NetGauge Games Platform: Connecting People and Place while Crowdsourcing Broadband Measurements

JARED DUVAL, COLE PENDERGRAFT, TAYLEN JOHNSON, OLIVIA VESTER, CHAITHANYA HEBLIKAR, HUNTER BEACH, SHELBY HAGEMANN, MD NAZMUL HOSSAIN, VARNA GOBBUR, and ALISON GRAHAM, Northern Arizona University, USA
MORGAN VIGIL-HAYES, Michigan State University, USA

Collecting measurements of mobile broadband (cellular Internet) performance is critical for efforts addressing digital disparities. However, collecting measurements at the necessary spatiotemporal scales to identify performance and coverage gaps is non-trivial. Crowdsourced data collection is a scalable approach to gathering mobile broadband performance data. However, existing platforms for crowdsourced mobile broadband measurements are not designed to engage workers over time or space, which can lead to spatial misrepresentation, stale data, and, ultimately, inequitable Internet access. With the insight that games and play offer naturally engaging frameworks for users, we iteratively implemented a game platform that collects Internet measurements–NetGauge Games–along with three co-designed broadband measurement games hosted within the platform. We conducted user evaluation of the platform across three locations with 25 participants and measured participants' connection to people and place using mixed methods. We found that each game's affordances yielded varied movement through space, connection to others, connection to environment, and amount of internet measurements, and we elucidate ultimate design particulars that support these often competing goals. Overall, the games performed significantly better regarding number of broadband measurements collected, connection to space, connection to others, radius of gyration, and game rankings when compared to the Speed Test control.

CCS Concepts: • **Networks** → **Network measurement**; • **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in collaborative and social computing**; • **Social and professional topics** → Geographic characteristics.

Additional Key Words and Phrases: mobile broadband measurement, crowdsourcing, play, rapid iterative testing and evaluation

1 INTRODUCTION

While Internet access is a human right [73], 2.6 billion people worldwide lack Internet access and only 48% of the world's rural population has access [42]. One technology that has rapidly increased global access to the Internet is *mobile broadband*, or Internet delivered to users over cellular networks (e.g., 3G, 4G/LTE, 5G). Mobile broadband is relatively inexpensive to deploy, smartphone devices are becoming increasingly affordable, and data plans tend to be less expensive than fixed Internet services. Indeed, at the time of writing, 92% of the world is "covered" by 4G/LTE networks [42]. However, even as mobile broadband coverage becomes increasingly pervasive, measurements of mobile broadband performance (e.g. "speed tests") reveal that coverage maps can hide areas where broadband infrastructure fails to provide usable Internet access [31, 54, 55, 68, 79]. The mismatch between reported coverage using standard practice radio propagation models and on-the-ground performance can be a major barrier to addressing Internet disparities, as it prevents network operators and policymakers from understanding where performance gaps prevent users from accessing usable Internet connectivity [1, 55, 64, 68].

Authors' Contact Information: Jared Duval, Jared.Duval@nau.edu; Cole Pendergraft, cp2462@nau.edu; Taylen Johnson, tsj85@nau.edu; Olivia Vester, orv9@nau.edu; Chaithanya Heblikar, csh342@nau.edu; Hunter Beach, hmb433@nau.edu; Shelby Hagemann, seh428@nau.edu; Md Nazmul Hossain, mh2752@nau.edu; Varna Gobbur, vg574@nau.edu; Alison Graham, akg352@nau.edu, Northern Arizona University, School of Informatics Computing and Cyber Systems, Flagstaff, Arizona, USA; Morgan Vigil-Hayes, vigilhay@msu.edu, Michigan State University, Department of Media & Information, East Lansing, Michigan, USA.

One way to address this problem is through the collection of mobile broadband measurement data using crowdsourcing. Indeed, 47% of The Organization for Economic Cooperation and Development (OECD) countries use crowdsourcing approaches to assess both the quality and availability of broadband from a user perspective with the goal of informing policy and infrastructure deployment efforts [39, 52, 55, 83]. Unfortunately, the crowdsourcing platforms used to collect mobile broadband measurements can be dry and unengaging [85], which prevent scaling crowdsourcing efforts to the spatial and temporal scales required to generate useful data about Internet service gaps. Therefore, with the insight that games and play offer naturally engaging frameworks for crowdsourcing data, in our previous work, we conducted five iterative participatory design sessions and created a speculative design catalog of 11 co-created broadband measurement games as well as an annotated mockup of a flexible platform for hosting all of the games, NetGauge [24].

In this work, we implemented the NetGauge Games platform with three of the games from the speculative design catalog to evaluate them along two goals. The first is to better understand the effectiveness of the games in their underlying utilitarian goal, which is to measure broadband across time and space. The second goal is to better understand the social-emotional experiential affordances of the games because their utilitarian goal cannot be met if the play experience is not intrinsically rewarding. Because the games are played outdoors, we were specifically interested in how these games connect people to each other and to their environments. Therefore, the specific research questions that guided this work are:

RQ1: *How well do NetGauge Games improve crowdsourced Internet measurement over space?*

RQ2: *To what extent do NetGauge Games foster a sense of community among players and environment?*

There are several key findings. When compared to the US Federal Communications Commission (FCC) inspired control, NetGauge Games had increased the average number of measurements collected per user by up to a factor of 2.21; increased distance covered by an average of 154m^2, and increased duration of exploration by an average of 891 seconds ($\sim$ 15 minutes). Experientially, when compared to the FCC-inspired control, NetGauge Games had an average increase in connectedness to other players of 124% and to their environment of 58%. Interestingly, there are significant correlations related to connection between players and environment among the various games and locations, explained by various game affordances, which are described in this paper. Overall, this work makes three contributions:

(1) *Artifact*: A mixed methods evaluation of a community-framework oriented platform for employing play to support the collection of crowdsourced broadband measurements.
(2) *Empirical*: An analysis of mobility metrics and quantity of broadband measurements, which are evaluated across varied geographical locations and between games.
(3) *Theory*: An elucidation of design particulars that inform varied player experiences related to social connectedness and place connection (our innate drive to be connected to a physical location for belonging, identity, and meaning).

The paper is organized as follows:

(1) In Section 2, we share background on the underlying motivations of the work, which is to reduce internet inequities by crowdsourcing broadband measurements for more accurate coverage maps using games to increase engagement.
(2) In Section 3, we describe and extend on our previous speculative design work [24] by sharing implementation details of the NetGauge platform, three supported games, and an FCC speed tester control.

(3) In Section 4, we present how the games, platform, and research protocol iteratively changed in parallel during our piloting with 13 participants using a Rapid Iterative Testing & Evaluation (RITE) approach [22].
(4) In Section 5, we provide an overview of the finalized mixed-methods research protocol, demographics of our 12 playtest participants, and ethical considerations
(5) In Section 6, guided by our research questions, we share results on the utilitarian goal of crowdsourcing broadband measurement times over time and space (*RQ1*) and the experiential results on NetGauge's ability to foster a sense of community among players and environment (*RQ2*).
(6) In Section 7, we elucidate design particulars [41], such as intentional ambiguity [81], that create intrinsically motivating playful experiences that also serve a utilitarian goal, facilitate connection to our environment, and foster social community.
(7) In Section 8, we present future plans to share our use of RITE [22] as a pedagogical tool for training student researchers, plans to host workshops for creating new community-driven games hosted by NetGauge, and plans to create an interactive data visualization supported by NetGauge's telemetry data to inform broadband policy and infrastructure investment.

2 BACKGROUND & RELATED WORK

This research centers on addressing Internet inequities and relates to work that uses human-centered approaches to measure Internet performance across different communities. Specifically, we employ games to foster community and connect us to space by employing intentional ambiguity, described in the following sections.

2.1 Location Based Games for Data Collection

Games and play are invaluable tools for transforming serious or mundane activities into engaging and enjoyable experiences [16, 25, 26, 53, 78, 80, 90]. Serious games have been recognized in the fields of health and education for their effectiveness in promoting positive behavioral change [15, 51, 89] and improved health outcomes [33, 35, 63, 65, 72, 72, 76, 89]. Gamification in particular has gained prominence [65, 76], emphasizing points, badges, and leaderboards [20]. However, while gamification may improve initial engagement, it does not necessarily precipitate long-term behavior change or significant outcomes [87]. Deep engagement in the design process and creating intrinsically rewarding experiences are key to lasting engagement and targeted change [14, 59]. Therefore, in our previous work [24], instead of adding leaderboards and badges to the FCC app, we conducted participatory design bodystorming sessions with 11 participants in three locations representing a spectrum between rural and urban sites to generate a catalog of 11 game concepts for crowdsourcing broadband measurements across time and space.

In the context of crowdsourcing mobile broadband performance data, location-based games (LBGs) can prompt players to move strategically around space [48, 49]. To investigate how people in a particular community envisioned LBGs as helping to promote crowdsourced mobile broadband measurement, we engaged in a series of iterative co-design sessions in various sites to bodystorm [57] "Internet measurement game" concepts with 11 participants [24]. Our co-design study resulted in a catalog of 11 game concepts that centered situated play design [5, 9, 10]. Analysis of the 11 game concepts revealed that all Internet measurement LBGs could be represented as hierarchical structures. This led to the speculative design of the NetGauge Games concept, which we elaborate on in Section 3.

2.2 Play and Games Fostering Community with People and Space via Intentional Ambiguity

2.2.1 Space and Environment. Play and games have always connected us to our environment [12, 13, 17, 37, 88, 92], with some obvious examples being playground activities like *Tag*, *Red Rover*, water balloon fights, and building a snowman—but digital games connect us to our environment too. This may be most evident in games like Pokemon Go [88], but console games like *Assassins Creed* have so accurately captured our cultural heritage and beloved cities that Notre-Dame is being rebuilt using the game's model as a reference for historians, architects, and researchers [37]. Games and play are often seen as children's activities, but all ages are connected to space through play. For example, for many golfers, "playing at Pebble Beach is often one of the highlights of their life" [17]. While it is clear play facilitates feeling connected to space, this research elucidates how the ultimate particulars [41] of play design affect this connection to space attachment.

To design for play that connects us to outdoor spaces our work employs Situated Play Design [2, 5, 9, 10, 24]. When designing for play, context is crucial because play is ephemeral, ever-changing, and spontaneous, making it difficult to systematize [2]. However, Situated Play Design has demonstrated success in creating playful sociotechnical systems in many spaces, including urban spaces [4], dinner tables [6], social media (e.g., TikTok [23] and Instagram [11]), and circus training gyms [26, 58]. Of particular relevance to this research is Situated Play Design's recent outdoor projects, including location-based games [48] that incentivize travel, six design probes that promote reflections in nature [8], joyful human-forest interactions [3], and a playful one-month backpacking adventure [7]. Because the utilitarian success of this research depends on broadband measurements across both time and space, we depend on methods that have a track record of playfully engaging people outdoors for sustained time periods and over long distances [7].

2.2.2 Community. Similar to connecting us to our environment, play and games have always connected us to each other [43]. This is particularly true for multiplayer games, where games like *Halo* can teach us leadership and collaboration [91]. Regardless of the implementation and whether intentional or not, the design of games afford "Socio-pleasure" derived from our natural enjoyment from interacting with others [66, p. 105]. This joy keeps us playing, which is a precursor to utilitarian goals, including collecting broadband measurements. There is incredible power in playing together, with countless utilitarian applications of games facilitating crowdsourced human computation that benefits society and science, including *Foldit* to better understand complex protein structures, *Phylo* for genomic sequencing, *Galaxy Zoo* for classifying galaxies, *ESP* for generating image labels, *TagATune* for describing audio, *Pebble It* for resource allocation, and *Photo City* for creating 3D models of our cities [19]. NetGauge Games takes inspiration from these citizen science games and uses telemetry [22] to crowdsource broadband measurements for accurate maps to inform policy, which represents the shared values of a community.

2.2.3 Ambiguity as a Design Resource. To accomplish our overarching goal of collecting broadband measurements across time and space, NetGauge must continually provide compelling experiences and high replayability. Key to working in many contexts by diverse players is to design for novel appropriation, similar to how somaesthetics and design probes (e.g., [60, 77, 84]) facilitate "Design through Use" opportunities [40]. During our Bodystorming [56, 57, 82] design sessions and development work, we regularly revisited our value of designing for *Multiplicity* [26] using a Value-Centered Approach [32], so that players can navigate intentional ambiguity [81]. Allowing players to reconcile play choices outside of the technology for unique appropriations is intended to afford a sense of ownership and connection to both other players and the play space itself, extending the *Magic Circle* [44, 46] of the NetGauge Platform to include non-screen-centric interactions with the environment and unscripted social affordances. Additionally, the NetGauge platform is made to support "Design after Design" through the Twine *Deck* paradigm (Described in Section 3) [24] that allows community members to contribute new games to the platform in addition to appropriating existing games. This expands the platform's capabilities to crowdsource both the broadband measurements and the games themselves through ambiguity as a design resource.

3 NETGAUGE GAMES DESIGN

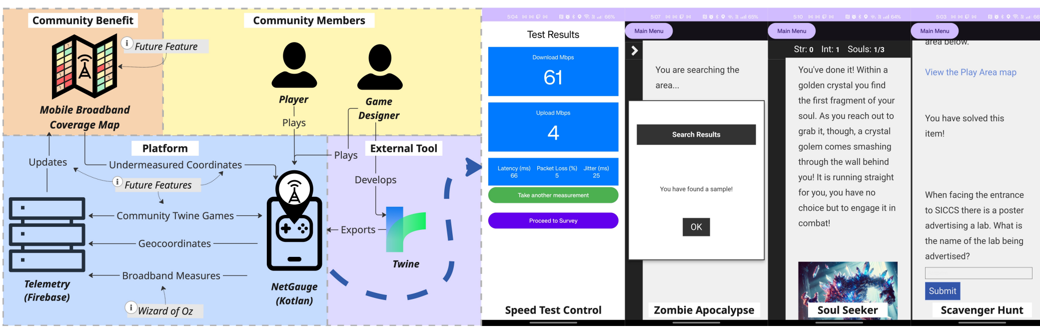


Fig. 1. An overview of the NetGauge platform. On the left, a diagram depicting how community members play and create games for crowdsourcing broadband measurements with annotated details about future features, system architecture, and data flow. On the right, annotated screenshots of the application.

This work extends an initial co-design study that engaged 11 participants in developing a catalog of location-based game concepts to support crowdsourced Internet measurement [24]. In that study, we found that all of the designed games shared a hierarchical structure and could be effectively represented as deck-based games [24]. Several concepts were culturally and geographically embedded within sovereign tribal communities, prompting us to envision a platform that could empower grassroots efforts to create community-specific games.

This vision led to the development of NetGauge Games, a mobile platform implemented as an Android application to prioritize open-sourcing—though our future efforts would look to porting a version for iOS. The platform integrates games built in Twine [47] with a Kotlin backend. Native application development is required to measure cellular network performance at the system level. Twine was selected for its accessibility to non-programmers and its ability to support immersive, community-driven storytelling. To align with participatory action research principles that emphasize community involvement at all stages of research and implementation [21, 45], we extended Twine's functionality to allow communities to author original location-based games. NetGauge supports modular Twine imports with custom broadband measurement hooks as the underlying game engine, enabling community-authored games for mobile broadband data collection. The hooks and extended functionality include geocoordinates of target destinations that can be dynamically fetched, haptic and textual feedback to guide player exploration to the target locations, described further in Section 4, and network performance metrics (e.g., latency, jitter, packet loss, download speed, upload speed) that can be used as conditions for Twine branching. The cloud back-end of NetGauge uses Firebase for the telemetry broadband measurements and geocoordinates, survey responses, and future community twine games. A diagram of the system architecture is shown in Figure 1. Figure 1 also details future features of the platform that are outside of this work's scope and research questions. Specifically, future work will explore the integration of NetGauge and dynamic broadband coverage maps that can be used to inform policy and suggest point of interest geocoordinates to measure that are underrepresented or stale.

For this study, we selected three games from the catalog that embodied relative endpoints on a range distinctive game dimensions, including: social vs. solitary; competitive vs. collaborative; asynchronous vs. synchronous; play-first vs. game first [24]. These dimensions correspond to how much each game requires interaction with other players, whether games can be played apart (both time and space) from other players, and the amount of structure versus ambiguity each game provides. To understand how LBGs impacted user mobility and connection to community compared to existing crowdsourced broadband measurement applications, we also implemented a replica of the US Federal Communications Commission (FCC) Speed Test app to function as a control activity in our study [28, 62]. We describe the LBGs and the control experiences in detail in Section 3.1.

3.1 Game Descriptions

All of the games used a Wizard of Oz approach to collecting network performance measurements. Specifically, we simulated the collection of typical quality of service metrics including upload speed, download speed, jitter, packet loss rate, and latency. Values for each metric were probabilistically generated from normal distributions of values corresponding to typical network performance that was measured at each of the sites using the FCC Speed Test app. All games were implemented as predominantly text-based and did not use audio. We provide a representative screenshot of each of the games in Figures 2 through 5. Each Twine game is available in our supplementary materials.

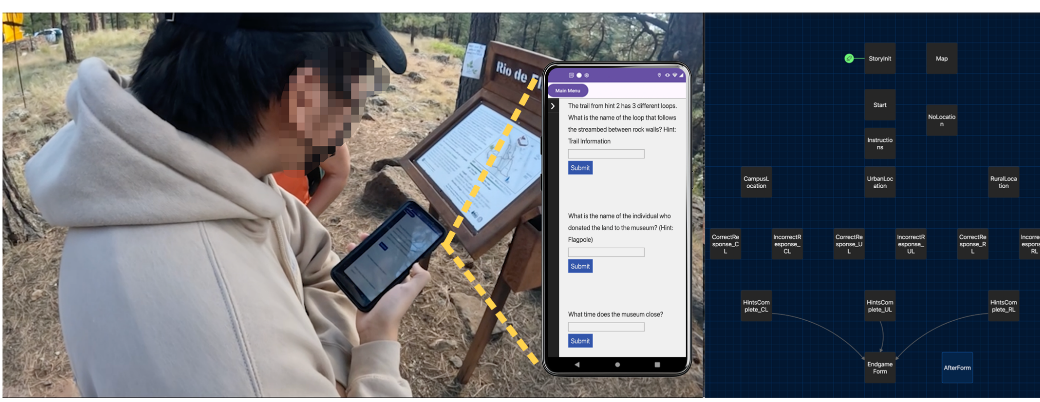


Fig. 2. A participant playing Scavenger Hunt at the rural location with an enlarged overlay depicting the game's user interface and a birds eye view of the Twine structure to the right.

3.1.1 Scavenger Hunt (SH). A treasure hunt game that was implemented to be competitive and asynchronous. Scavenger Hunt is a location-based Twine game designed to engage players in real-world environments. Players are presented with situated questions that require observation and interaction with their surroundings. Players input responses to these prompts, unlocking progressions through a structured sequence of location-aware challenges. The game dynamically adjusts its content based on detected player location and when a player checks whether their answer was correct, the game collects network performance measurements. The game ends when the player has successfully answered all clues and they receive a report on the total time and correct answers for an element of competition related to other players' performance at the site.

3.1.2 Soul Seeker (SS). A choose your own adventure game designed to be solitary and synchronous. Soul Seeker is a narrative Twine game that places the player in the role of a lost soul navigating a surreal and fragmented world in search of their identity. As players journey through strange realms, they encounter moral dilemmas, cryptic memories, and an enigmatic entity named Omni who challenges their perception of truth and self. The game unfolds through branching choices that influence both the tone of the journey and its outcome, ultimately leading to one of three possible endings. Throughout the experience, themes of memory, agency, and transformation are explored, making the player's decisions central to the narrative arc. Outcomes of the game are impacted by random dice rolls, players' dynamic strength and intelligence stats, and the choices they make. The soul fragments are located spread apart at each location. Network performance metrics are collected when a player is "searching" for their soul.

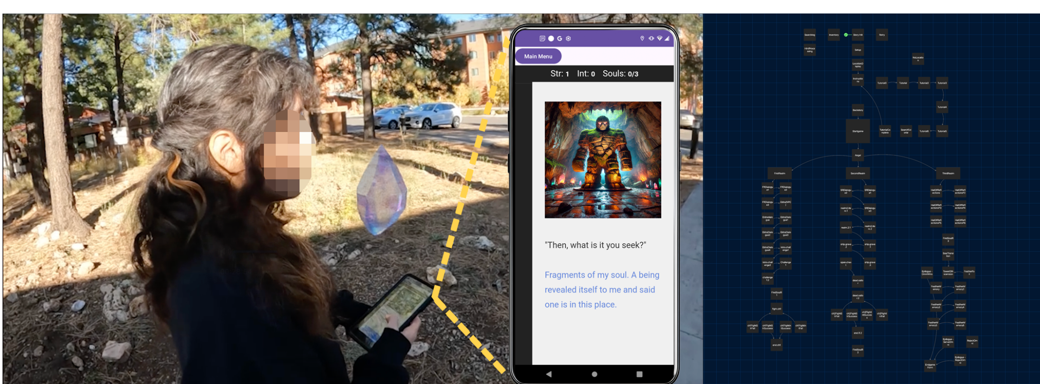


Fig. 3. A participant playing Soul Seeker at the campus location with an enlarged overlay depicting the game's user interface and a birds eye view of the Twine structure to the right.

3.1.3 Zombie Apocalypse (ZA). Zombie Apocalypse is a team-based, location-aware Twine game that blends mobile network measurement with competitive tag mechanics. Players choose to play as either a Scientist or a Zombie, with roles designed to be asymmetrical and emphasize different forms of interaction and mobility. Scientists move freely in search of antibody samples to create a zombie vaccine, using device vibration and orientation cues. Zombies, by contrast, use dice-rolled network measurements to determine their movement range and attack capabilities, choosing between "Speed" and "Strength" adaptations to pursue and infect Scientists. The game ends either when a Scientist successfully collects all required data samples and synthesizes a vaccine, or when Zombies manage to successfully attack a Scientist three times.

