinary Approaches to Freshwater Ecoacoustics. *Freshwater Science* 39, 2, 356-361. http://dx.doi.org/10.1086/709130
[4] Margot Brereton, Malavika Vasudevan, Tshering Dema, Jessica L. Cappadonna, Cara Wilson and Paul Roe. 2017. The Ambient Birdhouse: Bringing Birds inside to Learn About Birds Outside. In *Proceedings of 2017 ACM Conference Companion Publication on Designing Interactive Systems* ACM, 321-324. http://dx.doi.org/10.1145/3064857.3079184
[5] Richard Buchanan. 1992. Wicked Problems in Design Thinking. *Design Issues* 8, 2, 5-21. http://dx.doi.org/10.2307/1511637
[6] Jessica L. Cappadonna, Margot Brereton, David M Watson and Paul Roe. 2016. Calls from the Wild: Engaging Citizen Scientists with Animal Sounds. In *Proceedings of 2016 ACM Conference Companion Publication on Designing Interactive Systems* ACM, New York, NY, USA, 157-160. http://dx.doi.org/10.1145/2908805.2909413
[7] Mark Cartwright, Ayanna Seals, Justin Salamon, Alex Williams, Stefanie Mikloska, Duncan MacConnell, Edith Law, Juan P. Bello and Oded Nov. 2017. Seeing Sound: Investigating the Effects of Visualizations and Complexity on Crowdsourced Audio Annotations. *Proceedings of ACM Human-Computer Interaction* Vol 1, 1-21. http://dx.doi.org/10.1145/3134664
[8] Luigi Ceccaroni, James Bibby, Erin Roger, Paul Flemons, Katina Michael, Laura Fagan and Jessica L Oliver. 2019. Opportunities and Risks for Citizen Science in the Age of Artificial Intelligence. *Citizen Science: Theory and Practice* 4, 1, 1-14. http://dx.doi.org/10.5334/cstp.241
[9] Mark Constantine. 2006. *The Sound Approach to Birding: A Guide to Understanding Bird Sound*. The Sound Approach,
[10] Mark Cottman-Fields, Margot Brereton and Paul Roe. 2013. Virtual Birding: Extending an Environmental Pastime into the Virtual World for Citizen Science. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris, France, 2029-2032. http://dx.doi.org/10.1145/2470654.2466268
[11] Mark Cottman-Fields, Margot Brereton, Jason Wimmer and Paul Roe. 2014. Collaborative Extension of Biodiversity Monitoring Protocols in the Bird Watching Community. In *Proceedings of 13th Participatory Design Conference*. ACM, Windhoek, Namibia, 111-114. http://dx.doi.org/10.1145/2662155.2662193
[12] Joe Cox, Eun Young Oh, Brooke Simmons, Chris Lintott, Karen Masters, Anita Greenhill, Gary Graham and Kate Holmes. 2015. Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects. *Computing in Science & Engineering* 17, 4, 28-41. http://dx.doi.org/10.1109/MCSE.2015.65
[13] Tshering Dema, Margot Brereton, Jessica L. Cappadonna, Paul Roe, Anthony Truskinger and Jinglan Zhang. 2017. Collaborative Exploration and Sensemaking of Big Environmental Sound Data. *Computer Supported Cooperative Work* 26, 4-6, 693-731. http://dx.doi.org/10.1007/s10606-017-9286-9
[14] Tshering Dema, Margot Brereton, Michael Esteban, Alessandro Soro, Sherub Sherub and Paul Roe. 2020. Designing in the Network of Relations for Species Conservation: The Playful Tingtibi Community Birdhouse. In *Proceedings of 2020 CHI Conference on Human Factors in Computing Systems* Association for Computing Machinery, 1–14. http://dx.doi.org/10.1145/3313831.3376713
[15] Tshering Dema, Margot Brereton and Paul Roe. 2019. Designing Participatory Sensing with Remote Communities to Conserve Endangered Species. In *Proceedings of Conference on Human Factors in Computing Systems*. ACM, Glasgow, Scotland Uk, 1-16. http://dx.doi.org/10.1145/3290605.3300894