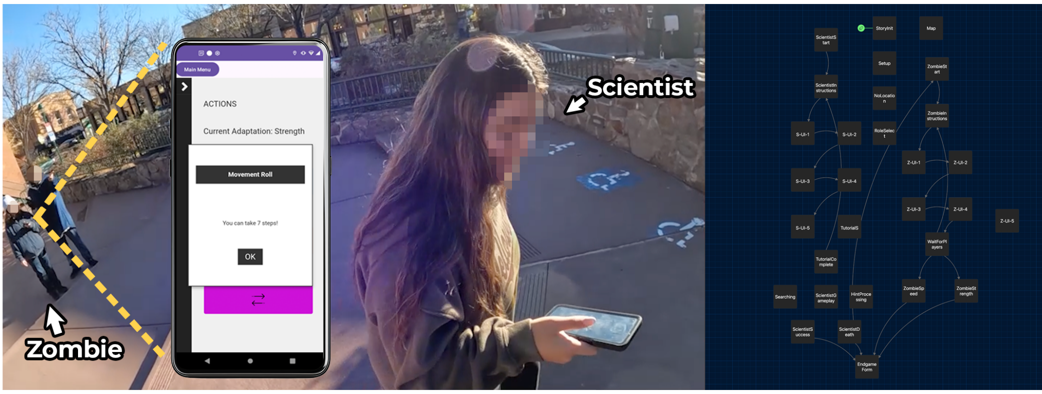


Fig. 4. Participants playing Zombie Apocalypse at the urban location with an enlarged overlay depicting the Zombie's user interface and a birds eye view of the Twine structure to the right.

3.1.4 Speed Test (ST). Represents a Wizard of Oz clone of the FCC Speed Test app. When a user presses a button to initiate a measurement, the resulting metrics are reported to the user in a screen report that mimics what the FCC Speed Test displays. The game ended whenever the player decided they wanted to end.

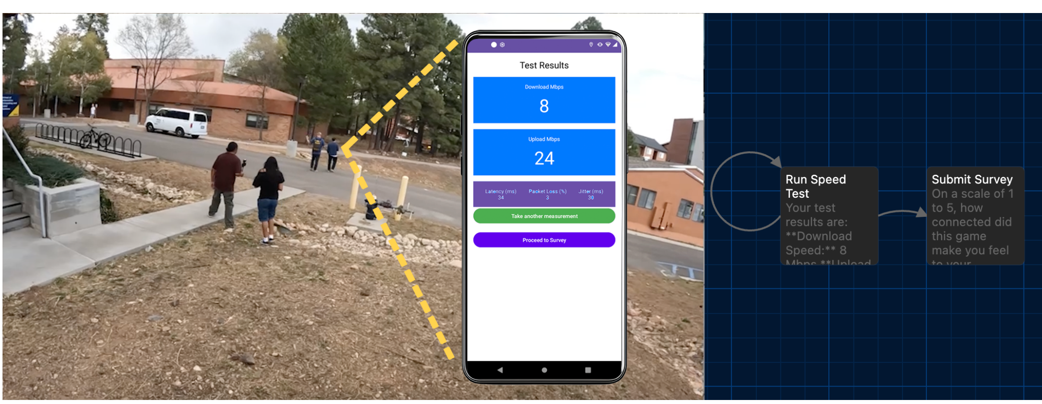


Fig. 5. Participants playing the Speed Test control at the campus location with an enlarged overlay depicting the user interface and a birds eye view of the Twine structure to the right.

4 RITE TESTING

We employed a Rapid Iterative Testing and Evaluation (RITE) approach [61, 70] to refine the design, technology, and research protocol of NetGauge Games. Across four sessions with 13 participants, we gathered feedback and implemented changes to improve usability and data collection. Participant demographics and session sites are detailed in Table 1, and game order was counterbalanced to reduce order effects (Table 2).

We used a provisional coding scheme to develop a codebook for analyzing game play [75], which we piloted and iterated on during RITE sessions prior to the final playtesting sessions. RITE iterations were conducted to ensure that neither technical limitations nor procedural breakdowns would inhibit the final phase of playtesting and data collection. The provisional coding scheme

Site	**Code**	**Age (years)**	**Gender**	**Race and Ethnicity**
Campus	R1A	19	Female	White; Hispanic
	R1B	22	Female	White; Non-Hispanic
	R1C	24	Female	White; Non-Hispanic
	R1D	36	Male	Asian/Pacific Islander; Non-Hispanic
Rural	R2A	20	Female	White; Hispanic
	R2B	20	Male	White; Non-Hispanic
	R2C	24	Female	Asian/Pacific Islander; Non-Hispanic
Urban	R3A	35	Male	Asian/Pacific Islander; Non-Hispanic
	R3B	41	Male	Black; Non-Hispanic
	R3C	Not given	Female	White; Non-Hispanic
Campus	R4A	21	Female	White; Non-Hispanic
	R4B	30	Male	White; Non-Hispanic
	R4C	18	Male	American Indian/Alaska Native; Non-Hispanic

Table 1. RITE participant overview.

Rural	Urban	Campus	Campus
FCC Speed Test	Scavenger	Zombie	Soul Seeker
Soul Seeker	Zombie	FCC Speed Test	Scavenger
Zombie	Soul Seeker	Scavenger	FCC Speed Test
Scavenger	FCC Speed Test	Soul Seeker	Zombie

Table 2. Order of Games at Each Location

was brainstormed collaboratively in Miro, grounded in *RQ2*, and is available here[1] and an export is included in supplementary materials. We used BORIS, an open source event-logging software for coding video and audio [29], to conduct our video analysis. Based on the Miro board, we created a BORIS ethogram [29] to annotate participant behaviors video recorded during gameplay. This scheme encompassed five broad categories, including *Interaction with Others*, which captured moments of collaborative problem-solving, logistical coordination, reflective dialogue, exploratory discussions, and expressions of joy or humor. We also coded instances of *Interaction with the Environment*, such as physical searching, spatial exploration, and responses to navigational or friction encountered during the game. The *Interaction with Technology* category included participant engagement with digital prompts, directional guidance, narrative content, and reflective responses within the game interface. Additionally, we included a *Metacontext* category to denote higher-level structural moments, such as training, transitions between physical locations, and post-game debriefing sessions. Finally, the *Quote* category was used to mark notable verbal expressions by participants that offered insight into their cognitive or emotional engagement. The final BORIS ethogram is provided in Section 5. Iterative changes to the BORIS ethogram based on RITE session outcomes are described in the following subsections alongside iterations to the games, protocol, and platform.

[1] https://miro.com/app/board/uXjVLdo5jC8=/?share_link_id=208996733772

4.1 RITE Session 1

This session took place at the rural site with three participants. Demographics of the participants are presented in Table 1. All participants in this session were selected because they had expertise in networks and information systems, including the collection of network traces, descriptive statistical analysis, network scientific analysis, and wireless network design and development. The rural location was selected because it represents a challenging case study for networks research.

4.1.1 Game Design Adjustments. Testers noted that it was difficult (and for some games, impossible) to play the game without external researcher direction. Thus, for the next iteration, we sought to have games *"speak for themselves"* by providing explicit how-to-instructions before each game. Navigation cues were renamed from "hints" to "request guidance" to signal they were an expected mechanic rather than an optional aid. Additionally, bugs in *Scavenger Hunt* responses were fixed and a feature was added allowing players to view the internet data they had collected.

4.1.2 Platform and Interface Updates. We adjusted Twine's default styling to improve legibility in outdoor settings by changing from light text on dark backgrounds to dark text on light backgrounds. To clarify spatial boundaries, diagrams of the play area were added. Directional language was revised to second person (e.g., "your left") to resolve confusion about phone orientation.

In addition to tester-based feedback, our research team also made adjustments to the platform based on data collection logistics. For instance, we noted that all of the telemetry data for the fourth game had not been sent to the backend due to data rate limitations on the data collection server. To address this, we reduced the GPS sampling rate from once per second to once every five seconds to improve reliability and reduce data loss.

4.1.3 Protocol and Data Collection. Given that we had issues with syncing telemetry data from lab phones to the data collection server, we decided to modify our data collection protocol to save copies of telemetry data locally on phone storage in addition to posting it to a data collection server in case of data loss when synchronizing with the Firebase server.

The initial BORIS ethogram was revised in several ways based on findings from the first RITE session. First, we standardized the "Ask at Stop" setting across all behaviors to ensure raters were prompted consistently, addressing issues of missing modifier data caused by inconsistent settings. Modifier groupings were then refined to streamline the coding experience and improve inter-rater alignment. For *Interact with Others*, we relocated the *Exploration* modifier to *Interact with Environment*, where it better fit the observed behaviors. Two new modifiers, *Debrief* and *Side Conversations*, were added inductively to capture reflective group discussions and informal in-game banter. For *Interact with Environment*, we inductively added the following modifiers: *Exploration*, *Friction*, *Mapping Physical Objects to Game*, *Problem Solving*, *Reflection*, and *Searching*. These additions reflected the complexity of spatial and embodied interaction as participants navigated outdoor play areas. In the *Interact with Technology* category, we added modifiers for *Direction*, *Friction*, *Problem Solving*, *Reflection*, and *Story* to account for both interaction patterns with Twine interfaces and affective narrative engagement. The *Metacontext* behavior was simplified by removing the *Internet Measurement* modifier, as these interactions were already captured automatically via telemetry and were not visible in the recorded video.

Modifier order was also standardized across behaviors to improve coder efficiency and consistency. All behaviors now followed the schema: (1) Metacontext, (2) Participants Involved in Interaction, and (3) Interaction Subtype. This change contributed to an improved coding experience and supported better alignment between raters. Preliminary inter-rater reliability analysis using only the metacontext modifier yielded high agreement ($\kappa = 0.925$) [50]. However, including all modifiers had low reliability ($\kappa = 0.011$) [50].

4.2 RITE Session 2

This session took place at the campus site with three participants. Demographics of the participants are presented in Table 1. All participants for this session were selected for their expertise in Human-Computer Interaction and game design, offering a complementary perspective to the network-focused participants in Session 1.

4.2.1 Game Design Adjustments. Participants informed us that Soul Seeker was far too text intensive and that it was hard to have to read through so much text on the screen at once—even if the plot of the game was interesting. We decided to modify the game by adding visuals to break up chunks of text and to present a maximum of two folds of content, including images and buttons, at once on the screen.

Participants also noted that the Zombie game required more coordination instructions to help players organize themselves into roles and understand the different play instructions for each role. We addressed this by adding a more robust tutorial with a screen for each button that describes what the button does and scenarios it may be useful in. Additionally, we made the buttons clearer with colors and icons corresponding to their function instead of using default Twine hyperlinks. Finally, we added a prompt at the end of the tutorial to check the readiness of all players so everyone has time to read the tutorial and get to their starting locations without rushing.

Participants noticed that the Speed Test control's wizard of oz broadband measurement values were so variable that they did not trust that it was actually measuring broadband. We added logic to the random number generator to output a normal distribution of values within reasonable ranges to avoid breaking player immersion and make it easier for players to suspend their disbelief.

4.2.2 Platform and Interface Updates. We fixed a bug that caused navigation vibration to persist if a game was closed before all points of interest were found. We also widened the radius for vibration intensity feedback, improving player experience when approaching objectives.

4.2.3 Protocol and Data Collection. We noted that using four cameras to capture mobile data across relatively wide areas could be challenging to synchronize for subsequent video data analysis. To help simplify this synchronization, we used a third-party synchronization application that allowed us to centrally control and synchronize all cameras while we were recording play test sessions. We also found that the battery capacity of each of the cameras was not always reliable for the entire play period. For two of the cameras, we needed to swap batteries mid-session, which was disruptive. To account for this, we decided that future sessions should couple cameras to external battery packs.

For RITE Session 1 and 2, we captured connection to people and environment during a debrief after each game, but did not want the order of participant responses to affect others' initial ratings. Therefore, we decided to more quantitatively capture how participants perceived connection by having them complete a survey built into the game platform once each game was completed prior to the qualitative debrief.

4.3 RITE Sessions 3 & 4

RITE Sessions 3 and 4 were held on the same day in consecutive sessions. Session 3 took place at the rural site and Session 4 took place at the urban site. Demographics of the participants are presented in Table 1. Participants for these sessions were re-invited from our previous study [27] so that they could provide perspectives as the original game concept designers.

4.3.1 Game Design Adjustments. During the debrief, participants reflected that the Zombie Apocalypse game left ambiguity as to whether or not the Zombie role had successfully "hit" the player with the scientist role. As a team, we decided to allow for this ambiguity to persist, as it did not fundamentally impede game play or data collection and allowed players to navigate balancing the game's roles.

4.3.2 Platform and Interface Updates. We did not note any major challenges with the platform at either site and confirmed that data appropriately synchronized despite limited and interrupted connectivity in each of the sites where these sessions took place.

4.3.3 Protocol and Data Collection. When all behaviors and modifiers were included there was still low reliability ($\kappa = 0.216$)[50]. One source of disagreement in the coding arose from tagging participants and researchers in interactions and limited direction on how to denote which of the camera frames codes referred to, which is important because participants often separated from one another. To address this, we changed our use of the subject feature in BORIS to tag which of the cameras each code is linked to and researchers and participants were tagged with behavior modifiers. In addition, we developed a guide to accompany video coding, which is available in our supplementary materials. Prior to analyzing the playtest videos, two coders from the larger group volunteered to complete all of the coding and attended a training to improve reliability. For the final coding of playtest sessions, which all used a standardized BORIS ethogram, inter-rater reliability for all behaviors improved significantly to a mean inter-rater reliability $\kappa = 0.739$ ($\sigma = 0.042$), which can be interpreted as substantial agreement [50], detailed further in Section 5.

5 METHODOLOGY

Table 3 provides an overview of the data we collected, the mixed methods used to collect the data, the research question associated with the data, and the type of analysis completed. We describe each of these data and methodologies in detail in this section.

Method	Data	Metric	RQ	Type	Analysis
Telemetry	GPS Route	Geocoordinate Timestamp	RQ1	Quant	Mobility
	Generated Broadband Metrics	Geocoordinate Timestamp	RQ1	Quant	Mobility
Video Coding	Interactions with Environment	Duration	RQ1	Quant	Temporal
		Subjects	RQ2	Quant	Network
		Theme	RQ2	Qual	Thematic
	Interactions with Technology	Duration	RQ1	Quant	Temporal
		Subjects	RQ2	Quant	Network
		Theme	RQ2	Qual	Thematic
	Interactions with People	Duration	RQ1	Quant	Temporal
		Subjects	RQ2	Quant	Network
		Theme	RQ2	Qual	Thematic
	Player Ethnography	Quotes	RQ2	Qual	Thematic
	Overall Game Impression	3 Adjectives	RQ2	Qual	Descriptive
Game Survey	Game-Specific Connection to Place	Likert	RQ2	Qual	Descriptive
	Game-Specific Connection to People	Likert	RQ2	Qual	Descriptive
Pre-Post	Game Ranking	Ordering	RQ2	Quant	Descriptive
	Connection to Place	Likert	RQ2	Qual	Descriptive
	Connection to People	Likert	RQ2	Qual	Descriptive
	Measuring Broadband Interest	Likert	RQ1	Qual	Descriptive

Table 3. Overview of Methodology

5.1 Telemetry

Telemetry is a common method in Games User Research [22] in which data is automatically streamed to a server. The nature of telemetry to produce a large corpus of quantitative data lends itself to identifying usage patterns, but telemetry benefits from paired qualitative methods to provide context [22]. Telemetry systems can provide numerous types of data, including screen interactions, sensor data, and event logs. For this research, we were interested in how participants moved within play spaces over time and in broadband metrics.

5.1.1 Mobility Metrics. To quantify the movement behaviors of participants during game play, we used a suite of mobility metrics derived from spatiotemporal GPS telemetry. This included measures of individual travel patterns (e.g., route length and spatial spread), and group-level spatial proximity patterns. All computations were implemented in Python using the scikit-mobility package [69], availible in our supplementary materials. We first filtered trajectory data by game type, ensuring that comparisons were based on matched gameplay contexts. For each participant trajectory, we calculated the route length, defined as the cumulative straight-line distance between consecutive GPS points. This provided a baseline measure of how far participants physically traveled during each game session, with higher values indicating broader movement through the play space.

Next, we computed the radius of gyration (r_g), a commonly used metric in mobility analysis that quantifies the characteristic spatial range of an individual's movement [34]. It is defined as the root mean square distance between each recorded location and the participant's center of mass (i.e., the average position of all GPS points). This measure provides insight into whether movement was tightly clustered (low r_g) or widely dispersed (high r_g).

We also incorporated two entropy-based metrics that characterize the predictability and diversity of participants' movements in bits. *Real entropy* estimates the theoretical limit of predictability by incorporating both the frequency and ordering of visited locations. Higher real entropy (in bits) indicates more complex and less predictable mobility patterns, while lower values suggest highly routine or stereotyped movements. *Random entropy*, in contrast, assumes that all visited locations are equally likely and independent, providing an upper bound on uncertainty. It thus reflects the maximum diversity a trajectory could achieve if location choices were entirely random. Across these entropy metrics, each additional bit corresponds to a doubling of potential uncertainty, enabling an interpretable scale for comparing mobility complexity across games and sessions. Full mobility statistics for each session are presented in Section 6.1.

5.2 Video Coding

To analyze participants' in-game behaviors and interactions, we employed a structured video coding process using the Behavioral Observation Research Interactive Software (BORIS)[29]. Each playtest session was recorded using multiple camera angles to capture spatially distributed activity in large outdoor play spaces. The camera angles were synchronized using timestamps in Adobe Premiere to create a four-frame compiled video used in BORIS. These recordings were annotated in BORIS using a standardized ethogram (Table 4) developed iteratively during RITE testing (see Section 4). Two researchers from the study team received dedicated training in the coding scheme and completed all final behavioral coding for the playtest sessions. A custom video coding guide and walkthrough was also developed and is included in the supplementary materials. To calculate inter-rater reliability (IRR) using Cohen κ, we set the sampling rate to the standard deviation of aggregated event durations, shown in Table 5. The overall mean IRR for all behaviors was $\kappa = 0.739$ ($\sigma = 0.042$), which can be interpreted as substantial agreement [50]. Fine-grain IRR is available in Tables 6 (excluding modifiers) and 7 (including modifiers). All coded behavioral data was exported to CSV format for subsequent analysis (see Section 6).

Behavior	**Modifier**	**Modifier Values**
Interact with others (i)	Metacontext of Interaction	FCC-Control (F) Scavenger Hunt (H) Soul Seeker (S) Zombie (Z) Consent (C) Moving to Play Space (M) Warmup (W)
	Participants Involved In Interaction	Researcher ID (list selection) Participant ID (list selection)
	Type of Interaction	Debrief (d) Friction (f) Training / Logistics (g) Moment of joy or laughter (j) Reflection (r) Problem Solving (s) Roleplaying (p) Side conversations (c)
Interact with Environment (e)	Metacontext Of Interaction	FCC-Control (F) Scavenger Hunt (H) Soul Seeker (S) Zombie (Z) Consent (C) Moving to Play Space (M) Warmup (W)
	Participants Involved In Interaction	Researcher ID (list selection) Participant ID (list selection)
	Type of Interaction	Exploration (e) Friction (f) Mapping Physical Objects to Game (m) Problem Solving (p) Reflection (r) Searching (s)
Interaction with Technology(t)	Metacontext Of Interaction	FCC-Control (F) Scavenger Hunt (H) Soul Seeker (S) Zombie (Z) Consent (C), Moving to Play Space (M) Warmup (W)
	Participants Involved In Interaction	Researcher ID (list selection) Participant ID (list selection)
	Type of Interaction	Direction (d) Friction (f) Problem Solving (p) Reflection (r) Story (s) Internet Measurement (i)
Metacontext (m)	Metacontext Of Interaction	FCC-Control (F) Scavenger Hunt (H) Soul Seeker (S) Zombie (Z) Consent (C) Moving to Play Space (M)
Quote (q)	Metacontext Of Interaction	FCC-Control (F) Scavenger Hunt (H) Soul Seeker (S) Zombie (Z) Consent (C Moving to Play Space (M) Warmup (W)
	Participants Involved In Interaction	Researcher ID (list selection) Participant ID (list selection)
	QuoteID	ID linked to external spreadsheet

Table 4. Final BORIS Ethogram. Hot keys are shown within parenthesis.

Session	**Sampling Rate (s)**	**Excluding Modifiers (κ)**	**Including Modifiers (κ)**
1	128	0.757	0.604
2	172	0.729	0.63
3	294	0.783	0.708
4	213	0.685	0.558
Mean	**201** (σ = 70)	**0.739** (σ = 0.042)	**0.625** (σ = 0.063)

Table 5. IRR for all behaviors. κ was calculated using a sampling rate, in seconds, based on the standard deviation (σ) of state event durations.

Session	**Sampling Rate (s)**	**Others (κ)**	**Tech (κ)**	**Enviornment (κ)**
1	128	0.784	0.693	0.635
2	172	0.859	0.837	0.744
3	294	0.734	0.914	0.775
4	213	0.691	0.812	0.711
Mean	**202** (σ = 71)	**0.767** (σ = 0.072)	**0.814** (σ = 0.092)	**0.716** (σ = 0.060)

Table 6. IRR of individual behaviors excluding modifiers. κ was calculated using a sampling rate, in seconds, based on the standard deviation (σ) of state event durations.

Session	**Sampling Rate (s)**	**Others (κ)**	**Tech (κ)**	**Environment (κ)**
1	128	0.375	0.666	0.547
2	172	0.54	0.846	0.751
3	294	0.637	0.918	0.638
4	213	0.398	0.769	0.546
Mean	**202** (σ = 71)	**0.488** (σ = 0.124)	**0.800** (σ = 0.108)	**0.621** (σ = 0.097)

Table 7. IRR of individual behaviors, including modifiers. κ was calculated using a sampling rate, in seconds, based on the standard deviation (σ) of state event durations.

5.3 Game Survey

After completing each game, participants were presented with a short survey embedded directly within the game platform. This post-game instrument was designed to capture players' sense of connection both to the environment in which the game took place and to the other participants. The survey included two Likert-scale questions: "How connected did this game make you feel to the spatial environment?" and "How connected did this game make you feel to other players?" For each question, participants selected from five ordinal response options: *Not connected at all*, *Not connected*, *Neutral*, *Connected*, and *Very connected*. Embedding the survey within the game ensured that responses were collected immediately after gameplay, minimizing recall bias and aligning the reflection process with the emotional and cognitive state of participants. These measures were later analyzed in relation to behavioral, mobility, and temporal metrics to examine how game structure influenced perceived environmental and social connection (see Section 6).

5.4 Pre and Post-Survey

In addition to per-game feedback, participants completed brief surveys at the beginning and end of each session. These surveys were designed to contextualize gameplay responses by capturing baseline familiarity with mobile games and platforms, as well as overall impressions across the full play session. For example, one item asked participants to report their interest in measuring broadband internet to see if the playtest had an effect. At the conclusion of the full session, participants completed a post-survey asking them to rank the games in order of personal preference. This ranking provided a comparative measure of overall experience and allowed us to triangulate game-specific engagement metrics with holistic user satisfaction. Responses were analyzed descriptively and used to inform design implications presented in Section 6.3.

5.5 Post-Game Debrief Discussions

Immediately following each game, participants engaged in a short debrief session facilitated by a member of the research team. These sessions served to elicit qualitative reflections on gameplay experience and expand upon the Likert-scale questions administered within the game platform. Participants responded to each question in rotating order with a different participant starting each time. Participants were first asked to share three adjectives that best described their experience with the game. These descriptors were collected in real-time and later visualized as weighted treemaps to surface patterns in emotional tone, engagement, and perceived challenge across games and sites. Following, participants responded to two open-ended questions designed to parallel the Likert items: “In what ways did this game make you feel connected or disconnected from the spatial environment?” and “How did this game affect your sense of connection to the other players?” These open-ended reflections provided qualitative insight into players’ affective and social experiences, offering depth and nuance to complement the quantitative survey results. Debrief data was extracted within BORIS using the “Quote” behavior within the ethogram (Section 5.2). Results are presented alongside survey and behavioral data in Section 6.3.

5.6 Exploratory Correlation Analysis

To explore relationships between all collected variables, including mobility metrics, survey responses, telemetry, and video coding, we conducted an exploratory correlation analysis using Spearman’s rank-order correlation coefficient (ρ). This nonparametric method was chosen due to the small sample size and the presence of non-normal distributions among several measures. All statistical computations were performed using SPSS. For each pairwise combination of variables, we computed Spearman’s ρ and corresponding two-tailed significance values. Only correlations reaching statistical significance ($p < .05$) are reported in relevant sections of the Results. All correlation coefficients, including non-significant findings, are available in the supplemental materials.

5.7 Participants

We recruited a total of 12 participants for playtesting by posting research recruitment posters on bulletin boards around campus and sending email recruitment announcements to list-servs, including the Flagstaff Public Library and the Northern Arizona University student clubs list-servs. 58% identified as female, 33% as male, and 8% as non-binary or other. Participants were on average 20.91 years old ($\sigma = 2.53$ years). Participants were compensated with a $25 Amazon gift card at the conclusion of the study. We report on the demography of participants in Table 8.

Site	Code	Age (years)	Gender	Race and Ethnicity
	E1A	24	Female	Asian/Pacific Islander; Non-Hispanic
Rural	E1B	20	Male	American Indian/Alaska Native; Non-Hispanic
	E1C	22	Female	Asian/Pacific Islander; Non-Hispanic
	E2A	22	Male	White; Non-Hispanic
Campus	E2B	18	Female	American Indian/Alaska Native; Non-Hispanic
	E2C	24	Non-binary/third gender	White; Non-Hispanic
	E3A	20	Male	White; Non-Hispanic
Campus	E3B	19	Female	American Indian/Alaska Native; Non-Hispanic
	E3C	18	Female	Asian/Pacific Islander; Non-Hispanic
	E4A	26	Male	Asian/Pacific Islander; Non-Hispanic
Urban	E4B	19	Female	White; Hispanic
	E4C	19	Female	White; Hispanic

Table 8. Playtest participant overview.

5.8 Ethics

This study was approved by the Institutional Review Board (IRB) at Northern Arizona University (IRB #1860713). All participants provided informed consent prior to participation and were briefed on the nature of the research, their rights as participants, and how their data would be used. Participants were informed that they could withdraw at any time without penalty. To protect participant privacy during gameplay and recording, each participant was assigned a pseudonymous code, and all data, including survey responses, GPS traces, and video recordings, were stored on encrypted drives accessible only to the research team. Participants were compensated for their time with a $25 gift card per session. In addition, we provided snacks and drinks to participants in recognition of the time and energy required to participate. We ensured that all sites were accessible via public transportation and would be generally accessible in case of restricted mobility.

5.8.1 Positionality. Because the positionality and presentation of researchers fundamentally shape the dynamics of field-based user studies and design-oriented evaluations, we briefly reflect on our roles in this work. The RITE sessions and playtest sessions were facilitated by members of a multidisciplinary research team including Vigil-Hayes and Duval, who are faculty researchers with longstanding commitments to participatory design and community-engaged research. Vigil-Hayes brings over a decade of experience examining digital inequities and co-designing interventions in rural and Indigenous contexts. Duval has a decade of experience leading participatory research in serious games and playful health technologies. Cole Pendergraft (Masters student) served as the lead developer, responsible for implementing the technical infrastructure. Taylen Johnson (First-year undergraduate) led the development and design for the Twine-based games. Hunter Beach (Third-Year undergraduate), Chaithanya Heblikar (Recent Masters graduate), and Alison Graham (Third-Year undergraduate) participated in the RITE video coding process. Olivia Vester (Third-Year undergraduate) participated in both the RITE video coding process and, alongside Pendergraft, led the final video coding for the full playtest sessions. Shelby Hagemann (PhD Student), Md Nazmul Hossain (PhD Candidate), and Varna Gobbur (Recent Masters graduate) extracted quotes from videos. Five members of the team identify as male, and six identify as female. Across the team, we adopted an intersectional feminist approach to design, emphasizing collaboration, reflexivity, inclusive environments, and reduction of researcher-participant hierarchies.

6 RESULTS

We organize our results around answers to our research questions. First, we report on how NetGauge performed in its utilitarian goal of crowdsourcing broadband measurements across time and space (*RQ1*) in Section 6.1. Then, we report how NetGauge Games impacted players' connection to each other and to space (*RQ2*) in Section 6.2. To contextualize statistical analyses, we briefly summarize the study design: Each participant played four different location-based games in one session (within-subjects). We ran four different sessions that varied the geographic context and order of game play (between-subjects). This mixed design allowed us to examine how mobility patterns and number of broadband measurements differed across games as well as how these effects interacted with the physical and social environment.

For *RQ1* in Section 6.1 we examined the following six metrics: number of broadband measurements, straight line distance, radius of gyration, random entropy, and real entropy. We conclude *RQ1* with an overview of area coverage. For *RQ2* in Section 6.2, we begin by presenting our temporal analysis via video coding, which is organized by each game. Then we examined the following two metrics: Likert survey responses, and game rankings. For each metric we analyzed for *RQ1* and *RQ2*, we begin by presenting the omnibus statistical analysis and descriptive statistics, then we share correlational analysis with participant reflections, and conclude with a brief statement of key takeaways. Finally, we report on players' overall perspectives and experiences on the games as a holistic experience in Section 6.3.

6.1 Impact on Measurement and Spatial Mobility (*RQ1*)

Measure	Game	Median [IQR]	Mean $\pm\sigma$	Omnibus Test	p	Effect Size
Number of Broadband Measures	ZA	0.0 [13]	10.58 ± 6.857	$F(3, 24) = 9.94$	**<.001**	$\eta_p^2 = .55$
	SS	34.5 [24]	34.5 ± 3.864			
	ST	5.5 [46]	23.75 ± 8.221			
	SH	6.5 [5]	8.33 ± 1.202			
Straight-Line Distance (km)	ZA	0.197 [0.225]	0.285 ± 0.067	$F(3, 24) = 6.87$	**.002**	$\eta_p^2 = .46$
	SS	0.302 [0.186]	0.348 ± 0.049			
	ST	0.064 [0.365]	0.170 ± 0.063			
	SH	0.226 [0.336]	0.299 ± 0.057			
Radius of Gyration (km)	ZA	0.012 [0.008]	0.0167 ± 0.003	$F(3, 24) = 11.31$	**<.001**	$\eta_p^2 = .59$
	SS	0.025 [0.014]	0.0222 ± 0.002			
	ST	0.009 [0.045]	0.0196 ± 0.007			
	SH	0.013 [0.029]	0.0243 ± 0.004			
Real Entropy (bits)	ZA	6.015 [1.769]	6.00 ± 0.967	$F(3, 6) = 42.34$	**<.001**	$\eta_p^2 = .96$
	SS	6.346 [0.874]	6.395 ± 0.704			
	ST	4.270 [3.281]	4.404 ± 1.707			
	SH	6.118 [1.103]	6.101 ± 0.759			
Random Entropy (bits)	ZA	6.887 [0.500]	6.893 ± 0.384	$F(3, 6) = 315.77$	**<.001**	$\eta_p^2 = .99$
	SS	7.483 [0.332]	7.546 ± 0.461			
	ST	4.508 [2.997]	4.654 ± 1.605			
	SH	6.555 [0.969]	6.437 ± 0.609			
Uncorrelated Entropy (bits)	ZA	0.981 [0.039]	0.976 ± 0.006	$F(3, 24) = 21.66$	**<.001**	$\eta_p^2 = .73$
	SS	0.979 [0.023]	0.979 ± 0.004			
	ST	1.000 [0.003]	0.997 ± 0.002			
	SH	0.997 [0.014]	0.993 ± 0.003			

Table 9. Descriptive statistics (median [IQR] and mean $\pm\sigma$) and omnibus test results for the number of broadband measures and mobility outcomes across games, with associated effect sizes.

To answer *RQ1* about how well NetGauge Games improve crowdsourced measurement over space, we report the impact of NetGauge Games on the number of measurements collected (Section 6.1.1) and player mobility through space (Sections 6.1.2-6.1.6). Table 9 provides a high-level overview of telemetry-derived descriptive statistics and omnibus test results for number of internet measurements and mobility across games. We discuss the implications of these results in Section 7.1.

6.1.1 *Number of Measurements Collected.*

Omnibus Analysis and Descriptive Statistics. We report the median and average number of measurements collected per game per user in Table 9, with totals in Figure 6. Across games, the median number of measurements was highest for Soul Seeker (34.5), followed by Scavenger Hunt (6.5), Speed Test (5.5), and Zombie Apocalypse (0.0). Soul Seeker therefore yielded 5.31× as many median measurements as the next highest game (Scavenger Hunt). Zombie Apocalypse's median of 0.0 reflected its asymmetric design, in which only the Scientist role collected data while Zombies did not. As can be seen in Figure 7, only one participant per session acted as the Scientist in Zombie Apocalypse, but the "tag" mechanic could generate large bursts of data. For example, participant E3C recorded 83 measurements within 40.6 minutes of gameplay. Scavenger Hunt, by contrast, produced the most consistent measurement counts across players ($\sigma = 1.2$).

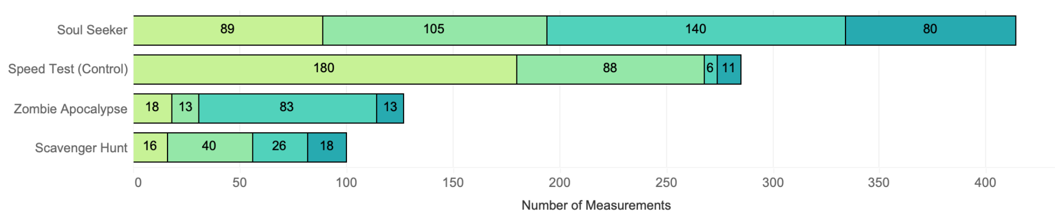


Fig. 6. Aggregate number of measurements per game across 4 sessions

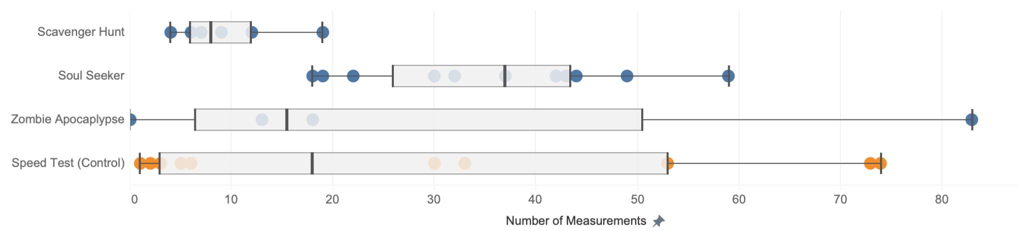


Fig. 7. Distributions of the the number of measurements taken each game

Because three of the four game conditions deviated from normality, as indicated by Shapiro–Wilk tests (Zombie Apocalypse $W = .515$, $p < .001$; Soul Seeker $W = .993$, $p = .413$; Scavenger Hunt $W = .782$, $p = .006$; Speed Test $W = .769$, $p = .004$), we analyzed the number of Internet measurements using a nonparametric aligned rank transform (ART) ANOVA (Table 10) that accommodated the mixed design (*Game* within-subjects; *Session* between-subjects). The test revealed a significant main effect of *Game*, $F(3, 24) = 9.94$, $p < .001$, $\eta_p^2 = .55$, and a significant *Game* × *Session* interaction, $F(9, 24) = 3.60$, $p = .006$, $\eta_p^2 = .57$. The main effect of *Session* was marginal, $F(3, 8) = 3.26$, $p = .081$, $\eta_p^2 = .55$.

ART ANOVA Number Broadband Measures					
Effect	F	df	df_{res}	p	η_p^2
Game	9.94	3	24	**<.001**	.55
Session	3.26	3	8	.081	.55
Game × Session	3.60	9	24	**.006**	.57
Pairwise Comparisons					
Contrast	Estimate	SE	t	p_{adj}	
ZA – SS	-23.00	4.95	-4.64	**<.001**	
ZA – ST	-14.25	4.95	-2.88	**.050**	
ZA – SH	-1.08	4.95	-0.22	1.000	
SS – ST	8.75	4.95	1.77	.540	
SS – SH	21.92	4.95	4.43	**.001**	
ST – SH	13.17	4.95	2.66	.083	

Table 10. Omnibus and pairwise Bonferroni-adjusted ART ANOVA results for number of Internet measurements using Type III Wald F tests with Kenward–Roger df.

Session E1						**Session E2**					
Contrast	Estimate	SE	df	t-ratio	p	Contrast	Estimate	SE	df	t-ratio	p
ZA – SS	0.667	9.36	24	0.071	1.0000	ZA – SS	-8.667	9.36	24	-0.926	1.0000
ZA – ST	-27.000	9.36	24	-2.884	**0.0489**	ZA – ST	-9.667	9.36	24	-1.033	1.0000
ZA – SH	-2.667	9.36	24	-0.285	1.0000	ZA – SH	-15.333	9.36	24	-1.638	0.6868
SS – ST	-27.667	9.36	24	-2.955	**0.0414**	SS – ST	-1.000	9.36	24	-0.107	1.0000
SS – SH	-3.333	9.36	24	-0.356	1.0000	SS – SH	-6.667	9.36	24	-0.712	1.0000
ST – SH	24.333	9.36	24	2.599	0.0944	ST – SH	-5.667	9.36	24	-0.605	1.0000
Session E3						**Session E4**					
Contrast	Estimate	SE	df	t-ratio	p	Contrast	Estimate	SE	df	t-ratio	p
ZA – SS	-13.000	9.36	24	-1.389	1.0000	ZA – SS	3.333	9.36	24	0.356	1.0000
ZA – ST	21.000	9.36	24	2.243	0.2062	ZA – ST	18.000	9.36	24	1.923	0.3987
ZA – SH	-3.667	9.36	24	-0.392	1.0000	ZA – SH	-5.667	9.36	24	-0.605	1.0000
SS – ST	34.000	9.36	24	3.632	**0.0080**	SS – ST	14.667	9.36	24	1.567	0.7816
SS – SH	9.333	9.36	24	0.997	1.0000	SS – SH	-9.000	9.36	24	-0.961	1.0000
ST – SH	-24.667	9.36	24	-2.635	0.0870	ST – SH	-23.667	9.36	24	-2.528	0.1107

Table 11. Simple Game Effects Within Session (Bonferroni-adjusted pairwise contrasts; Kenward–Roger degrees of freedom).

Bonferroni-adjusted contrasts showed that Soul Seeker yielded significantly more Internet measurements than both Zombie Apocalypse ($p < .001$) and Scavenger Hunt ($p = .001$). Zombie Apocalypse produced fewer measurements than Speed Test ($p = .05$). Other comparisons were not significant ($p > .05$).

Simple effects analyses (Table 11) confirmed that differences were driven by Session 1 (Rural) and Session 3 (Campus). In Session 1 (Rural), Speed Test yielded significantly more measurements than both Zombie Apocalypse and Soul Seeker (both $p < .05$). In Session 3 (Campus), Soul Seeker produced significantly more measurements than Speed Test ($p = .008$). By contrast, no reliable differences were found in Session 2 (Campus) or Session 4 (Downtown). Overall, Soul Seeker generated the highest number of Internet measurements on average, followed by Speed Test, while Zombie Apocalypse and Scavenger Hunt yielded fewer measurements. However, given the significant interaction, interpretation of marginal averages should be made cautiously, as the rank ordering of games varied by session. Full descriptive statistics (SPSS), Shapiro–Wilk test (SPSS), and ART ANOVA (R script) are provided in the Supplementary Materials.

Correlational Analysis and Qualitative Reflections. Participants collected broadband measurements in very different ways depending on the game structure. The Speed Test replicated the FCC app, allowing participants to initiate tests at will. This often led to a minimal and repetitive experience, reflected in its median of only 5.5 measurements per player and wide variance (Table 9). One participant summarized the activity bluntly: "*It was horrible. I just walked around and checked out different internet connection measurements throughout the environment*" (E1B, Speed Test). Speed Test reduced measurement to a utilitarian task, lacking intrinsic engagement: "*It wasn't a video game, it was a tool*" (E1A, Speed Test).

The number of measurements collected during Speed Test was strongly and negatively correlated with the order in which the game was played ($\rho = -0.758$, $p = .004$), indicating that participants who encountered Speed Test later in their session took substantially fewer measurements than those who played it earlier. In other words, when positioned after more engaging games like Zombie Apocalypse or Soul Seeker, Speed Test suffered from reduced player effort and attention. By contrast, games like Zombie Apocalypse and Soul Seeker embedded measurement into narrative and role-based mechanics, turning data collection into part of the play experience. As one participant noted:

> "*I feel like I was kind of in my own world a little bit more than the other games just because I really get into the story of video games. It was kind of like immersing yourself in the fake world that it had set up for you. But also trying to guess how the story would end*" (E1A, Soul Seeker).

Mechanics that require proximity checks, exploration, or role-play transform taking measurements into moments of fun and competition, which may explain why Soul Seeker produced over five times more measurements than Scavenger Hunt, and why Zombie Apocalypse, despite asymmetries in who could measure, still generated bursts of intense data collection by scientists under pursuit.

Key Takeaways. Soul Seeker consistently drove the highest and most stable number of measurements, producing over five times as many as Scavenger Hunt and significantly more than Zombie Apocalypse. Scavenger Hunt yielded moderate and consistent counts, while Zombie Apocalypse showed bursts of intense measurement from a single role but little activity otherwise. The Speed Test control produced relatively few measurements and declined sharply when played later in the session, underscoring its lack of engagement. Overall, embedding measurement into narrative, role-based, or exploratory mechanics sustained player effort and yielded richer data than utilitarian testing alone. However, significant *Game* × *Session* interactions highlight that both game mechanics and geographic context shaped measurement outcomes.

6.1.2 Straight Line Distance.

Omnibus Analysis and Descriptive Statistics. We plot the distribution of straight-line distance traveled per player per game across all sessions in Figure 8. We observe that Speed Test had the shortest median straight-line distance (0.06 km), whereas Soul Seeker had the longest median straight-line distance (0.30 km). Compared to Speed Test, Soul Seeker distances were nearly five times longer on the median scale. Importantly, Speed Test exhibited the widest spread of distances (IQR = 0.34 km), indicating more variability across players than for the other games.

Because three of the four game conditions deviated from normality, as indicated by Shapiro–Wilk tests (Zombie Apocalypse $W = .77$, $p = .004$; Soul Seeker $W = .84$, $p = .026$; Scavenger Hunt

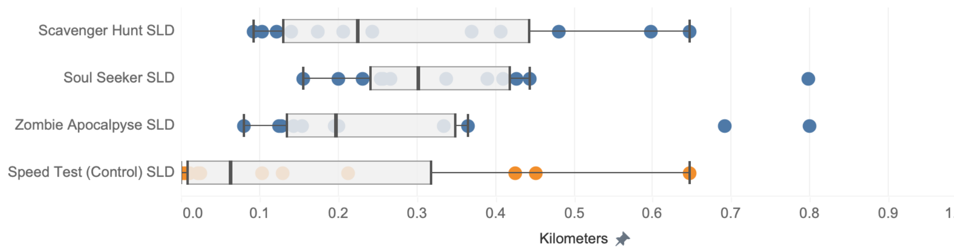


Fig. 8. Distributions of the straight-line distance associated with each game.

$W = .79$, $p = .006$; Speed Test $W = .89$, $p = .108$), we analyzed straight-line distance using a nonparametric aligned rank transform (ART) ANOVA that accommodated the mixed design. The omnibus tests revealed a significant main effect of *Game*, $F(3, 24) = 6.87$, $p = .002$, $\eta_p^2 = .46$, a significant main effect of *Session*, $F(3, 8) = 18.63$, $p < .001$, $\eta_p^2 = .87$, and a robust *Game* × *Session* interaction, $F(9, 24) = 6.22$, $p < .001$, $\eta_p^2 = .70$. Both the game being played and the session context strongly influenced the distances that participants traveled, with game effects varying across sessions.