[16] Sebastian Deterding, Staffan L. Björk, Lennart E. Nacke, Dan Dixon and Elizabeth Lawley. 2013. Designing Gamification: Creating Gameful and Playful Experiences. In *Proceedings of CHI '13 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, Paris, France, 3263–3266. http://dx.doi.org/10.1145/2468356.2479662
[17] Sebastian Deterding, Dan Dixon, Rilla Khaled and Lennart Nacke. 2011. From Game Design Elements to Gamefulness: Defining "Gamification". In *Proceedings of 15th International Academic MindTrek Conference: Envisioning Future Media Environments* Association for Computing Machinery, 9–15. http://dx.doi.org/10.1145/2181037.2181040
[18] Carl DiSalvo, Illah Nourbakhsh, David Holstius, Ayça Akin and Marti Louw. 2008. The Neighborhood Networks Project: A Case Study of Critical Engagement and Creative Expression through Participatory

Design. In *Proceedings of Tenth Anniversary Conference on Participatory Design 2008* Indiana University, 41–50. http://dx.doi.org/10.5555/1795234.1795241
[19] Peter Dulačka, Jakub šimko and Mária Bieliková. 2012. Validation of Music Metadata Via Game with a Purpose. In *Proceedings of 8th International Conference on Semantic Systems*. Association for Computing Machinery, Graz, Austria, 177–180. http://dx.doi.org/10.1145/2362499.2362526
[20] Ilker Etikan, Sulaiman Abubakar Musa and Rukayya Sunusi Alkassim. 2016. Comparison of Convenience Sampling and Purposive Sampling. *American journal of theoretical and applied statistics* 5, 1, 1-4. http://dx.doi.org/10.11648/j.ajtas.20160501.11
[21] Alexandra Eveleigh, Charlene Jennett, Stuart Lynn and Anna L. Cox. 2013. "I Want to Be a Captain! I Want to Be a Captain!": Gamification in the *Old Weather* Citizen Science Project. In *Proceedings of First International Conference on Gameful Design, Research, and Applications*. ACM, Toronto, Ontario, Canada, 79-82. http://dx.doi.org/10.1145/2583008.2583019
[22] Tracy Fullerton. 2014. Chapter 9: Playtesting. In *Game Design Workshop: A Playcentric Approach to Creating Innovative Games, Third Edition*, A K Peters/CRC Press LLC, New York, United States, 271-304. http://dx.doi.org/10.1201/b16671
[23] William Gaver. 2002. Designing for Homo Ludens. *I3 Magazine No. 12*, June 2002, 2-6.
[24] William Gaver, John Bowers, Andrew Boucher, Andy Law, Sarah Pennington and Brendan Walker. 2007. Electronic Furniture for the Curious Home: Assessing Ludic Designs in the Field. *International Journal of Human–Computer Interaction* 22, 1-2, 119-152. http://dx.doi.org/10.1080/10447310709336958
[25] William W. Gaver, John Bowers, Andrew Boucher, Hans Gellerson, Sarah Pennington, Albrecht Schmidt, Anthony Steed, Nicholas Villars and Brendan Walker. 2004. The Drift Table: Designing for Ludic Engagement. In *Proceedings of CHI '04 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, Vienna, Austria, 885–900. http://dx.doi.org/10.1145/985921.985947
[26] Rory Gibb, Ella Browning, Paul Glover-Kapfer and Kate E. Jones. 2018. Emerging Opportunities and Challenges for Passive Acoustics in Ecological Assessment and Monitoring. *Methods in Ecology and Evolution* 10, 169– 185. http://dx.doi.org/10.1111/2041-210X.13101
[27] Jonathan Gilmurray. 2017. Ecological Sound Art: Steps Towards a New Field. *Organised Sound* 22, 1, 32-41. http://dx.doi.org/10.1017/S1355771816000315
[28] Sue Gould. 2018. Vicki Hallett and the Elephant Listening Project. *AUDIOWINGS: Journal of the Australia Wildlife Sound Recording Group Inc* 2, 2, 28.