Pairwise comparisons averaged across sessions (Bonferroni-adjusted) showed that Speed Test elicited significantly shorter distances than Zombie Apocalypse ($p = .041$), Soul Seeker ($p = .001$), and Scavenger Hunt ($p = .021$). No reliable differences were observed among Zombie Apocalypse, Soul Seeker, and Scavenger Hunt (all $p > .10$). Effect sizes for the significant contrasts were moderate-to-large ($r \approx .50$–$.70$). However, given the significant interaction, interpretation of marginal comparisons should be made cautiously.

Simple effects analyses within each session revealed that the pattern of differences varied across location. In Session 1 (Rural), Speed Test distances were significantly shorter than Zombie Apocalypse, Soul Seeker, and Scavenger Hunt (all $p < .005$). In Session 2 (Campus), no significant differences emerged among games (all $p = 1$). In Session 3 (Campus), results closely mirrored Session 1 (Rural): Speed Test distances were again significantly shorter than Zombie Apocalypse, Soul Seeker, and Scavenger Hunt (all $p < .05$). In Session 4 (Downtown), no reliable differences were observed. Speed Test consistently produced the shortest travel trajectories, while Soul Seeker and Zombie Apocalypse tended to elicit longer distances, although the strength of these differences was session-dependent. Straight line distance calculations (Python script using SciKit), Shapiro–Wilk normality tests (SPSS), ART ANOVA computations (R script), and Bonferroni-adjusted pairwise comparisons (R script) are provided in the supplementary materials.

Correlational Analysis and Qualitative Reflections. Straight line distance measures revealed several significant behavioral and spatial patterns across different conditions, such as Speed Test's distance showing an exceptionally strong positive correlation with Speed Test's total duration ($\rho = 0.979$, $p < 0.001$). Movement unpredictability was also highly correlated with distance, as evidenced by the strong relationship between Scavenger Hunt distance and Scavenger Hunt Distance real entropy ($\rho = 0.965$, $p < 0.001$), suggesting that participants who traveled greater straight-line distances also exhibited more unpredictable and variable movement patterns. Participants' reflections illustrate how game mechanics drove these relationships. One described the pursuit of clues in Zombie Apocalypse: "*As a zombie, my goal was to get closer to everybody*" (E2A). Another reflected on the Scavenger Hunt:

ART ANOVA Straight Line Distance				
Effect	F	df	df_{res}	p
Game	6.87	3	24	**.002**
Session	18.63	3	8	**<.001**
Game × Session	6.22	9	24	**<.001**
Pairwise Comparisons				
Comparison	Estimate	SE	t	p_{adj}
ZA – SS	-7.33	5.34	-1.37	1.000
ZA – ST	15.83	5.34	2.96	**.041**
ZA – SH	-1.50	5.34	-0.28	1.000
SS – ST	23.17	5.34	4.34	**.001**
SS – SH	5.83	5.34	1.09	1.000
ST – SH	-17.33	5.34	-3.24	**.021**

Table 12. Straight line distance ART ANOVA and pairwise Bonferroni-adjusted results for using Type III Wald F tests with Kenward–Roger df.

Session E1						**Session E2**					
Contrast	Estimate	SE	df	t-ratio	p-value	Contrast	Estimate	SE	df	t-ratio	p-value
ZA – SS	-6.000	8.040	24	-0.746	1.000	ZA – SS	-9.667	8.040	24	-1.203	1.000
ZA – ST	-37.333	8.040	24	-4.645	**< .001**	ZA – ST	-2.000	8.040	24	-0.249	1.000
ZA – SH	3.333	8.040	24	0.415	1.000	ZA – SH	-7.000	8.040	24	-0.871	1.000
SS – ST	-31.333	8.040	24	-3.898	**.004**	SS – ST	7.667	8.040	24	0.954	1.000
SS – SH	9.333	8.040	24	1.161	1.000	SS – SH	2.667	8.040	24	0.332	1.000
ST – SH	40.667	8.040	24	5.059	**< .001**	ST – SH	-5.000	8.040	24	-0.622	1.000
Session E3						**Session E4**					
Contrast	Estimate	SE	df	t-ratio	p-value	Contrast	Estimate	SE	df	t-ratio	p-value
ZA – SS	6.333	8.040	24	0.788	1.000	ZA – SS	-2.667	8.040	24	-0.332	1.000
ZA – ST	32.000	8.040	24	3.981	**.003**	ZA – ST	0.667	8.040	24	0.083	1.000
ZA – SH	-1.000	8.040	24	-0.124	1.000	ZA – SH	-3.333	8.040	24	-0.415	1.000
SS – ST	25.667	8.040	24	3.193	**.023**	SS – ST	3.333	8.040	24	0.415	1.000
SS – SH	-7.333	8.040	24	-0.912	1.000	SS – SH	-0.667	8.040	24	-0.083	1.000
ST – SH	-33.000	8.040	24	-4.105	**.002**	ST – SH	-4.000	8.040	24	-0.498	1.000

Table 13. Straight line distance simple game effects within session (Bonferroni-adjusted pairwise contrasts; Kenward–Roger degrees of freedom).

> "*I think it was pretty interactive. It made me really like look into the things that were in the environment. I walked around here like everyday for two years and I [pointing at the direction of the object] didn't even know that was there*" (E1B).

Following and searching behaviors pushed players to take more varied and less predictable routes, resulting in both increased distances and higher movement entropy. In contrast, one player summarized the Speed Test as, "*Walked around and checked out different internet measurements*" (E2A), capturing the shorter and more structured paths of the control condition.

Key Takeaways. Across conditions, Speed Test consistently produced the shortest travel trajectories, while Soul Seeker and Zombie Apocalypse elicited the longest routes, with Scavenger Hunt falling in between. Soul Seeker distances were nearly five times longer than Speed Test on the median scale, though differences varied by session context. Longer distances were strongly tied to session duration and movement unpredictability, reinforcing that exploratory or pursuit-based mechanics (e.g., clue-following, chasing) pushed players into more diverse and less predictable paths. By contrast, Speed Test's utilitarian structure constrained movement, yielding shorter and more repetitive routes.

6.1.3 Radius of Gyration.

Omnibus Analysis and Descriptive Statistics. We plot the distribution of the radius of gyration per player per game across all sessions in Figure 9. Median values were small overall (reflecting tightly bounded movement), with Speed Test yielding the shortest median radius of gyration (0.009km) and Soul Seeker the largest (0.025km). Speed Test exhibited the widest dispersion, with an interquartile range (IQR) of 0.045. In contrast, Zombie Apocalypse is tightly clustered near .01 km with some outliers.

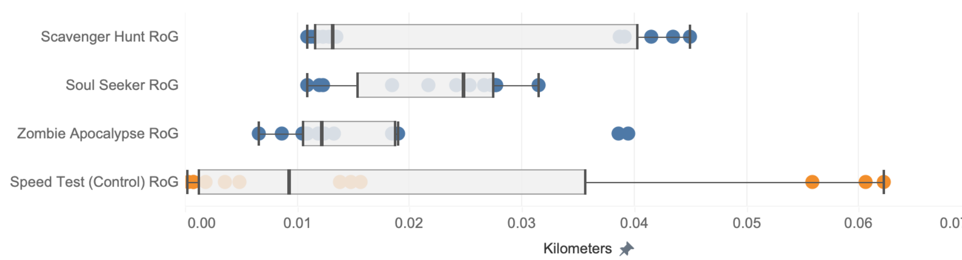


Fig. 9. Distributions of radius of gyration associated with each game

Because three of the four game conditions deviated from normality, as indicated by Shapiro–Wilk tests (Zombie Apocalypse $W = .74$, $p = .002$; Soul Seeker $W = .88$, $p = .087$; Scavenger Hunt $W = .723$, $p = .001$; Speed Test $W = .73$, $p = .002$), we analyzed radius of gyration using a nonparametric aligned rank transform (ART) ANOVA that accommodated the mixed design. The omnibus tests revealed a significant main effect of *Game*, $F(3, 24) = 11.31$, $p < .001$, $\eta_p^2 = .59$, a significant main effect of *Session*, $F(3, 8) = 20.83$, $p < .001$, $\eta_p^2 = .89$, and a robust *Game* × *Session interaction*, $F(9, 24) = 13.69$, $p < .001$, $\eta_p^2 = .84$. Both the game being played and the session context strongly influenced participants' dispersion of movement, with game effects varying across sessions.

Pairwise comparisons averaged across sessions (Bonferroni-adjusted) showed that Soul Seeker elicited significantly higher gyration than Zombie Apocalypse ($p = .011$), while Scavenger Hunt elicited significantly higher gyration than both Speed Test ($p = .002$) and Zombie Apocalypse ($p < .001$). No reliable differences were observed between Speed Test and Soul Seeker ($p > .10$). Effect sizes for these contrasts were large ($r \approx .60$-.85). However, given the significant interaction, interpretation of marginal averages should be made cautiously.

Simple effects analyses confirmed session-dependent patterns. In Session 1 (Rural), Speed Test produced significantly higher gyration than Zombie Apocalypse and Soul Seeker, while Scavenger Hunt was higher than Soul Seeker but lower than Speed Test (all $p < .05$). In Session 2 (Campus), only the Scavenger Hunt vs. Speed Test contrast reached significance ($p = .031$). In Session E3, all three main games showed significantly greater gyration than Speed Test (Speed Test $p < .001$; Soul Seeker $p = .004$; Scavenger Hunt $p < .001$). No other comparisons were significant. Radius of gyration (Python script using ScyKit), Shapiro–Wilk normality tests (SPSS), ART ANOVA computations (R

ART ANOVA Radius of Gyration				
Effect	F	df	df_{res}	p
Game	11.31	3	24	**<.001**
Session	20.83	3	8	**<.001**
Game × Session	13.69	9	24	**<.001**
Pairwise Comparisons				
Comparison	Estimate	SE	t	p_{adj}
ZA – SS	-16.20	4.65	-3.50	**.011**
ZA – ST	-4.50	4.65	-0.97	1.000
ZA – SH	-24.20	4.65	-5.22	**<.001**
SS – ST	11.80	4.65	2.53	0.110
SS – SH	-8.00	4.65	-1.72	0.587
ST – SH	-19.80	4.65	-4.25	**.002**

Table 14. Omnibus and pairwise Bonferroni-adjusted ART ANOVA results for radius of gyration using Type III Wald F tests with Kenward–Roger df.

Session E1						**Session E2**					
Contrast	Estimate	SE	df	t-ratio	p-value	Contrast	Estimate	SE	df	t-ratio	p-value
ZA – SS	-18.67	6.07	24	-3.077	**.031**	ZA – SS	-5.00	6.07	24	-0.824	1.000
ZA – ST	-39.00	6.07	24	-6.430	< **.001**	ZA – ST	7.33	6.07	24	1.209	1.000
ZA – SH	3.00	6.07	24	0.495	1.000	ZA – SH	-11.33	6.07	24	-1.868	0.444
SS – ST	-20.33	6.07	24	-3.352	**.016**	SS – ST	12.33	6.07	24	2.033	0.319
SS – SH	21.67	6.07	24	3.572	**.009**	SS – SH	-6.33	6.07	24	-1.044	1.000
ST – SH	42.00	6.07	24	6.924	< **.001**	ST – SH	-18.67	6.07	24	-3.077	**.031**
Session E3						**Session E4**					
Contrast	Estimate	SE	df	t-ratio	p-value	Contrast	Estimate	SE	df	t-ratio	p-value
ZA – SS	10.00	6.07	24	1.649	0.674	ZA – SS	4.33	6.07	24	0.714	1.000
ZA – ST	33.67	6.07	24	5.550	< **.001**	ZA – ST	20.00	6.07	24	3.297	**.018**
ZA – SH	-4.33	6.07	24	-0.714	1.000	ZA – SH	9.33	6.07	24	1.539	0.822
SS – ST	23.67	6.07	24	3.902	**.004**	SS – ST	15.67	6.07	24	2.583	0.098
SS – SH	-14.33	6.07	24	-2.363	0.159	SS – SH	5.00	6.07	24	0.824	1.000
ST – SH	-38.00	6.07	24	-6.265	< **.001**	ST – SH	-10.67	6.07	24	-1.759	0.548

Table 15. Radius of gyration simple game effects within session (Bonferroni-adjusted pairwise contrasts; Kenward–Roger degrees of freedom).

script), and Bonferroni-adjusted pairwise comparisons (R script) are provided in the supplementary materials.

Correlational Analysis and Qualitative Reflections. As expected, radius of gyration correlates strongly with distance traveled, ($\rho = 0.993$, $p < 0.001$), random entropy ($\rho = 0.921$, $p < 0.001$) and total time ($\rho = 0.965$, $p < 0.001$), as evidenced by Speed Test. Interestingly, radius of gyration also showed a strong positive correlation with Likert results about connection to environment ($\rho = 0.869$, $p < 0.001$). Participant reflections reveal how different games shaped movement spread. In Zombie Apocalypse, one participant explained, "*For me, It felt like a game of tag from my perspective*" (E2A) and the scientist player recalled, "*I have to pay attention to my phone and where it's buzzing and* *where I have to go. But I always had to keep an eye out for those two [zombies]*" (E4C), showing how evasion and pursuit mechanics stretched trajectories across the play space.

Key Takeaways. Radius of gyration results showed that movement dispersion was strongly shaped by both game mechanics and session context. Soul Seeker and Scavenger Hunt generally elicited broader dispersal than Zombie Apocalypse, while Speed Test produced the smallest median gyration but with the widest variability. *Session*-level effects were substantial, with contrasts flipping across rural, campus, and downtown contexts. Importantly, larger gyration values were strongly associated with longer sessions, greater movement unpredictability, and stronger reported connection to the environment. Narrative, exploratory, and pursuit mechanics stretch players' spatial ranges, while utilitarian tasks constrain movement to smaller, more repetitive areas.

6.1.4 Random Entropy.

Omnibus Analysis and Descriptive Statistics. Across sessions, random entropy (Table 10) totals showed that Soul Seeker had the highest overall value (M = 7.55, σ = 0.46), followed by Zombie Apocalypse (M = 6.89, σ = 0.38) and Scavenger Hunt (M = 6.44, σ = 0.61), while Speed Test produced the lowest random entropy (M = 4.65, σ = 1.60), indicating much greater variability and reduced randomness compared to the other games.

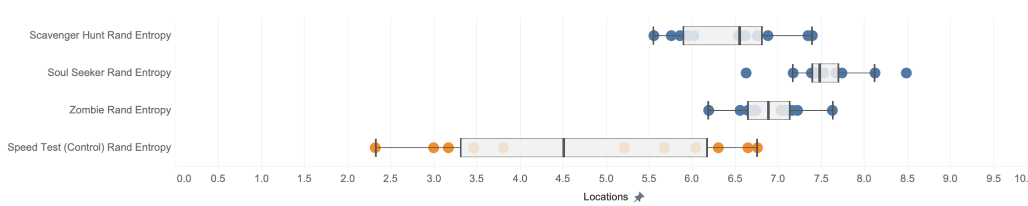


Fig. 10. Distributions of the random entropy associated with the control and each game.

Because none of the four game conditions deviated from normality, as indicated by Shapiro–Wilk tests (Zombie Apocalypse $W = .97$, $p = .85$; Soul Seeker $W = .93$, $p = .39$; Scavenger Hunt $W = .94$, $p = .49$; Speed Test $W = .89$, $p = .13$), we analyzed random entropy using a two-way mixed-design ANOVA (Table 16). Multivariate tests revealed a highly significant main effect of *Game*, Pillai's Trace = .99, $F(3, 6) = 315.77$, $p < .001$, $\eta_p^2 = .99$, as well as a significant *Game* × *Session* interaction, Pillai's Trace = 1.84, $F(9, 24) = 4.25$, $p = .002$, $\eta_p^2 = .61$. Univariate within-subjects tests confirmed a robust main effect of *Game*, $F(3, 24) = 177.70$, $p < .001$, $\eta_p^2 = .96$, and a strong *Game* × *Session* interaction, $F(9, 24) = 31.49$, $p < .001$, $\eta_p^2 = .92$. Between-subjects effects indicated a significant main effect of *Session*, $F(3, 8) = 5.36$, $p = .026$, $\eta_p^2 = .67$.

Post hoc Bonferroni-adjusted pairwise comparisons showed that Speed Test entropy (M = 4.65 bits, 95% CI [4.34, 4.97]) was significantly lower than all other games (all $p < .001$). Soul Seeker (M = 7.55, 95% CI [7.35, 7.74]) was significantly higher than both Zombie Apocalypse (M = 6.89, 95% CI [6.74, 7.04], $p = .005$) and Scavenger Hunt (M = 6.44, 95% CI [6.19, 6.68], $p < .001$). In contrast, Zombie Apocalypse and Scavenger Hunt did not differ significantly from one another ($p = .116$). Random entropy (Python script using ScyKit), Shapiro–Wilk normality tests (SPSS), Mauchly's test of sphericity (SPSS), mixed-design ANOVA (SPSS), and Bonferroni-adjusted pairwise comparisons (SPSS), are provided in the supplementary materials.

Correlational Analysis and Qualitative Reflections. Spatial dispersion measures showed robust associations with random entropy, as evidenced by the strong positive correlation between Speed

Mixed ANOVA Random Entropy				
Effect	*F*	df	*p*	η_p^2
Game	315.77	(3,6)	**<.001**	.99
Session	5.36	(3,8)	**.026**	.67
Game × Session	4.25	(9,24)	**<.002**	.61
Pairwise Comparisons				
Comparison	Estimate	SE	p_{adj}	
ZA – SS	-0.65	0.12	**.005**	
ZA – ST	2.24	0.18	**<.001**	
ZA – SH	0.46	0.16	.116	
SS – ST	2.89	0.10	**<.001**	
SS – SH	1.11	0.08	**<.001**	
ST – SH	-1.78	0.12	**<.001**	

Table 16. Mixed ANOVA and Bonferroni-adjusted pairwise comparisons for random entropy.

Session E1				**Session E2**			
Contrast	Estimate	SE	*p*-value	Contrast	Estimate	SE	*p*-value
ZA – SS	-0.833	0.247	.059	ZA – SS	0.098	0.247	1.000
ZA – ST	0.061	0.360	1.000	ZA – ST	1.517	0.360	**.018**
ZA – SH	0.906	0.313	.120	ZA – SH	0.804	0.313	.199
SS – ST	0.894	0.203	**.014**	SS – ST	1.419	0.203	**<.001**
SS – SH	1.739	0.151	**<.001**	SS – SH	0.707	0.151	**.010**
ST – SH	0.845	0.249	.057	ST – SH	-0.712	0.249	.127
Session E3				**Session E4**			
Contrast	Estimate	SE	*p*-value	Contrast	Estimate	SE	*p*-value
ZA – SS	-0.839	0.247	.057	ZA – SS	-1.034	0.247	**.018**
ZA – ST	4.061	0.360	**<.001**	ZA – ST	3.320	0.360	**<.001**
ZA – SH	0.139	0.313	1.000	ZA – SH	-0.023	0.313	1.000
SS – ST	4.900	0.203	**<.001**	SS – ST	4.354	0.203	**<.001**
SS – SH	0.978	0.151	**.001**	SS – SH	1.010	0.151	**<.001**
ST – SH	-3.922	0.249	**<.001**	ST – SH	-3.344	0.249	**<.001**

Table 17. Random entropy simple game effects within session (Bonferroni-adjusted pairwise contrasts).

Test random entropy and gyration ($\rho = 0.921, p < 0.001$), suggesting that more dispersed movement patterns aligned with higher theoretical randomness. This result is supported by a strong correlation in Speed Test between random entropy and number of broadband measurements taken ($\rho = 0.896$, $p < 0.001$). Random entropy demonstrated strong negative relationships with task ordering, as Speed Test random entropy was negatively correlated with Speed Test game order played ($\rho = -0.898$, $p < 0.001$). Additionally, cross-entropy relationships emerged, with Zombie Apocalypse's random entropy showing strong positive correlations with both real entropy ($\rho = 0.895$, $p < 0.001$) and uncorrelated entropy ($\rho = 0.832$, $p < 0.001$), revealing systematic relationships between baseline theoretical randomness and actual movement entropy measures. Participant quotes reveal insights into why some movement patterns are predictable while others are not: "*Honestly for me I kind of went solo. I did not really pay attention to what the other zombie was doing. I was just trying to get the scientist*" (E1C, Zombie Apocalypse).

Key Takeaways. Random entropy was highest for Soul Seeker, indicating the most unpredictable movement patterns, and lowest for Speed Test, which produced shorter, more repetitive routes. Both Zombie Apocalypse and Scavenger Hunt fell in between, with no significant difference between them. Entropy strongly correlated with broader spatial dispersal and higher measurement counts, suggesting that games encouraging exploration or pursuit led to more varied and less predictable paths. Speed Test entropy decreased sharply when played later in the session, underscoring how utilitarian mechanics were more vulnerable to fatigue and disengagement. In contrast, narrative- or role-based mechanics sustained unpredictability by embedding randomness into the play experience.

6.1.5 Real Entropy.

Omnibus Analysis and Descriptive Statistics. Across sessions, Soul Seeker exhibited the highest overall real entropy (M = 6.40 bits, SD = 0.70). Scavenger Hunt (M = 6.10, SD = 0.76) and Zombie Apocalypse (M = 6.00, SD = 0.97) were similarly high. In contrast, Speed Test displayed the lowest and most variable entropy (M = 4.40, SD = 1.71).

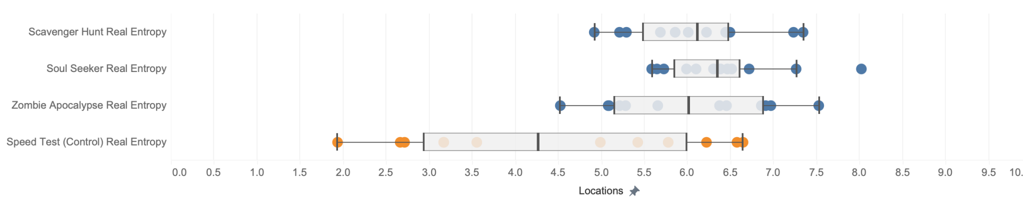


Fig. 11. Distributions of the real entropy associated with the control and each game.

Because none of the four game conditions deviated from normality, as indicated by Shapiro–Wilk tests (Zombie Apocalypse $W = .93$, $p = .37$; Soul Seeker $W = .91$, $p = .19$; Scavenger Hunt $W = .96$, $p = .80$; Speed Test $W = .90$, $p = .16$), we analyzed real entropy using a two-way mixed-design ANOVA. Multivariate tests revealed a significant main effect of *Game*, Pillai's Trace = .96, $F(3, 6) = 42.34$, $p < .001$, $\eta_p^2 = .96$, and a significant *Game* × *Session* interaction, Pillai's Trace = 1.43, $F(9, 24) = 2.41$, $p = .041$, $\eta_p^2 = .48$. Univariate within-subjects tests confirmed a strong main effect of *Game*, $F(3, 24) = 31.98$, $p < .001$, $\eta_p^2 = .80$, and a robust *Game* × *Session* interaction, $F(9, 24) = 13.64$, $p < .001$, $\eta_p^2 = .84$.

Between-subjects effects further indicated a reliable main effect of *Session*, $F(3, 8) = 14.13$, $p = .001$, $\eta_p^2 = .84$. Post hoc pairwise comparisons (Bonferroni-adjusted) showed that Speed Test had significantly lower real entropy than all other games (all $p < .001$), whereas Zombie Apocalypse, Soul Seeker, and Scavenger Hunt did not significantly differ from one another. Estimated marginal means indicated that Soul Seeker (M = 6.40 bits, 95% CI [5.93, 6.86]) and Scavenger Hunt (M = 6.10, 95% CI [5.76, 6.44]) were consistently high, Zombie Apocalypse was moderately high (M = 6.00, 95% CI [5.78, 6.21]), and Speed Test was markedly lower (M = 4.40, 95% CI [4.07, 4.74]). Real entropy (Python script using ScyKit), Shapiro–Wilk normality tests (SPSS), Mauchly's test of sphericity (SPSS), mixed-design ANOVA (SPSS), and Bonferroni-adjusted pairwise comparisons (SPSS), are provided in the supplementary materials.

Correlational Analysis and Qualitative Reflections. Movement unpredictability was strongly linked to spatial dispersion, as demonstrated by the robust positive correlation between Scavenger Hunt real entropy and distance ($\rho = 0.965$, $p < 0.001$). Real entropy in Speed Test was strongly associated with social duration ($\rho = 0.892$, $p < 0.001$). In Zombie Apocalypse, real entropy demonstrated a strong positive correlation with uncorrelated entropy ($\rho = 0.972$, $p < 0.001$), revealing consistency

Mixed ANOVA Real Entropy				
Effect	F	df	p	η_p^2
Game	31.98	(3,24)	**<.001**	.80
Session	14.13	(3,8)	**.001**	.84
Game × Session	13.64	(9,24)	**<.001**	.84
Pairwise Comparisons				
Comparison	Estimate	SE	p_{adj}	
ZA – SS	-0.40	0.22	.643	
ZA – ST	1.59	0.19	**<.001**	
ZA – SH	-0.11	0.20	1.000	
SS – ST	1.99	0.28	**<.001**	
SS – SH	0.29	0.28	1.000	
ST – SH	-1.70	0.14	**<.001**	

Table 18. Mixed ANOVA and Bonferroni-adjusted pairwise comparisons for real entropy.

Session E1				**Session E2**			
Contrast	Estimate	SE	p-value	Contrast	Estimate	SE	p-value
ZA – SS	-1.108	0.440	.215	ZA – SS	0.156	0.440	1.000
ZA – ST	-1.583	0.384	**.020**	ZA – ST	1.207	0.384	.083
ZA – SH	-0.401	0.397	1.000	ZA – SH	0.323	0.397	1.000
SS – ST	-0.475	0.561	1.000	SS – ST	1.051	0.561	.587
SS – SH	0.707	0.560	1.000	SS – SH	0.168	0.560	1.000
ST – SH	1.182	0.282	**.018**	ST – SH	-0.883	0.282	.084
Session E3				**Session E4**			
Contrast	Estimate	SE	p-value	Contrast	Estimate	SE	p-value
ZA – SS	0.138	0.440	1.000	ZA – SS	-0.781	0.440	.681
ZA – ST	4.213	0.384	**< .001**	ZA – ST	2.533	0.384	**.001**
ZA – SH	0.165	0.397	1.000	ZA – SH	-0.508	0.397	1.000
SS – ST	4.075	0.561	**< .001**	SS – ST	3.314	0.561	**.002**
SS – SH	0.027	0.560	1.000	SS – SH	0.273	0.560	1.000
ST – SH	-4.049	0.282	**< .001**	ST – SH	-3.041	0.282	**< .001**

Table 19. Real entropy simple game effects within session (Bonferroni-adjusted pairwise contrasts).

between different entropy measures and movement patterns. Quotes show how mechanics generated cycles of movement. In Zombie Apocalypse, one player described, "*The phone is vibrating and it tells me when I'm close*" (E2C), pointing to system feedback that guided repeated visits. Another emphasized, "*It made me walk around to different spaces in the area*" (E2A), capturing how the game scaffolded cyclical exploration.

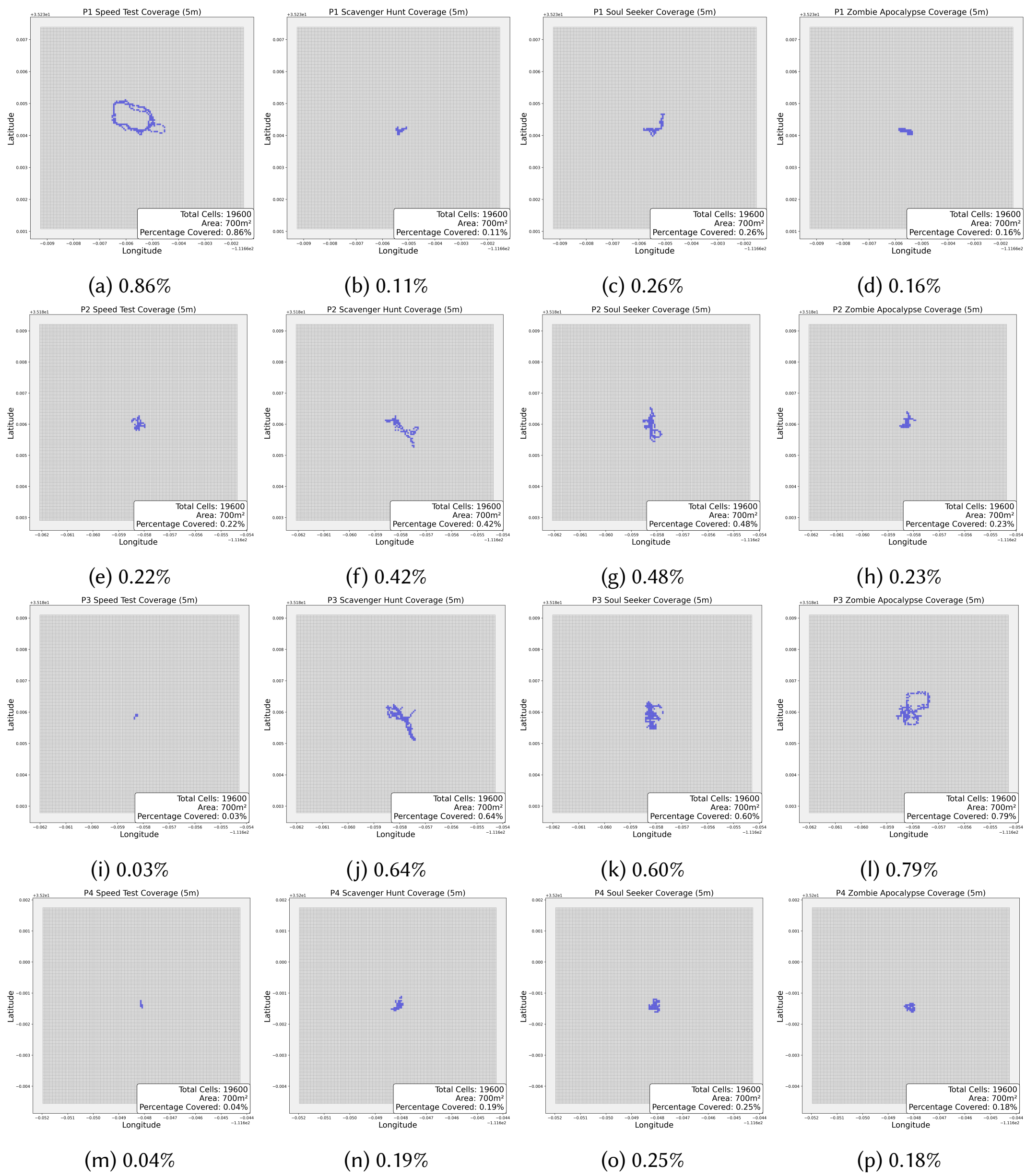
Key Takeaways. Real entropy was significantly higher in Soul Seeker, Scavenger Hunt, and Zombie Apocalypse compared to Speed Test, which produced the lowest and most variable unpredictability. Differences among the three narrative-driven games were small, indicating that each supported similarly complex and non-repetitive movement patterns. Session context shaped these outcomes, with contrasts shifting depending on location. Strong correlations between real entropy, spatial dispersion, and session duration further support how exploratory and pursuit mechanics generated cyclical yet unpredictable movement, while Speed Test constrained players to more repetitive paths. Participant reflections highlight how system feedback and task design in games like Zombie Apocalypse sustained ongoing, varied exploration. Together with random entropy, the real entropy results show that narrative and role-based games sustained higher and more unpredictable movement patterns than Speed Test, reinforcing that unpredictability in play emerges both in theoretical randomness and in actual movement behaviors.

6.1.6 Area Coverage. Using GPS coordinates we collected during play sessions, we plotted the area coverage afforded by each game in each session (E1-E4) by the players in aggregate in Figure 12. To best represent how the games encourage players to move with their phones around an area the size of an H3 resolution 7 area[2], we calculated coverage against a 700m^2 area partitioned by 5m^2 squares, which is the horizontal area at which GPS is accurate under open sky[3]. Thus, in Figure 12, each blue square in a subplot represents a 5m^2 space where at least one player during a session moved through during a session. In general, the players covered more area when engaging with the games than with the Speed Test experience. The exception to this trend was E1, which took place in a rural place that was unfamiliar to most participants and Speed Test was the first game that was played during the session. To illustrate how each locations' geography shaped player movement for each game, we provide maps of each player's movement paths in Figures 13-16.

We observed the highest median coverage values across all games in E2 and E3 which both took place in the campus environment and represented spaces that participants were typically the most familiar with (0.33% and 0.62%, respectively). We observed the highest median coverage value for the Soul Seeker game with a median coverage of 0.37%. This complements other mobility assessments of the games, where Soul Seeker also offered the greatest median values for straight line distance and radius of gyration.

Key Takeaways. Across sessions, games generally expanded spatial coverage well beyond the Speed Test control, with a clear exception in E1 (rural, unfamiliar, and played first) where the control covered comparatively more. Campus sessions (E2–E3) yielded the highest median coverage ($\approx$ 0.33% and $\approx$ 0.62%), reflecting participants' familiarity with the setting. Soul Seeker produced the highest median coverage across games ($\approx$ 0.37%), aligning with its leading medians for straight-line distance and radius of gyration. The grid maps illustrate how geography and mechanics co-produce coverage: pursuit and clue-following dispersed players broadly, while device-centric, utilitarian play concentrated movement into smaller pockets.

[2]The area at which mobile broadband measurements can be used to challenge the FCC National Broadband Map.
[3]https://www.gps.gov/systems/gps/performance/accuracy/



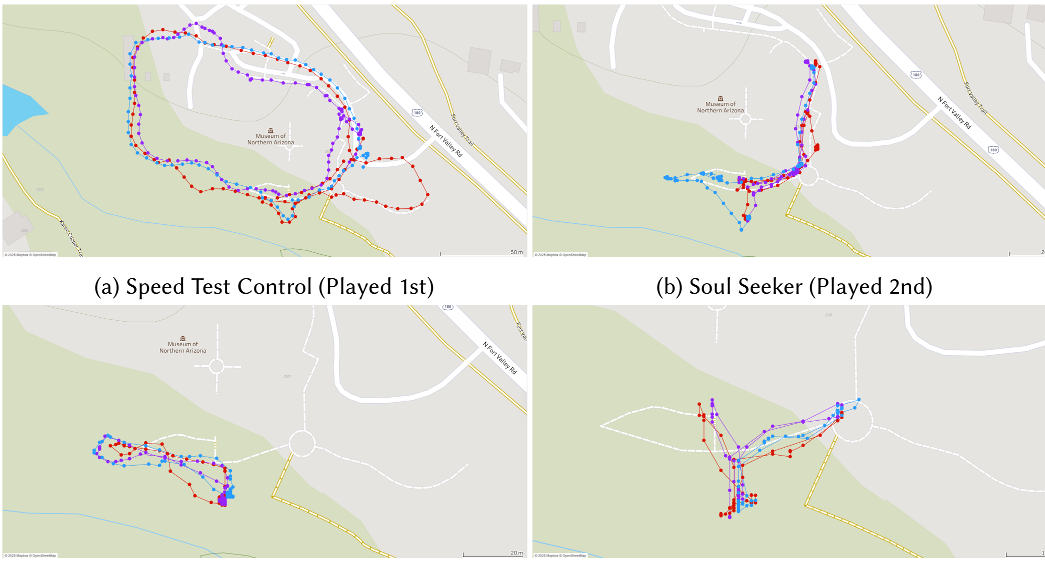
(a) 0.86% (b) 0.11% (c) 0.26% (d) 0.16%

(e) 0.22% (f) 0.42% (g) 0.48% (h) 0.23%

(i) 0.03% (j) 0.64% (k) 0.60% (l) 0.79%

(m) 0.04% (n) 0.19% (o) 0.25% (p) 0.18%

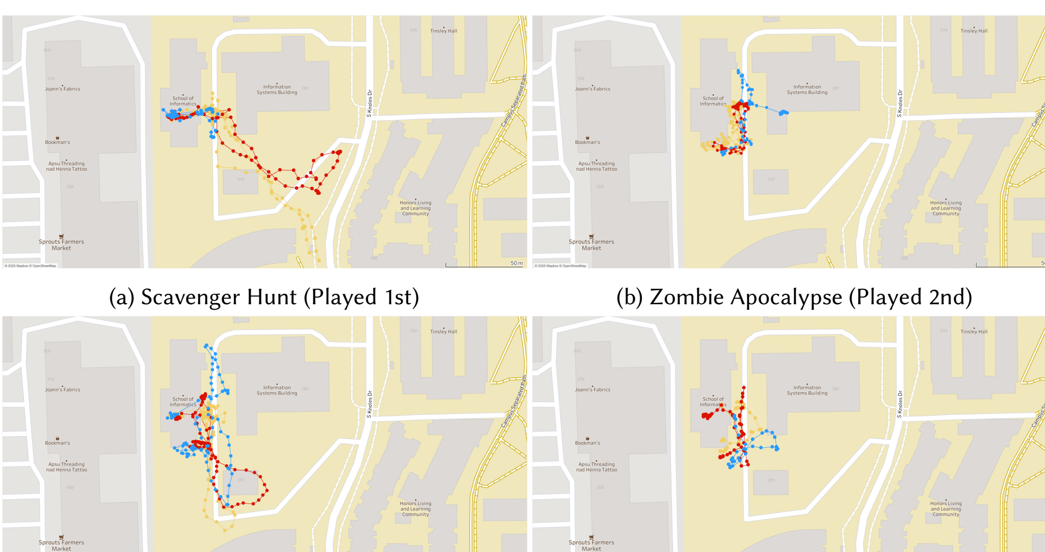
Fig. 12. Area coverage visualizations for all players for all sessions P1-P4 for (a, e, i, m) Speed Test; (b, f, j, n) Scavenger Hunt; (c, g, k, o) Soul Seeker; and (d, h, l, p) Zombie Apocalypse. Resolution is $5m^2$ over a total area of $700m^2$.



(a) Speed Test Control (Played 1st) (b) Soul Seeker (Played 2nd)

(c) Zombie Apocalypse (Played 3rd) (d) Scavenger Hunt (Played 4th)

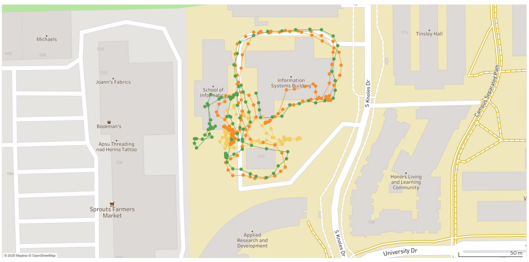
Fig. 13. Session E1 map depicting how players move and around the geographical area for each game (a,b,c,d). Each color is a player, dot is a GPS coordinate, and dots are connected via timestamp to show paths.



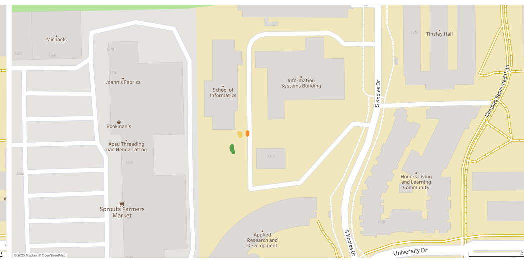
(a) Scavenger Hunt (Played 1st) (b) Zombie Apocalypse (Played 2nd)

(c) Soul Seeker (Played 3rd) (d) Speed Test Control (Played 4th)

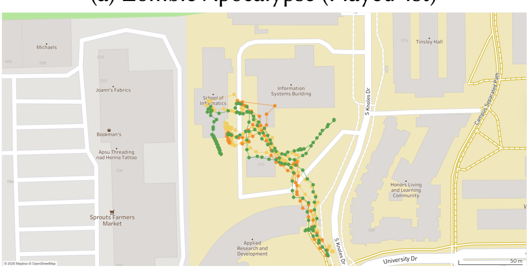
Fig. 14. Session E2 map depicting how players move around the geographical area for each game (a,b,c,d). Each color is a player, dot is a GPS coordinate, and dots are connected via timestamp to show paths.



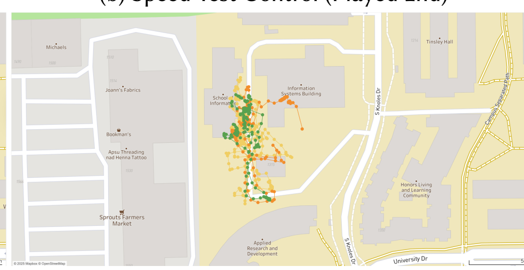
(a) Zombie Apocalypse (Played 1st)



(b) Speed Test Control (Played 2nd)



(c) Scavenger Hunt (Played 3rd)

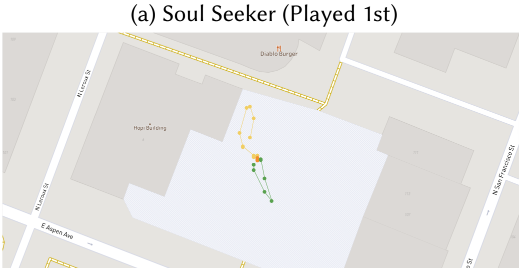


(d) Soul Seeker (Played 4th)

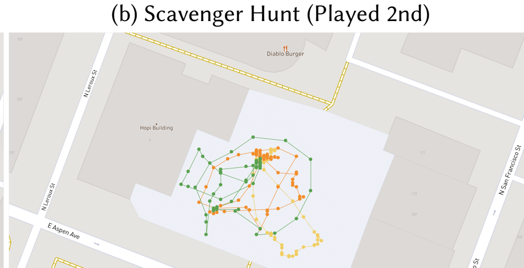
Fig. 15. Session E3 map depicting how players move within the geographical area for each game (a,b,c,d). Each color is a player, dot is a GPS coordinate, and dots are connected via timestamp to show paths.

(a) Soul Seeker (Played 1st)

(b) Scavenger Hunt (Played 2nd)



(c) Speed Test Control (Played 3rd)



(d) Zombie Apocalypse (Played 4th)

Fig. 16. Session E4 map depicting how players move within the geographical area for each game (a,b,c,d). Each color is a player, dot is a GPS coordinate, and dots are connected via timestamp to show paths.

6.2 Community among players and environment (*RQ2*)

To answer *RQ2* about how well NetGauge Games impacted players' connection to each other and to space, we begin by reporting the overall outcomes of our video coding (Section 6.2.1), which provides thematic insight into how players interact with the environment, technology, and other people. Then, we discuss results specific to players' connections to each other, and how game affordances impacted these in Section 6.2.2. Next, we discuss results specific to player's connection with space, and how game affordances impacted these in Section 6.2.3. We discuss the implications of these results in Section 7.2.