[29] Thomas Gruber. 2007. Ontology of Folksonomy: A Mash-up of Apples and Oranges. *International Journal on Semantic Web and Information Systems* 3, 1, 1-11. http://dx.doi.org/10.4018/jswis.2007010101
[30] Andrew P. Hill, Peter Prince, Jake L. Snaddon, C. Patrick Doncaster and Alex Rogers. 2019. Audiomoth: A Low-Cost Acoustic Device for Monitoring Biodiversity and the Environment. *HardwareX* 6, e00073. http://dx.doi.org/10.1016/j.ohx.2019.e00073
[31] Maarten Houben, Rens Brankaert, Saskia Bakker, Gail Kenning, Inge Bongers and Berry Eggen. 2020. The Role of Everyday Sounds in Advanced Dementia Care. In *Proceedings of 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Honolulu, HI, USA, 1–14. http://dx.doi.org/10.1145/3313831.3376577
[32] Corey Jackson, Kevin Crowston, Carsten Østerlund and Mahboobeh Harandi. 2018. Folksonomies to Support Coordination and Coordination of Folksonomies. *Computer Supported Cooperative Work* 27, 647–678. http://dx.doi.org/10.1007/s10606-018-9327-z
[33] Bryan Kim, Yi Cheng, Zixuan Li, Ruoxi Li, Chenchen Tan, Shuo Wang, Yifeng Shi and Jessica Hammer. 2019. Game with a Purpose to Collect Home Audio Data. In *Proceedings of Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts*. Association for Computing Machinery, Barcelona, Spain, 445–452. http://dx.doi.org/10.1145/3341215.3356270
[34] Christopher Kullenberg and Dick Kasperowski. 2016. What Is Citizen Science? – a Scientometric Meta-Analysis. *PLOS ONE* 11, 1, e0147152. http://dx.doi.org/10.1371/journal.pone.0147152
[35] Christophe Lalanne and Jean Lorenceau. 2004. Crossmodal Integration for Perception and Action. *Journal of Physiology-Paris* 98, 1, 265-279. http://dx.doi.org/10.1016/j.jphysparis.2004.06.001
[36] Edith Law and Luis von Ahn. 2009. Input-Agreement: A New Mechanism for Collecting Data Using Human Computation Games. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems* Association for Computing Machinery, 1197–1206. http://dx.doi.org/10.1145/1518701.1518881

[37] Edith L.M. Law, Luis von Ahn, Roger B. Dannenberg and Mike Crawford. 2007. Tagatune: A Game for Music and Sound Annotation. In *Proceedings of International Society of Music Information Retrieval*. Vienna, Austria, 361-364.
[38] Michael I Mandel and Daniel P.W. Ellis. 2008. A Web-Based Game for Collecting Music Metadata. *Journal of New Music Research* 37, 2, 151-165. http://dx.doi.org/10.1080/09298210802479300
[39] Merriam-Webster. 2020. Pareidolia. from https://www.merriam-webster.com/dictionary/pareidolia.
[40] Jessica L. Oliver, Margot Brereton, Selen Turkay, David M. Watson and Paul Roe. 2020. Exploration of Aural & Visual Media About Birds Informs Lessons for Citizen Science Design. In *Proceedings of 2020 ACM on Designing Interactive Systems Conference*. Association for Computing Machinery, Eindhoven, Netherlands, 1687–1700. http://dx.doi.org/10.1145/3357236.3395478
[41] Jessica L. Oliver, Margot Brereton, David M. Watson and Paul Roe. 2018. Visualisations Elicit Knowledge to Refine Citizen Science Technology Design: Spectrograms Resonate with Birder. In *Proceedings of 30th Australian Conference on Computer-Human Interaction* ACM, 133-144. http://dx.doi.org/10.1145/3292147.3292171