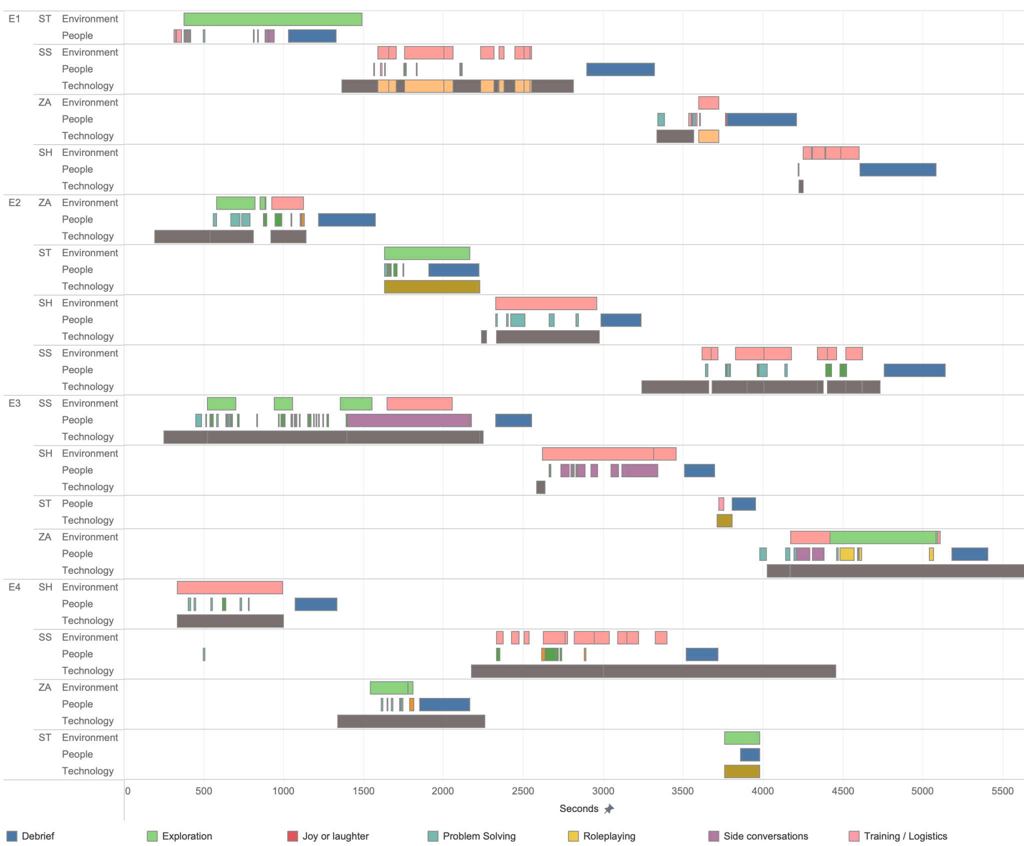


Fig. 17. Overview of video coding

6.2.1 Video Coding.

Figure 17 provides an overview of coded video interactions across all four games, segmented by environment, people, and technology themes. Below, we describe the thematic analysis of each game and include game-specific video coding figures that are more fine-grained and colorblind friendly.

Scavenger Hunt. The video coding in Figure 18 shows Scavenger Hunt as an outward-facing, place-seeking activity: long stretches of *Interact with Environment* punctuated by brief glances to the device to confirm clue progress, with little *Interact with Others*. This rhythm matches the game's

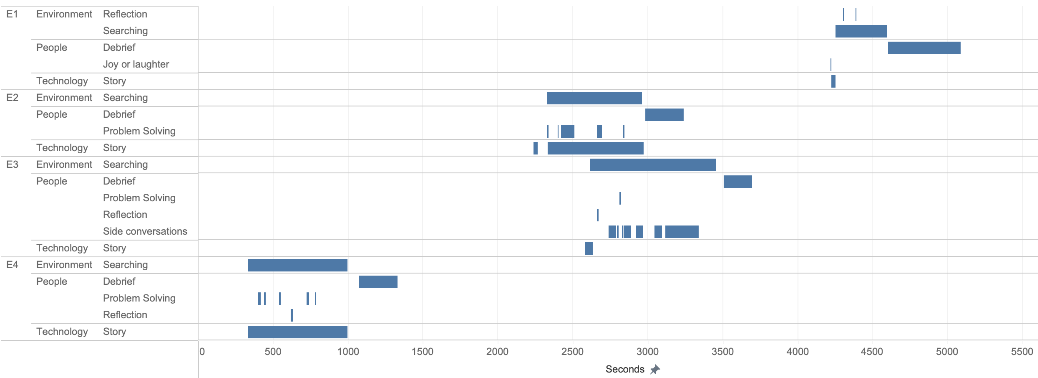


Fig. 18. Overview of Scavenger Hunt video coding

clue-following loop—heads-up scanning and walking dominate, while phone use is instrumental rather than focal. Longer sessions tracked stronger place connection, suggesting that time spent hunting translated into situated noticing (ρ = .730, p = .0255). Participants' accounts mirror these coded patterns of environmental attention and low peer interaction:

> "*Even if the place is very familiar, it was really hard for me to find out the things and all, and it was actually very interesting to find out what exactly and all, and this game makes the place even more familiar for me to understand, and it's actually very interesting.*" (E4A).
>
> "*There wasn't any, like, connection. I just played it by myself*" (E4B).

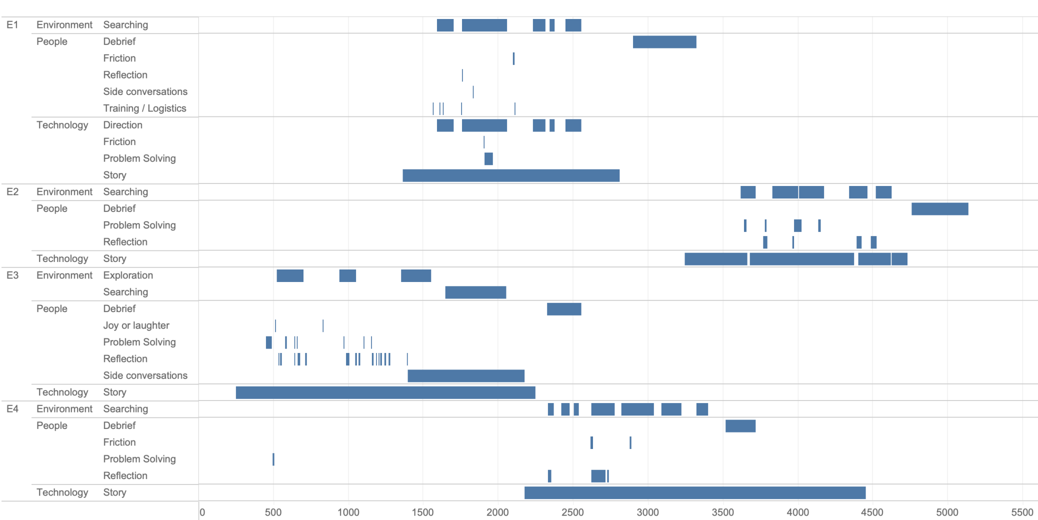


Fig. 19. Overview of Soul Seeker video coding

Soul Seeker. In Soul Seeker, the coding in Figure 19 indicates a device-centered, introspective pacing: prolonged *Interact with Technology* for reading and making choices, with comparatively sparse *Interact with Others*; *Interact with Environment* appears as wayfinding driven by the fiction rather than by inspecting physical affordances. Quantitatively, Soul Seeker time using technology strongly tracked session length, consistent with reading-heavy and choice-heavy play ($\rho = .839$, $p < .001$). As a complementary (cross-task) indicator of the solo tenor of Soul Seeker, players' social connection ratings aligned with how they felt in the also-solitary control, Speed Test ($\rho = .719$, $p = .019$). Participants' language emphasizes immersion in on-screen content and attention tethered to the device:

> "*The distraction part really comes down to like, you're reading a lot of text and you are not paying attention to your surroundings. Cause like the one thing I noticed was when we were over there, there was like two people waiting. I didn't even notice*" (E3A).

Noticing the solitary nature of Soul Seeker, one participant suggested ways to make the game more social:

> "*I think it would be cool if there was an option to play as a team. Where, like in D&D, your goal is to do something with your team. So we could all be finding our spearate soul fragements and at the end, depending on what everybody else chose....and because everyone has their own mini-tasks to do, we don't have to rely too much on others to figure out where we are going and it would be a little more interesting and then I'd be like 'okay this is my story and I can figure out where I'm going*" (E3B).

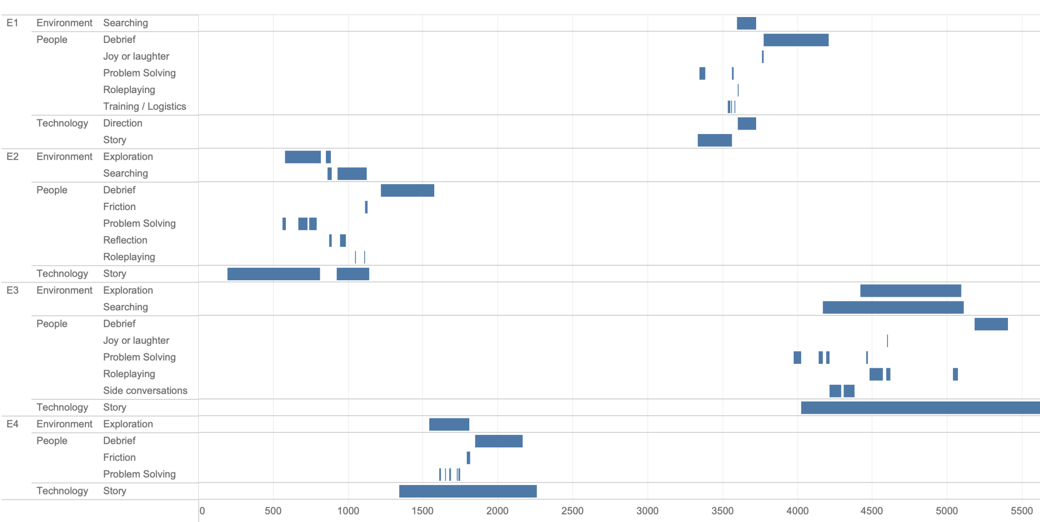


Fig. 20. Overview of Zombie Apocalypse video coding

Zombie Apocalypse. Zombie Apocalypse's video coding in Figure 20 reveals tight coupling between embodied chase (frequent *Interact with Environment*) and moment-to-moment social awareness (*Interact with Others*), with the device acting as a tactical aid (bursts of *Interact with Technology* for scans or status). Time in environment was positively associated with players' felt social connection ($\rho = .825$, $p = .006$), consistent with head-up searching, hiding, and coordinating in sight of others. Players described vigilant scanning for others and strategic movement:

> "*I avoided them like the plague. As the person who was the scientist, I tried to make sure that I was outside of a certain range, weaving between them to make sure. I did not know if they had to physically touch me or if they just had to try to attack me (E1A)*"

> "*As a zombie, my goal is to kind of like getting closer to everybody...closing the radius towards the scientist*" (E1B).

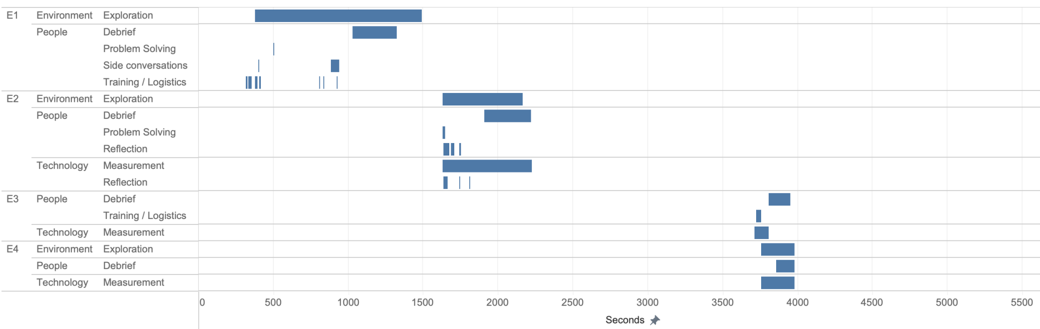


Fig. 21. Overview of Speed Test video coding

Speed Test. Speed Test's video coding in Figure 21 exhibits the most technology-forward profile: extended *Interact with Technology* while standing still or walking short segments, minimal *Interact with Others*, and functional *Interact with Environment* tied to finding test spots rather than discovering features. In line with that, time spent engaging the physical space was utilitarian and tightly coupled to the measurement task: Time in environment correlated with number of measurements ($\rho = .876, p > .001$). Participants framed the experience as a precise, tool-like routine with occasional informational payoff:

> "*I thought it was pretty informative to see the different download and upload speeds compared in different locations*" (E2A).
>
> "*The UI was kind of basic. I wouldn't say it really was a game. It didn't really feel like a game to me, so that's my only complaint*" (E3A).

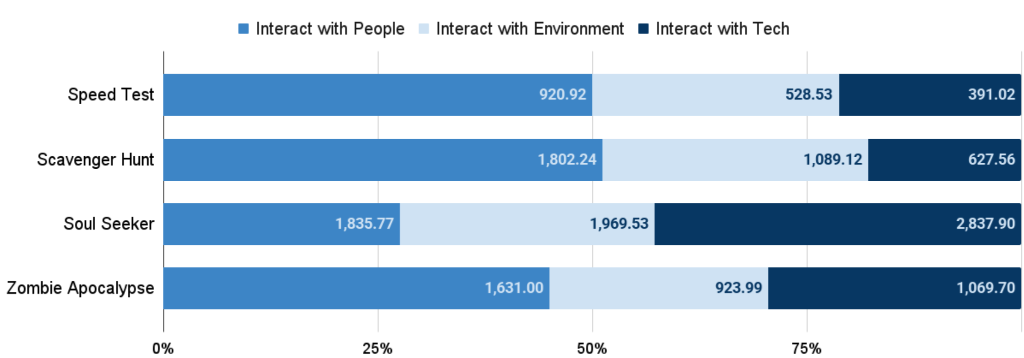


Fig. 22. Aggregate percentage of time spent interacting with people, environment, and technology across each game. We report aggregate duration in seconds within each bar.

Game Affordances and Aggregate Thematic Analysis. Across all games, aggregate durations further emphasize how each design distributed attention across people, environment, and technology (Figure 22). Scavenger Hunt and Soul Seeker yielded the longest overall sessions, but in different ways: Scavenger Hunt balanced substantial environment interaction (1,089s) with steady peer discussion (1,802s), while Soul Seeker shifted that time heavily into device-based story play (2,838s). Zombie Apocalypse showed the most social intensity, with over 1,600s coded as interaction with others, coupled with roughly equal distribution of environment (924s) and technology (1,070s). By contrast, Speed Test was shortest and most utilitarian, with limited people interaction (921s) and environment engagement (529s) compared to its sustained measurement use of technology (391s).

Key Takeaways. Video coding highlights how each game's mechanics distinctly shaped patterns of attention and interaction. Scavenger Hunt emphasized environmental exploration with light device use and limited peer exchange. Soul Seeker was device-centered, dominated by reading and narrative choices, with minimal social engagement. Zombie Apocalypse uniquely blended embodied searching with frequent social awareness and collaboration, balancing device cues with active pursuit. Speed Test was the most utilitarian, characterized by repetitive device-based measurement and little sustained connection to others or the environment. Aggregate durations confirm these contrasts: Soul Seeker maximized technology engagement, Zombie Apocalypse foregrounded social play, Scavenger Hunt emphasized environment, and Speed Test minimized all three.

6.2.2 Connection to Other Players.

Omnibus Analysis and Descriptive Statistics. Participants' self-reported feelings of social connection varied widely across the four games, as shown in Figure 23. Most responses to the Speed Test control clustered at the lowest categories ("not connected at all"), yielding an average score of 1.4. Soul Seeker also trended toward weak social connection, with an average of 1.8 and many participants selecting low values despite some neutral or connected ratings. Scavenger Hunt occupied a more intermediate position, with responses spread across the scale and an average of 2.7, reflecting occasional collaboration but mostly individual searching. Zombie Apocalypse stood apart: nearly all participants reported feeling connected or very connected to others, producing an average of 4.7. These contrasting distributions set the stage for formal ranking analysis and underscore how game mechanics directly shaped whether players experienced the session as a shared social activity or as a largely solitary task.

Because all four games' ranking data deviated from normality, as indicated by Shapiro–Wilk tests (Zombie Apocalypse $W = .327$, $p < .001$; Speed Test $W = .465$, $p < .001$; Scavenger Hunt $W = .754$, $p = .003$; Soul Seeker $W = .828$, $p = .020$), and the discrete, ordinal nature of preference rankings, the Plackett-Luce model was selected. The Plackett-Luce analysis revealed substantial differences in game preferences for social connection. Table 20 presents the overall worth estimates, which sum to 1.0 and represent relative preference strengths. Zombie Apocalypse demonstrated overwhelming preference, accounting for 89.5% of total preference weight. The remaining games received substantially lower preferences: Speed Test (7.8%), Scavenger Hunt (2.1%), and Soul Seeker (0.6%). Worth estimates by location are presented in Table 21, but location did not have a significant impact or interaction on ranking.

Correlational Analysis and Qualitative Reflections. Correlational analyses show that social connection was supported differently across the games. In Scavenger Hunt, connection to others increased with more time spent engaging the environment ($\rho = .940$, $p < .001$). In Speed Test, reported social connection was associated with total time ($\rho = .661$, $p = .027$). Zombie Apocalypse showed connection to others was positively linked to time coded as environmental searching ($\rho = .825$, $p = .006$). In Soul Seeker, connection to others tracked with uncorrelated entropy ($\rho = .651$, $p = .041$). Participant interviews revealed many insights related to the degree of connection to others afforded by each game. For Scavenger Hunt, there were mixed reactions based on whether

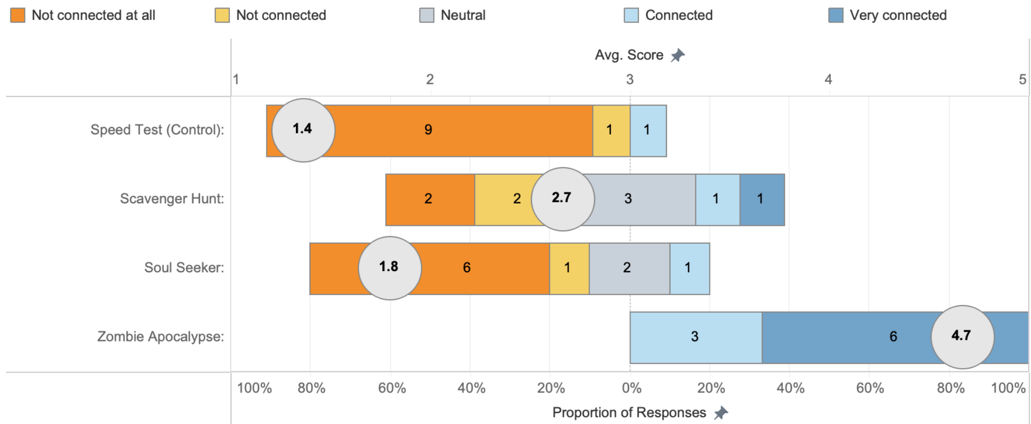


Fig. 23. Likert survey responses to *"How connected did this game make you feel to other players?"*. Graph uses dual axis where proportions of responses are shown centered at neutral with raw counts within each proportion and average score in gray circles. Not all participants answered all questions (e.g., Speed Test had 11/12 participants respond).

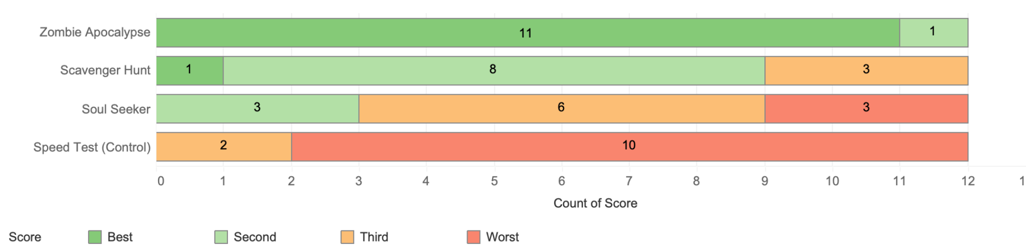


Fig. 24. Player rankings for each game across social dimension, sorted by overall ranking.

or not participants viewed the game as a solo challenge or a competition, which is left intentionally ambiguous:

> "*There was no connection at all...it was individual playing kind of thing. They were working on their [phone] and I was playing on my own [phone].*"(E4A).

> "*I was trying to be hyper alert where they were at all times. I wanted so bad to win.*"(E1A).

For Soul Seeker, there was unanimous agreement that it was not a very socially connected game beyond looking to where others are finding the soul fragments or comparing the endings everyone got:

> "*I didn't really see any interaction. I was more invested into my own story. The only interaction I would say we have is finding out each other's endings.*"(E3A).

> "*"The first two times, I was on my own, but by the third time I already had it in my mind that maybe I need to go in that particular direction because I could see the other players going in that direction, so I just felt like I know where I should go even before the played the game"*"(E3C).

Overall Game Worth Estimates				
Game	Worth			
ZA	0.895			
ST	0.078			
SH	0.021			
SS	0.006			
Plackett-Luce Model Comparisons				
Model Comparison	LR	df	df_{res}	p
Game × Location vs Overall	0.000	0	–	1.000
Stratified vs Pooled	7.918	9	–	0.542
Overall Pairwise Comparisons (Holm-adjusted)				
Contrast	Log-Worth Diff	SE	z	p_{adj}
ZA – SS	5.08	1.09	4.66	**<.001**
ZA – SH	3.74	0.995	3.76	**<.001**
SS – ST	-2.64	0.786	-3.37	**.003**
ZA – ST	2.44	1.04	2.34	.057
SS – SH	-1.34	0.726	-1.84	.086
ST – SH	1.31	0.646	2.02	.086

Table 20. Game worth estimates, model comparisons, and Holm-adjusted pairwise contrasts for social connection rankings.

Location 1						**Location 2**					
Contrast	Log-Worth Diff	SE	z	p	p_{adj}	Contrast	Log-Worth Diff	SE	z	p	p_{adj}
ZA – SS	3.75	1.84	2.04	0.041	0.247	ZA – SS	3.94	1.86	2.12	0.034	0.170
ZA – ST	0.99	1.23	0.80	0.424	0.849	ZA – ST	1.96	2.06	0.95	0.341	0.683
ZA – SH	1.58	1.23	1.29	0.198	0.790	ZA – SH	4.46	1.92	2.32	0.020	0.121
SS – ST	-2.76	1.76	-1.57	0.116	0.581	SS – ST	-1.98	1.39	-1.42	0.154	0.463
SS – SH	-2.17	1.76	-1.23	0.218	0.790	SS – SH	0.52	1.17	0.45	0.656	0.683
ST – SH	0.60	1.11	0.54	0.590	0.849	ST – SH	2.50	1.47	1.70	0.090	0.358
Location 3						**Location 4**					
Contrast	Log-Worth Diff	SE	z	p	p_{adj}	Contrast	Log-Worth Diff	SE	z	p	p_{adj}
ZA – SS	4.96	2.21	2.25	0.025	0.148	ZA – SS	4.81	2.14	2.24	0.025	0.149
ZA – ST	2.33	1.77	1.32	0.187	0.562	ZA – ST	2.62	2.00	1.31	0.190	0.569
ZA – SH	2.79	1.76	1.59	0.112	0.561	ZA – SH	3.38	1.96	1.73	0.084	0.421
SS – ST	-2.63	1.72	-1.53	0.127	0.561	SS – ST	-2.18	1.45	-1.50	0.133	0.531
SS – SH	-2.17	1.71	-1.26	0.206	0.562	SS – SH	-1.43	1.40	-1.02	0.307	0.614
ST – SH	0.46	1.09	0.43	0.670	0.670	ST – SH	0.76	1.17	0.65	0.515	0.614

Table 21. Game preference contrasts within each location (Holm-adjusted pairwise comparisons on log-worth scale).

For Zombie Apocalypse, multiplayer affordances were explicitly included in the instructions, and the mechanics forced players to be aware of each other and interact as well as make strategic choices about when to use different power-ups:

> "*It was very connecting because I have to see whoever is near or the way the scientist is going and all. And that was actually a very good connection with them.*"(E4A).

> "*I also like looked for the other people...Like if he's there and I'm there I can probably like hit the scientist, so I make sure I have the strength power instead of like the speed.*"(E4B).

For Speed Test, participants did not interact with each other:

> "*I feel like most of the time you're spent looking at your phone trying to press the button to collect the data. If we could have a clicker, that way it'll work wirelessly for the phone. So we don't have to continue to look at our phone. We can just click a button.*"(E3A).

Key Takaways. Zombie Apocalypse was the only game that consistently fostered strong feelings of social connection, with players reporting high awareness, collaboration, and strategic interaction. Scavenger Hunt produced mixed experiences—some players treated it as an individual challenge while others engaged competitively. Soul Seeker emphasized personal storytelling and thus offered minimal social connection beyond comparing outcomes. Speed Test elicited the lowest connection, as participants were largely absorbed in individual data collection. Overall, social connection outcomes aligned directly with whether multiplayer affordances were intentionally built into the game design. In contrast with utilitarian outcomes related to *RQ1* (Section 6.1), there was not a significant interaction between *Session* and *Game*, which is interesting because it suggests that both context and game affordances matter for utilitarian outcomes, but only game affordances matter for fostering connection to others.

6.2.3 Connection to Space.

Omnibus Analysis and Descriptive Statistics. Participants' reported sense of connection to the physical environment also varied across games, as shown in Figure 25. Speed Test again scored lowest, with most responses falling in the "not connected at all" or "not connected" categories and an average of 2.5. Soul Seeker produced slightly higher ratings (average = 3.0), with many participants neutral and a few noting moments of attentiveness despite the story-heavy design. In contrast, both Scavenger Hunt and Zombie Apocalypse were rated highly for environmental connection, each averaging 4.3.

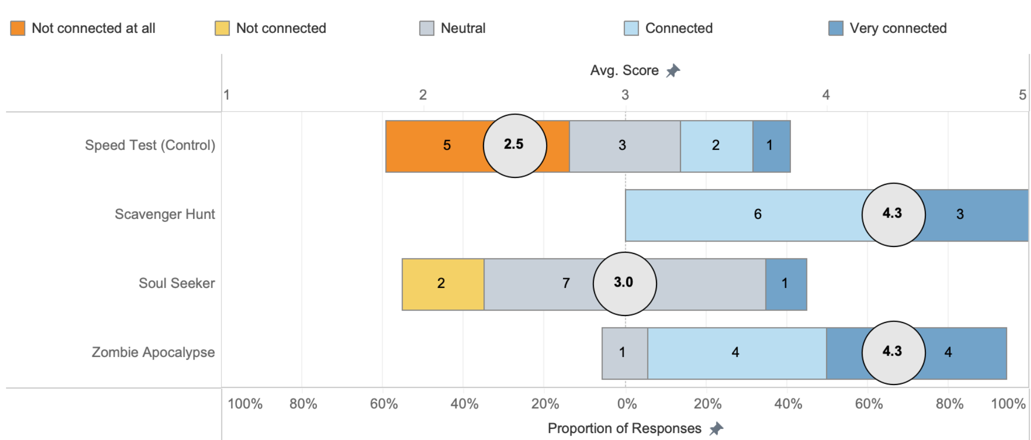


Fig. 25. Likert survey responses to *"How connected did this game make you feel to the spatial environment?"*. Graph uses dual axis where proportions of responses are shown centered at neutral with raw counts within each proportion and average score in gray circles. Not all participants answered all questions (e.g., Speed Test had 11/12 participants respond).

Because all four games' ranking data deviated from normality, as indicated by Shapiro–Wilk tests (Zombie Apocalypse $W = .809$, $p < .012$; Speed Test $W = .731$, $p < .002$; Scavenger Hunt $W = .552$,

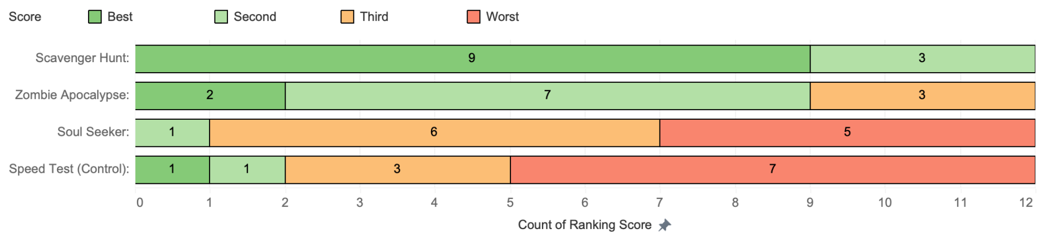


Fig. 26. Player rankings for each game across environmental dimension, sorted by overall ranking.

$p < .001$; Soul Seeker $W = .784$, $p = .006$), and the discrete, ordinal nature of preference rankings, the Plackett-Luce model was selected. The Plackett-Luce analysis revealed substantial differences in game preferences for environmental connection. Table 22 presents the overall worth estimates, which sum to 1.0 and represent relative preference strengths. Scavenger Hunt demonstrated overwhelming preference, accounting for 75.0% of total preference weight. The remaining games received substantially lower preferences: Zombie Apocalypse (18.9%), Speed Test (3.5%), and Soul Seeker (2.6%). Worth estimates by location are presented in Table 23, but location did not have a significant impact or interaction on ranking.

Correlational Analysis and Qualitative Reflections. Correlational analyses linking reported connection to the environment with behavioral measures showed meaningful differences across games. In Scavenger Hunt, players who spent longer in the session also reported stronger environmental connection ($\rho = .730$, $p = .025$), suggesting that extended clue-following translated into greater place-based attentiveness. In Zombie Apocalypse, higher entropy in players' spatial paths was positively associated with environmental connection ($\rho = .721$, $p = .028$), and self-reported social connection also tracked with environmental connection ($\rho = .700$, $p = .036$), highlighting how scanning and movement were intertwined with both spatial and social awareness. In Speed Test, environmental connection was driven by number of broadband measures ($\rho = .818$, $p = .002$) and real entropy ($\rho = .830$, $p = 002$). Soul Seeker did not have any statistically significant correlations. Participant interviews revealed many insights related to the degree of connection to environment afforded by each game. Scavenger Hunt's clue finding mechanic forced players to look closely at the environment and learn about their surroundings:

> "*This one actually kind of a significant amount because you're actually looking in the environment so you're actually observing it you have to actually observe it and actually learn about other facts i didn't know about so i would say a little bit more.*"(E3A).

> "*I noticed things I wouldn't, like, have cared about the area. Like, over there, it talks about how, like, Flagstaff was trying to, like, help the butterflies. I wouldn't know that if I didn't play the game.*"(E4B).

For Soul Seeker, the mechanics largely mediated attention through the device rather than the setting, which limited how strongly players felt tied to their surroundings. Participants noted both the constrained physical footprint of the game and how sustained text reading kept their focus on the phone rather than on the park. As one explained:

> "*I just wish the space was a little bigger. Just because it would have been nice to be walking around through the park area and feeling like you are going on your own adventure. And it would be nice to also find some spots where there is seating, because I know you have to read a lot of text while you're standing up and there's sunlight so*

Overall Game Worth Estimates				
Game	Worth			
SH	0.750			
ZA	0.189			
ST	0.035			
SS	0.026			
Plackett-Luce Model Comparisons				
Model Comparison	LR	df	df_{res}	p
Game × Location vs Overall	0.000	0	–	1.000
Stratified vs Pooled	7.911	9	–	0.543
Overall Pairwise Comparisons (Holm-adjusted)				
Contrast	Log-Worth Diff	SE	z	p_{adj}
SS – SH	-3.36	0.772	-4.35	**<.001**
SH – ST	3.07	0.740	4.15	**<.001**
ZA – SS	1.98	0.658	3.01	**.010**
ZA – ST	1.69	0.620	2.73	**.019**
ZA – SH	-1.38	0.755	-1.82	.137
SS – ST	-0.29	0.641	-0.45	.654

Table 22. Game worth estimates, model comparisons, and Holm-adjusted pairwise contrasts for environmental connection rankings.

Location 1						**Location 2**					
Contrast	Log-Worth Diff	SE	z	p	p_{adj}	Contrast	Log-Worth Diff	SE	z	p	p_{adj}
ST – SH	2.54	1.32	1.93	0.054	0.325	ST – SH	3.23	1.50	2.15	0.031	0.188
ZA – SS	0.04	1.09	0.04	0.971	1.000	ZA – SH	2.72	1.43	1.90	0.058	0.288
ZA – ST	-1.44	1.19	-1.21	0.225	1.000	SS – ST	-2.71	1.44	-1.89	0.059	0.288
ZA – SH	1.10	1.11	1.00	0.320	1.000	ZA – SS	2.19	1.36	1.61	0.107	0.321
SS – ST	-1.48	1.30	-1.14	0.256	1.000	ZA – ST	-0.52	1.52	-0.34	0.733	1.000
SS – SH	1.06	1.23	0.86	0.388	1.000	SS – SH	0.52	1.35	0.39	0.698	1.000
Location 3						**Location 4**					
Contrast	Log-Worth Diff	SE	z	p	p_{adj}	Contrast	Log-Worth Diff	SE	z	p	p_{adj}
SS – ST	-3.75	1.84	-2.04	0.041	0.247	SS – ST	-4.46	1.92	-2.32	0.020	0.121
ZA – SS	2.76	1.76	1.57	0.116	0.581	ST – SH	3.94	1.86	2.12	0.034	0.170
SS – SH	-2.17	1.76	-1.23	0.218	0.790	ZA – SS	2.50	1.47	1.70	0.090	0.358
ST – SH	1.58	1.23	1.29	0.198	0.790	ZA – SH	1.98	1.39	1.42	0.154	0.463
ZA – ST	-0.99	1.23	-0.80	0.424	0.849	ZA – ST	-1.96	2.06	-0.95	0.341	0.683
ZA – SH	0.60	1.11	0.54	0.590	0.849	SS – SH	-0.52	1.17	-0.45	0.656	0.683

Table 23. Game preference contrasts within each location (Holm-adjusted pairwise comparisons on log-worth scale).

> *you have to kind of like move your body in the way of the phone just so you can view the text. Overall I just wish it was a much bigger playspace, I wish it was optimized so we could have seating arrangements.*" (E3A).

Others were even more direct in emphasizing how vibration cues, rather than place features, oriented their attention:"*I wasn't really connected to the environment. I was more looking for the vibration of a soul fragment.*" (E4B).

For Zombie Apocalypse, the mechanics of chasing and hiding fostered connection to the environment via heightened spatial awareness as they navigated paths, obstacles, and designated play zones. For example, one player noted, "*I feel like it was still interactive, where I felt like every place I went, I'm almost like visualizing in my head that I am a scientist like maybe digging up samples or something wherever I am.*" (E1A). Others emphasized how the physical constraints of the space shaped their awareness: '*I would say it made me aware of my surroundings. I didn't like wanna go into the rock falls … could only go in one direction. So I would just wanna try to stay on the path.*" (E1C).

For Speed Test, the mechanics encouraged device-centered measurement rather than environmental engagement, though participants varied in how much this constrained their awareness of place. Some described the repetitive task of tapping the phone as distracting, reducing opportunities to notice surroundings:

> "*I feel like most of the time you're spent looking at your phone trying to press the button to collect the data. If we could have a clicker, that way it'll work wirelessly for the phone. So we don't have to continue to look at our phone. We can just click a button.*" (E3A).

Others, however, experienced the simplicity of the task as freeing, allowing them to look up and scan their environment:

> "*I kind of feel the opposite because it was more of a cookie clicker. I wasn't really paying attention to much on my phone. I'm not sure if I was supposed to. But because it was so simple to do and there wasn't really much instruction and I didn't feel like I had to pay attention to the phone as much, I really got to look around and actually understand where I'm going and what's going on.*" (E3C).

Yet others emphasized that the minimal feedback kept them disengaged from the setting, with one player concluding, "*If I knew how to read like [what in there], I would probably be like 'Oh, the WiFi is good or bad'. But I was just clicking. Yeah, that's all!*" (E4B).

Key Takeaways. Scavenger Hunt and Zombie Apocalypse most strongly fostered connection to place, with Scavenger Hunt overwhelmingly preferred in rankings. These preferences were stable across locations (no *Game* × *Location* effect), indicating that environmental connection is driven by game design rather than context. Correlational patterns align with this: in Scavenger Hunt, longer sessions predicted stronger place connection; in Zombie Apocalypse, higher path entropy and social connection tracked with environmental connection; Speed Test's place connection rose mainly with more measurements and movement entropy, and Soul Seeker showed no significant behavioral ties. Overall, mechanics that require clue-following, scanning, and spatial strategy amplify place-based attentiveness, whereas device-centric play dampens it.

6.3 Players' Perspectives on Game Experience

In this section, we foreground players' own accounts of the games. First, we summarize the three adjectives participants used to describe each game and contextualize these with representative quotes that highlight how mechanics shaped emotion, attention, and sociality (Section 6.3.1). Then, we assess whether participating in the session shifted interest in broadband measurement using a pre–post design (Section 6.3.2), providing a quantitative complement to participants' game reflections. We discuss implications in Section 7.

6.3.1 Participant Game Reflections and Descriptions. At the end of each game in each session, we asked each player to provide three adjectives to describe their experience playing a particular game. We created a treemap of the adjectives most frequently used by participants in Figure 27.

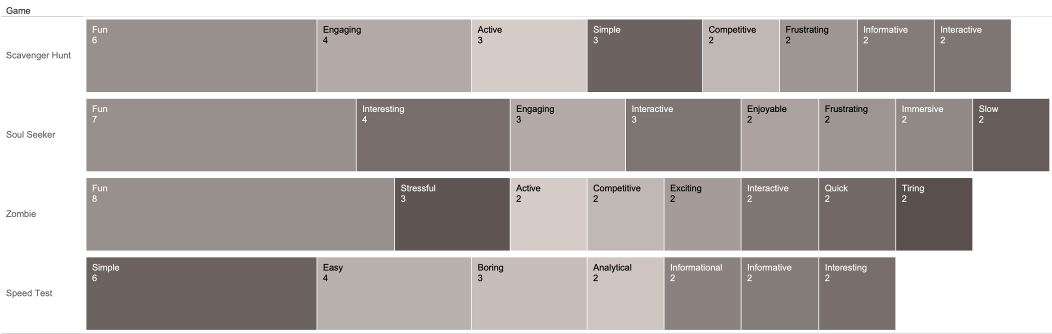


Fig. 27. Participant-Adjective Frequency by Game. This treemap visualization displays the most common adjectives participants used to describe each game. The size of each block represents the number of participants who used that adjective.

For Scavenger Hunt, Soul Seeker, and Zombie Apocalypse, the most frequently used adjective was "fun"—mentioned by 6, 7, and 8 participants respectively. In contrast, the most frequently used adjective to describe Speed Test was "simple," cited by 6 participants. Other commonly used adjectives included "engaging" (4) for Scavenger Hunt, "interesting" (4) for Soul Seeker, and "stressful" (3) for Zombie Apocalypse, reflecting both the enjoyment and cognitive demands of each game. Speed Test was also described as "easy" (4) and "boring" (3), supporting the benefits of using games for broadband measurement. Participant interviews revealed a range of experiences across the four games, with players highlighting different emotional, spatial, and technological aspects that shaped their engagement. Below, we present quotes grouped by game to illustrate the diversity of player responses.

Scavenger Hunt. Scavenger Hunt's clue-finding mechanics encouraged players to look closely at their surroundings, uncovering details they might have otherwise overlooked. Many participants emphasized how the game helped them see the area with fresh eyes:

> "*I was learning the area a little bit better*" (E1A).

> "*I also have been to this area several times, so I knew some of them, but a lot of them I didn't. And it helped me feel more connected.*" (E4C).

Soul Seeker. Soul Seeker evoked comparisons to fantasy games and popular location-based experiences:

> "*it was interesting and the story at the start was a lot of information I was bombarded with that could be figured out along the way and the game was simple enough to understand, the story at the start was a little bit boring but the ending was interesting and I felt it was the best part*" (E3A).

> "*It reminds me of Pokemon Go. And that's what I like about it. That's why I was like, I want to, I want to skip through it so I can like go find all my Pokemon.*" (E3C).

Zombie Apocalypse. Zombie Apocalypse was experienced as energetic and immersive, blending physical activity with moments of tension and competition. Several players noted that the game could support stronger social connection, especially with familiar teammates:

> "*I feel like in the future we could communicate better but ultimately I love the attention to really like connect you with the people you are playing with in general.*" (E2A).

> "*Competitive - that's probably just a me thing. I am already a pretty competitive person. I like the competitive aspect of it. Almost like nerve wrecking....because don't wanna be dead*" (E1A).
>
> "*One thing, if I knew the players in the zombie game, it probably would have been fun, because you're probably interacting, yelling at your friend, being like, don't get me, don't get me. Like, since I don't know them, I'm just like, oh, I've got to win.*" (E4B).

Players also commented on specific mechanics they enjoyed or wished to improve:

> "*I liked the competitive aspect of it*" (E2C).
>
> "*I do like the buzz mechanic*" (E1A).
>
> "*I also kind of wish that instead of like telling me how many steps I need to take it was like a timer like I wouldn't have to like keep looking down instead like the timer would go off and it would like vibrate and tell me like you can't walk anymore.*" (E3C).

Speed Test. Speed Test elicited notably different reactions, with many describing it as task-focused, solitary, and less engaging than the other games. While some appreciated its clarity and simplicity, others commented on the lack of meaningful social interaction:

> "*We just walked by each other. We didn't really acknowledge each other.*" (E3A).
>
> "*Oh boy!!!...I would say interesting. I don't know if it's just because I'm interested in network stuff but definitely interesting...informative.*" (E1B).
>
> "*So the complexity has to do with a lot of me not knowing the terms of this. Like, what is latency? What does upload mean in this situation like around here?*" (E3A).

6.3.2 Interest in Broadband Measurement (Pre vs. Post). To evaluate whether participation in the broadband measurement games affected participants' interest in the topic of internet measurement, we collected interest ratings both before and after the game session. Participants were asked to rate their interest in broadband measurement on a 5-point Likert scale (1 = Not at all interested, 5 = Extremely interested) during the pre-test and again at the end of the session.

Tests of normality indicated that pre- and post-interest scores were not normally distributed (Shapiro–Wilk: Pre $W = 0.753$, $p = .003$; Post $W = 0.809$, $p = .012$), with both distributions showing deviations from normality despite relatively symmetric skewness values. Because the data were paired and ordinal-level with violated assumptions of normality, a nonparametric Wilcoxon Signed-Rank test was applied to assess changes in interest from pre- to post-intervention. The Wilcoxon test showed that post-interest scores (Median = 4.0, IQR = 1) were not significantly higher than pre-interest scores (Median = 4.0, IQR = 1), $Z = 1.73$, $p = .083$. The associated effect size was medium-to-large ($r = .50$), suggesting a trend toward increased interest after the intervention, though not statistically significant at the $\alpha = .05$ level. All SPSS computations, including normality tests, variance checks, and stepwise outputs, are included in the supplementary files.

Key Takeaways. Players consistently described Scavenger Hunt, Soul Seeker, and Zombie Apocalypse as "fun," highlighting how narrative, exploration, and competition sustained engagement, while Speed Test stood out as "simple" and often "boring." Reflections showed that game mechanics directly shaped emotional tone—Scavenger Hunt encouraged discovery, Soul Seeker emphasized story, Zombie Apocalypse blended excitement with stress, and Speed Test reduced play to task repetition. Despite these differences, overall interest in broadband measurement showed only a non-significant upward trend from pre- to post-session, suggesting that while games shaped players' immediate experiences and perceptions, short-term participation did not reliably translate into broader shifts in interest.

7 DISCUSSION

This section examines the implications of our findings in relation to the two research questions, each addressed through a consistent structure: broader field implications, design research implications, methodological insights, future work, and limitations. We begin with *RQ1*, on crowdsourcing broadband measurements across space (7.1). We then turn to *RQ2*, which explores how the platform fosters community among players and with place (7.2).

7.1 *RQ1*: Crowdsourcing Internet Measurements over Space

7.1.1 Broadband Measurement (Broader Field Implications). NetGauge offers a novel contribution to the networking research community by reimagining broadband measurement as a participatory, gameplay-driven process. Traditional approaches—such as drive tests or passive background logging—often result in spatially biased datasets, overrepresenting major roadways and affluent urban centers while neglecting rural or underserved areas [67]. As scholars have argued, network measurement must adapt to real-world, user-centered conditions to remain representative and actionable [71]. NetGauge supports this shift by introducing a transparent and consent-driven sensing model: rather than silently logging performance, it visibly ties measurements to meaningful player actions—advancing ethical data practices [30]. Its modular design also allows for integration with testbeds like M-Lab [30], or for deployment in underserved regions via community-driven initiatives, aligning with broader goals of digital equity [67]. As networking researchers seek scalable, ethical, and community-aligned sensing strategies, NetGauge demonstrates how play can function as a vehicle for engagement and infrastructure—one that reshapes how, where, and by whom broadband access is measured.