[42] Jessica L. Oliver, Margot Brereton, David M. Watson and Paul Roe. 2019. Listening to Save Wildlife: Lessons Learnt from Use of Acoustic Technology by a Species Recovery Team. In *Proceedings of 2019 Designing Interactive Systems Conference*. Association for Computing Machinery, San Diego, CA, USA, 1335-1348. http://dx.doi.org/10.1145/3322276.3322360
[43] Maria V. Palacin-Silva, Antti Knutas, Maria Angela Ferrario, Jari Porras, Jouni Ikonen and Chandara Chea. 2018. The Role of Gamification in Participatory Environmental Sensing: A Study in the Wild. In *Proceedings of 2018 CHI Conference on Human Factors in Computing Systems*. ACM, Montreal QC, Canada, 1-13. http://dx.doi.org/10.1145/3173574.3173795
[44] Leo Natan Paschoal, Lucas Fernandes Turci, Tayana Uchôa Conte and Simone R. S. Souza. 2019. Towards a Conversational Agent to Support the Software Testing Education. In *Proceedings of XXXIII Brazilian Symposium on Software Engineering*. Association for Computing Machinery, Salvador, Brazil, 57–66. http://dx.doi.org/10.1145/3350768.3352456
[45] Anne Marie Piper, Nadir Weibel and James D Hollan. 2014. *Designing Audio-Enhanced Paper Photos for Older Adult Emotional Wellbeing in Communication Therapy*. Academic Press, Inc., http://dx.doi.org/10.1016/j.ijhcs.2014.01.002
[46] Michael J. O. Pocock, John C. Tweddle, Joanna Savage, Lucy D. Robinson and Helen E. Roy. 2017. The Diversity and Evolution of Ecological and Environmental Citizen Science. *PLOS ONE* 12, 4, e0172579. http://dx.doi.org/10.1371/journal.pone.0172579
[47] Marisa Ponti, Thomas Hillman and Igor Stankovic. 2015. *Science and Gamification: The Odd Couple?* Association for Computing Machinery, London, United Kingdom. http://dx.doi.org/10.1145/2793107.2810293
[48] Jennifer Preece. 2016. Citizen Science: New Research Challenges for Human–Computer Interaction. *International Journal of Human–Computer Interaction* 32, 8, 585-612. http://dx.doi.org/10.1080/10447318.2016.1194153
[49] Jennifer Preece. 2017. How Two Billion Smartphone Users Can Save Species! *interactions* 24, 2, 26-33. http://dx.doi.org/10.1145/3043702
[50] Danial Qaurooni, Ali Ghazinejad, Inna Kouper and Hamid Ekbia. 2016. Citizens for Science and Science for Citizens: The View from Participatory Design. In *Proceedings of 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose, California, USA, 1822-1826. http://dx.doi.org/10.1145/2858036.2858575
[51] Hayes Raffle, Cati Vaucelle, Ruibing Wang and Hiroshi Ishii. 2007. Jabberstamp: Embedding Sound and Voice in Traditional Drawings. In *Proceedings of 6th international conference on Interaction design and children*. Association for Computing Machinery, Aalborg, Denmark, 137–144. http://dx.doi.org/10.1145/1297277.1297306
[52] Ute Ritterfeld, Michael J. Cody, Peter Vorderer and (Eds.). 2009. *Serious Games: Mechanisms and Effects*. Routledge, London. http://dx.doi.org/10.4324/9780203891650
[53] Hermann Rorschach. 1942. *Psychodiagnostics: A Diagnostic Test Based on Perception*. (P Lemkau & B. Kronenberg (Trans): Verlag Hans Huber, http://dx.doi.org/10.1037/11537-000
[54] Holly Rosser and Andrea Wiggins. 2018. *Tutorial Designs and Task Types in Zooniverse*. Association for Computing Machinery, Jersey City, NJ, USA. http://dx.doi.org/10.1145/3272973.3274049

[55] Jodi JL Rowley, Corey T Callaghan and Timothy Cutajar. 2019. FrogID: Citizen Scientists Provide Validated Biodiversity Data on Frogs of Australia. *Herpetological Conservation and Biology* 14, 1, 155-170.
[56] Paul Rudman and Mary Zajicek. 2006. Autonomous Agent as Helper - Helpful or Annoying? In *Proceedings of IEEE/WIC/ACM International Conference on Intelligent Agent Technology*. IEEE Computer Society, 170–176. http://dx.doi.org/10.1109/iat.2006.41
[57] Mangalam Sankupellay, Anna Kalma, Sean Magin, Jessica L Cappadonna, Paul Roe and Margot Brereton. 2017. Birdsound: Enticing Urban Dwellers to Engage with Local Birds around Their Home. In *Proceedings of 29th Australian Conference on Computer-Human Interaction* ACM, 172-181. http://dx.doi.org/10.1145/3152771.3152790
[58] Nirwan Sharma, Greaves Sam, Laura Colucci-Gray, Advaith Siddharthan and René van der Wal. 2019. From Citizen Science to Citizen Action: Analysing the Potential for a Digital Platform to Cultivate Attachments to Nature. *Journal of Science Communication* 18, 1, 1-35. http://dx.doi.org/10.22323/2.18010207
[59] Jennifer L Shirk, Heidi L Ballard, Candie C Wilderman, Tina Phillips, Andrea Wiggins, Rebecca Jordan, Ellen McCallie, Matthew Minarchek, Bruce V Lewenstein and Marianne E Krasny. 2012. Public Participation in Scientific Research: A Framework for Deliberate Design. *Ecology and Society* 17, 2, 29. http://dx.doi.org/10.5751/ES-04705-170229
[60] Alessandro Soro, Margot Brereton, Tshering Dema, Jessica L Oliver, Min Zhen Chai and Aloha May Hufana Ambe. 2018. The Ambient Birdhouse: An Iot Device to Discover Birds and Engage with Nature. In *Proceedings of Conference on Human Factors in Computing Systems* ACM, 1-13. http://dx.doi.org/10.1145/3173574.3173971
[61] Kyle A. Swiston and Daniel J. Mennill. 2009. Comparison of Manual and Automated Methods for Identifying Target Sounds in Audio Recordings of Pileated, Pale-Billed, and Putative Ivory-Billed Woodpeckers. *Journal of Field Ornithology* 80, 1, 42-50. http://dx.doi.org/10.1111/j.1557-9263.2009.00204.x
[62] Cristina Sylla, Sérgio Gonçalves, Pedro Branco and Clara Coutinho. 2013. Peter Piper Picked a Peck of Pickled Peppers: An Interface for Playful Language Exploration. In *Proceedings of CHI '13 Extended Abstracts on Human Factors in Computing Systems*. Association for Computing Machinery, Paris, France, 3127–3130. http://dx.doi.org/10.1145/2468356.2479627
[63] Virginia Thomas. 2012. National Recovery Plan for Eastern Bristlebird *Dasyornis Brachypterus*. 05 Jun, 2016 from https://www.environment.gov.au/resource/national-recovery-plan-eastern-bristlebird-dasyornis-brachypterus.
[64] Ramine Tinati, Markus Luczak-Roesch, Elena Simperl and Wendy Hall. 2017. An Investigation of Player Motivations in Eyewire, a Gamified Citizen Science Project. *Computers in Human Behavior* 73, 527-540. http://dx.doi.org/10.1016/j.chb.2016.12.074
[65] Anthony Truskinger, Mark Cottman-Fields, Daniel Johnson and Paul Roe. 2013. Rapid Scanning of Spectrograms for Efficient Identification of Bioacoustic Events in Big Data. In *Proceedings of 2013 Institute of Electrical and Electronics Engineers Ninth International Conference on eScience* Institute of Electrical and Electronics Engineers (IEEE), 270-277. http://dx.doi.org/10.1109/eScience.2013.25
[66] Anthony Truskinger, Ian Newmarch, Mark Cottman-Fields, Jason Wimmer, Michael Towsey, Jinglan Zhang and Paul Roe. 2013. Reconciling Folksonomic Tagging with Taxa for Bioacoustic Annotations. In *Web Information Systems Engineering – Wise 2013: 14th International Conference, Nanjing, China, October 13-15, 2013, Proceedings, Part I*, Xuemin Lin et al. Eds. Springer, Berlin, Heidelberg, 292-305. http://dx.doi.org/10.1007/978-3-642-41230-1_25
[67] Kellie Vella, Jessica L. Oliver, Tshering Dema, Margot Brereton and Paul Roe. 2020. Ecology Meets Computer Science: Designing Tools to Reconcile People, Data, and Practices. In *Proceedings of 2020 CHI Conference on Human Factors in Computing Systems* Association for Computing Machinery, 1–13. http://dx.doi.org/10.1145/3313831.3376663
[68] Luis von Ahn and Laura Dabbish. 2008. Designing Games with a Purpose. *Communications of the ACM* 51, 8, 58-67. http://dx.doi.org/10.1145/1378704.1378719
[69] Sarah Webber, Marcus Carter, Sally Sherwen, Wally Smith, Zaher Joukhadar and Frank Vetere. 2017. Kinecting with Orangutans: Zoo Visitors' Empathetic Responses to Animals? Use of Interactive Technology. In *Proceedings of 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver, Colorado, USA, 6075-6088. http://dx.doi.org/10.1145/3025453.3025729