Measurement as Mechanic. While passive internet measurement via apps like navigation tools can generate large-scale data, we argue that embedding measurement within game mechanics offers distinct utilitarian and experiential advantages. In NetGauge, tests were triggered by intentional in-game actions—such as narrative decisions in Soul Seeker, combat events in Zombie Apocalypse, and spatial reach goals in Scavenger Hunt. These mechanics foregrounded player agency, framing measurement as a meaningful, embodied contribution rather than as invisible background logging. NetGauge responds to calls to examine whether games and play can deepen civic engagement [74], and our results provide evidence that they can: all three games outperformed the Speed Test control on key utilitarian outcomes, including significantly greater distance traveled and more internet measurements collected. By aligning sensing with action, NetGauge sustained motivation and can direct movement toward infrastructural blind spots. While passive approaches may yield higher measurement volume, game-based systems can dynamically shape *where* and *how* that data is collected—an important distinction in contexts where spatial equity and data granularity matter. As broadband access remains a Wicked Problem [18], our findings contribute to a growing body of literature that sees play not as a distraction from civic work, but as a design strategy for doing it better [36, 38].

7.1.2 Ultimate Design Particulars and Platform Utility (Design Research Implications). Each NetGauge game enacted a different rhythm of movement through space, revealing how subtle shifts in play mechanics produce patterned differences in how utilitarian tasks (like broadband measurement) are performed. Scavenger Hunt encouraged fast, goal-oriented traversal between known points, maximizing spread and surface coverage. Soul Seeker fostered slower, meandering paths shaped by introspective narrative cues, often leading players to peripheral or quiet zones. Zombie Apocalypse generated urgency and looping through its predator-prey logic, creating spatial redundancy and tight clustering. These differences were not incidental but designed: specific affordances—time pressure, ambiguity, spatial density—shaped movement tempos and coverage patterns. Framed through the lens of ultimate design particulars [86], these cases show how even small changes in pacing, goal structure, or narrative framing materialize as infrastructural labor in the world. For designers working at the intersection of play and serious purposes, the implication is clear: spatial outcomes must be designed as intentionally as mechanics or narrative arcs. Effective broadband measurement is not just about collecting data, but about choreographing movement to align with infrastructural blind spots, signal edge conditions, or social variability in connectivity.

Situated Play Design and Environmental Context. NetGauge made clear that game mechanics do not operate independently of place; rather, they are modulated—and sometimes reinterpreted—by the physical, infrastructural, and social textures of the environment. Across our deployments, the same mechanics produced different spatial behaviors depending on context: in semi-urban campuses, players in Soul Seeker wandered freely through open greens and quiet paths, supporting wide dispersal and high entropy. In dense urban areas, Zombie Apocalypse produced tighter movement loops, as players navigated pedestrian traffic and constrained pathways. In rural settings, Scavenger Hunt revealed drop-offs in pace and engagement due to long gaps between landmarks and lack of visual cues. These frictions and affordances were not failures of design, but expressions of what Situated Play Design predicts: that play is always co-constructed with place and the magic circle [2, 46]. For broadband sensing, this demands more than spatially neutral game logic. Designers must tune interaction density, narrative pacing, and spatial layout to the specific conditions of deployment—whether to motivate coverage across long rural stretches, avoid redundancy in urban grids, or scaffold movement through ambiguous terrain. Effective measurement depends on what mechanics are used, where they are situated, and what kinds of movement and meaning they afford in context.

Design Ambiguity for Space Exploration. Ambiguity in NetGauge was intentionally designed for as a spatial design instrument—deployed to resist overdetermined movement and encourage emergent, player-driven navigation. Zombie Apocalypse introduced tactical ambiguity: zombie start points and cure locations were unknown, forcing improvisation, risk assessment, and looped traversal. Scavenger Hunt presented a fixed objective ("find four hidden locations") but withheld maps, requiring players to infer structure through exploration. Across these designs, ambiguity disrupted optimization and seeded divergent paths—producing broader spatial coverage without explicit instruction. In HCI, ambiguity is often framed as a resource for interpretation and appropriation [81]; here, it functioned as a mechanism for spatial decentralization, countering redundancy and player convergence. For civic sensing applications, this suggests that ambiguity can be strategically layered—not as confusion, but as a means of diffusing player trajectories across unmeasured or under-engaged zones. Future designs might experiment with probabilistic rewards, ambiguous affordances, or decoy signals to calibrate exploration while retaining interpretive flexibility. In this view, ambiguity is play aesthetic and a tool for infrastructural tuning.

Dimensions of Play. Categorizing NetGauge Games across core play dimensions—such as synchronous vs. asynchronous, competitive vs.collaborative, solo vs. social—revealed how structural affordances shaped broadband measurement outcomes [24]. These dimensions encoded how movement unfolded and where sensing occurred. Viewing games through this lens enables designers to more precisely align play structures with sensing goals, or intentionally vary them to balance coverage. Just as importantly, mapping across these dimensions reveals gaps—underutilized combinations like cooperative, asynchronous, or high-chance play—that may offer new leverage points for exploration. In this way, play dimensions function as both design parameters and analytic tools for engineering more adaptive, context-sensitive civic sensing systems.

7.1.3 Methodological Insights.

Mobility Analysis For Broadband Measurement Collection. To assess how well NetGauge Games facilitated spatial broadband measurement, we analyzed mobility patterns using five core metrics: straight-line distance, radius of gyration, random entropy, real entropy, and area coverage. Each metric captured a distinct aspect of spatial behavior relevant to sensing goals. Straight-line distance measured overall reach from the starting location; radius of gyration captured how concentrated or dispersed movement was around a central point. Random entropy quantified the number of unique locations visited, while real entropy accounted for both uniqueness and sequence, reflecting the unpredictability of player paths. Area coverage estimated the total geographic surface traversed. These metrics allowed us to compare spatial behaviors across game conditions. For instance, Soul Seeker yielded significantly greater straight-line distances and higher entropy than the Speed Test control, showing that open-ended, narrative-driven play led to more exploratory and spatially varied measurement. Mobility metrics are both evaluative and generative: they provide actionable insights into how different play structures produce different sensing outcomes. For broadband measurement, this suggests that designers should align game mechanics with desired spatial patterns—maximizing entropy to reach underrepresented zones or adjusting gyration to avoid oversampling specific areas.

Mobility Analysis Usefulness in Play Design. While developed to evaluate broadband measurement coverage, our mobility metrics also surfaced latent dynamics of spatial interaction across game designs. Patterns of clustering, dispersion, and entropy revealed how mechanics scaffold or constrain movement, often in ways invisible to direct observation. Zombie Apocalypse's tight gyration and low entropy highlighted how reactive mechanics reinforced spatial redundancy—effective for repeated sampling, but limiting reach. Soul Seeker's higher entropy and distance reflected how narrative ambiguity encouraged emergent, self-directed exploration. These patterns were not simply reflections of game goals, but diagnostics of how players made sense of space under design constraints. Even when movement isn't the core mechanic, mobility metrics can uncover friction points, reveal affordance mismatches, or flag emergent inequities—such as clustering near start points due to unclear cues or environmental discomfort. For situated play design, this makes mobility analysis both evaluative and speculative: a tool for interrogating how spatial experience unfolds across bodies, environments, and rule systems, and for anticipating how design decisions might resonate differently across contexts.

RITE Piloting. RITE served both as a usability tool and as a systems-level tuning mechanism for the entire research pipeline. Iterative testing with real players in diverse environments surfaced failures we could not have anticipated in lab contexts—missing GPS logs in Zombie Apocalypse, dead camera batteries during outdoor sessions, and low initial inter-rater reliability in video coding. These breakdowns prompted immediate changes to both platform and protocol: refining how data were buffered, ensuring equipment readiness, and restructuring coder training. RITE also exposed analytic blind spots, leading us to revise coding schemas and add post-game surveys. This recursive refinement loop where gameplay, sensing, and analysis co-evolved positions RITE as more than a usability method. In situated civic technologies like NetGauge, it functions as a methodological backbone: surfacing misalignments between design, infrastructure, and evaluation before they scale, and enabling more robust deployment in the wild.

Mixed Methods. Beyond triangulation, our mixed-methods approach exposed tensions between what players did and what designers intended. Telemetry and mobility metrics showed that Soul Seeker achieved the broadest and most unpredictable movement, yet surveys revealed that players did not always associate this with "doing well" in the game. Conversely, Zombie Apocalypse produced tightly clustered paths that looked inefficient from a measurement standpoint but felt purposeful to players chasing cures or evading threats. These divergences surfaced design levers invisible to any single method—e.g., how ambiguity sustains exploration without clear extrinsic reward. For experts designing situated systems, this highlights the value of mixed methods not for revealing misalignments between system behavior, user perception, and utilitarian outcomes.

7.1.4 Future Work. Future work will expand NetGauge in three directions: measurement integration, platform scalability, and civic impact. First, we plan to incorporate actual network performance metrics—such as download speed, upload speed, jitter, and packet loss—into gameplay outcomes. This will allow game mechanics to respond dynamically to local network conditions, transforming infrastructure quality into a core variable of player experience (e.g., slower speeds could affect character mobility, challenge intensity, or scoring). Second, we aim to support dynamic point generation, enabling the platform to populate game objectives in real time based on network gaps—scaling NetGauge across geographies and contexts. Extending authoring capabilities to non-experts, particularly through platforms like Twine, will facilitate crowdsourcing of both measurements and the games themselves, fostering participatory design and local ownership. Lastly, by using telemetry to generate high-resolution coverage maps and training undergraduates through the RITE process, we aim to support both policy-aligned infrastructure planning and broader inclusion in civic tech research.

7.1.5 Limitations. While our findings demonstrate the potential of game-based approaches to broadband measurement, several limitations constrain generalizability and operational scalability. First, the platform currently simulates measurement triggers without integrating real network performance data (e.g., speed, latency), limiting our ability to evaluate sensing accuracy or variation across connectivity conditions. Second, the spatial behavior observed reflects short-term play sessions in specific geographic and environmental contexts; longer or repeated engagements may yield different movement patterns. Third, our study did not assess battery drain, data usage, or device variability—all of which could affect deployment feasibility at scale. Finally, our analysis emphasized spatial coverage as a proxy for sensing utility, but did not yet evaluate data quality or redundancy across repeated measurements. These constraints suggest opportunities for future iterations to integrate real performance data, test across broader environmental conditions, and expand evaluation frameworks to include sensing efficacy, user retention, and equity of representation.

7.2 *RQ2*: Community among Players and Environment

7.2.1 Community-Engaged Research Sustainability (Broader Field Implications). Sustained civic sensing requires more than one-off participation—it requires systems that foster a sense of ownership, continuity, and local relevance. NetGauge offers a model for sustainable engagement by making participants more than data sources—as co-performers in infrastructural inquiry. Unlike traditional crowdsourcing models that extract data through passive participation, NetGauge embeds sensing in playful, repeatable experiences that players can return to, share with their friends, adapt, or create themselves. By supporting modifiable games through platforms like Twine, the system invites communities to shape the narratives, mechanics, and spatial logics that structure participation. This opens a pathway from user to author, from player to steward—anchoring participation in place, identity, and iterative meaning-making. Sustainability then surpasses retention or reusability to be reconfigured and reinterpreted by the communities they serve. For civic tech and community-based research, this positions play not as an ephemeral hook, but as a scaffold for long-term engagement, local accountability, and infrastructural literacy.

Modular Platform. NetGauge's modular architecture supports a wide range of game genres, making it adaptable to the social dynamics, cultural preferences, and spatial constraints of different communities. By decoupling sensing logic from narrative and mechanic layers, the platform enables designers—or players themselves—to build experiences that resonate with local values and modes of interaction. Some games emphasize competition (Zombie Apocalypse), others reflection (Soul Seeker), or discovery (Scavenger Hunt), each fostering different forms of social connection—from playful rivalry to quiet co-presence. This flexibility allows NetGauge to scaffold community through a pluralistic ecosystem of play styles. For communities historically excluded from civic technology initiatives, this opens new entry points: groups can select or co-design genres that reflect their own aesthetics, rhythms, and relationships to infrastructure. In this way, modularity is not just a technical feature but a social affordance—enabling civic platforms to be tailored, localized, and made meaningful across diverse contexts.

Community Benefit. NetGauge demonstrates how "play for good" can extend societal benefit beyond data collection, fostering connection among people and between people and place. In contrast to extractive sensing models that instrumentalize participation, NetGauge embeds measurement within shared narratives, physical co-presence, and localized exploration. Players reported moments of recognition—passing familiar landmarks in unfamiliar ways, encountering others mid-game, or reflecting on the quality of their digital infrastructure in relation to their lived environment. These experiences reframe broadband sensing as a relational and situated act. This aligns with broader HCI work on playful civic engagement, which shows that games can make invisible systems legible and meaningful through embodied interaction [57]. NetGauge has the potential to foster awareness, conversation, and shared stakes in infrastructure beyond crowdsourcing measurements. The platform offers community benefit by supporting better measurement *and* by creating moments of civic intimacy, infrastructural literacy, and place-based solidarity.

7.2.2 Ultimate Design Particulars Fostering Community (Design Research Implications). Outdoor play facilitates community to environments—cultivating both spatial awareness and a sense of belonging. In NetGauge, players' interactions with physical landscapes were shaped by local knowledge, environmental affordances, and personal interpretations of place. As Situated Play Design argues, physical and cultural contexts are not passive backdrops but active co-constructors of experience [2]. Open spaces encouraged relaxed, exploratory movement; denser areas imposed constraints that invited strategic adaptation. These interactions made infrastructure legible while embedding play in everyday geographies. Critically, when players move with intent through familiar environments, they reaffirm presence and stake in that place. Outdoor civic play thus becomes a scaffold for local ownership: a means of turning infrastructural participation into a felt connection to the spaces players inhabit and the systems that serve them.

Intentional Design Ambiguity. Intentional ambiguity in NetGauge functioned as a spatial, social, and relational design tool. By leaving goals open to interpretation—such as emotional prompts in Soul Seeker or the unstated logic of zombie behavior in Zombie Apocalypse—the games invited players to project meaning, negotiate strategy, and reflect individually within a shared structure. This aligns with HCI's framing of ambiguity as a resource for dialogue, reflection, and situated appropriation [81]. In NetGauge, ambiguity supported "design through use" as much as "design through design"—creating space for players to make the game their own through how they moved, reacted, or narrated their experience. It also enabled multiplicity: players could engage competitively or contemplatively, treat objectives as puzzles or metaphors, and still contribute to shared sensing goals. This pluralism is critical for fostering community, as it supports diverse forms of participation without enforcing uniformity. In civic contexts, ambiguity offers creative play and it becomes a mechanism for inclusivity, allowing people with different motivations, backgrounds, and relationships to place to engage meaningfully on their own terms.

Dimensions of Play. Play structures scaffold interaction—they shape who feels invited, how relationships form, and what kinds of community take hold. Across NetGauge Games, core dimensions such as synchrony, cooperation, and structural framing acted as subtle social architectures. In Zombie Apocalypse, real-time threats generated a shared sense of urgency—even among strangers—through indirect coordination and spatial awareness of others. Soul Seeker and Scavenger Hunt emphasized asynchronous and solitary play, yet players described a quiet sense of togetherness through parallel movement, shared goals, or awareness of being part of a larger civic effort. These affective traces of community—fleeting or sustained—emerged through the tempo, rhythm, and openness of the game structures themselves. For designers of civic technologies, this highlights the role of play dimensions as levers for cultivating belonging: tuning for engagement, efficiency, and the social texture of participation. Play becomes a way to do sensing work *and* to feel part of something shared.

7.2.3 Methodological Insight.

Situated Play Design. NetGauge applied the three steps of Situated Play Design—uncovering play potentials, augmenting them with technology, and deploying to understand—to explore how civic play might foster community in context [2]. We grounded our designs in our prior work's bodystorming catalog and layered them with mobile technology to support location-aware triggers and broadband measurement. Through deployment in diverse environments, we observed how community surfaced in varied ways—via shared spatial rhythms, ambient awareness of others, and a collective goal of making broadband measurement fun. These findings extend Situated Play Design by showing how community can be catalyzed through the relational dynamics of co-present play in public space. Moreover, our use of mobility analysis offers a novel method for Situated Play Design: by quantifying movement patterns across contexts, it becomes possible to evaluate how players engage with space *and* how play structures facilitate—or inhibit—emergent forms of social and place attachment.

RITE. Beyond iterative refinement of gameplay and sensing protocols, the RITE process proved valuable as a pedagogical scaffold for designing community-centered, situated play. By embedding undergraduate researchers directly into rapid cycles of field testing, NetGauge created opportunities for students to engage deeply with how design decisions affect social dynamics and spatial experience. RITE provided a structure through which participants could iteratively learn about how usability, technical performance, intentional ambiguity, pacing, and environmental context shape players' sense of connection—to one another and to place. This hands-on exposure to breakdowns, redesign, and real-world feedback encouraged reflexive thinking about community as something that is designed through situated conditions, not just interaction mechanics. As a result, RITE became more than a method for prototyping—it served as an on-ramp for cultivating civic-minded designers capable of thinking relationally about infrastructure, play, and social belonging.

Mixed Methods. Our mixed-methods approach was a way of making community visible across different registers of interaction, space, and perception. Telemetry and mobility data surfaced patterned proximity and overlapping trajectories that would otherwise go unnoticed in individually played games. Video coding revealed how players responded to one another's movement without formal social mechanics—through peripheral vision, synchronized pauses, or subtle route mirroring. Surveys captured emotional and cognitive traces of collective experience, such as feeling "part of something bigger" despite minimal interaction in some games. Community is not a fixed output, but an emergent property of shared spatial rhythms, narrative alignment, and environmental co-presence. This methodological layering points to a broader implication: understanding how civic technologies foster belonging requires tools that operate across behavioral, environmental, and affective dimensions. Platforms like NetGauge won't generate community by design alone—but mixed-method analysis can help identify the conditions under which community becomes possible, felt, and durable. For researchers and designers, this means using instrumentation not just to log actions, but to detect traces of alignment, attunement, and shared civic orientation in the wild.

7.2.4 Future Work. Future work will extend NetGauge's capacity to foster community through civic play by scaling both its reach and its co-creative potential. First, we plan to formalize the RITE process as a pedagogical tool—training students to design for community by iteratively observing how social dynamics emerge through space, interaction, and narrative framing. Second, we aim to support game creation workshops with community members, educators, and designers to expand the platform's catalog and diversify the types of communal experiences it can support. As the platform scales, we will explore how persistent gameplay across neighborhoods or regions can surface new forms of local cohesion or infrastructural awareness. Finally, we are beginning to investigate whether AI-generated or AI-co-created games—based on prompts, place data, or community-authored themes—can support relational, meaningful play without sacrificing situatedness or inclusivity. These directions aim to move beyond platform engagement toward community authorship and infrastructural stewardship through play.

7.2.5 Limitations. While NetGauge was designed to support community engagement through situated play, several limitations constrain how these outcomes can be interpreted. First, community connection in our study was primarily inferred from behavioral patterns, self-report data, and observational cues—none of which provide direct evidence of sustained social bonds. Second, the short duration and episodic nature of gameplay limited our ability to assess long-term or repeated engagement with others or place. Finally, the diversity of player backgrounds and contexts was limited, raising questions about how these dynamics might vary across cultures, geographies, or levels of digital literacy. Future work will need to develop more robust instruments for capturing relational outcomes, and explore how game structure, community authorship, and longer-term engagement influence the depth and durability of civic connection.

8 CONCLUSION

NetGauge demonstrates how broadband measurement can be reimagined as a situated, participatory, and socially meaningful activity through game-based design. Across two research questions, we examined how games shape crowdsourced spatial sensing outcomes (*RQ1*) and how they foster community among players and environments (*RQ2*). Our empirical results show that game-based approaches outperformed a Speed Test control in several key areas: our games produced significantly greater movement distance, number of broadband measurements, time exploring, and sense of connection to players and space. These outcomes were driven by game mechanics that embedded measurement into intentional design affordances. These findings extend Situated Play Design and civic technology literature by showing how intentional ambiguity, environmental affordances, and game structure co-produce both data and connection. Methodologically, our use of mobility analysis, RITE, and mixed methods provided actionable insight across spatial, technical, and relational dimensions of play. As debates around digital equity, infrastructure accountability, and participatory sensing evolve, NetGauge offers a model for rethinking who gets to generate data, how that process unfolds, and what kinds of civic relationships it can support. Future work will scale the platform, explore co-authored and AI-assisted game creation, and continue to treat play not as an add-on, but as an engine for civic design, community-building, and infrastructural imagination.

REFERENCES

[1] Vivek Adarsh, Michael Nekrasov, Udit Paul, Tarun Mangla, Arpit Gupta, Morgan Vigil-Hayes, Ellen Zegura, and Elizabeth Belding. 2021. Coverage is Not Binary: Quantifying Mobile Broadband Quality in Urban, Rural, and Tribal Contexts. In *2021 International Conference on Computer Communications and Networks (ICCCN)*. IEEE, New York City, 1–9. https://doi.org/10.1109/ICCCN52240.2021.9522152

[2] Ferran Altarriba Bertran. 2021. *Situated Play Design: Co-creating the Playful Potential of Future Technology*. Ph. D. Dissertation. UC Santa Cruz.

[3] Ferran Altarriba Bertran, Oğuz 'Oz Buruk, Velvet Spors, and Juho Hamari. 2023. Playful Inspiration for a New Wave of Joyful Forest Technology. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. ACM, Pittsburgh PA USA, 1886–1903. https://doi.org/10.1145/3563657.3596015

[4] Ferran Altarriba Bertran, Jared Duval, Laura Bisbe Armengol, Ivy Chen, Victor Dong, Binaisha Dastoor, Adrià Altarriba Bertran, and Katherine Isbister. 2021. A Catalog of Speculative Playful Urban Technology Ideas: Exploring the Playful Potential of Smart Cities. In *Academic Mindtrek 2021*. ACM, Tampere/Virtual Finland, 60–71. https://doi.org/10.1145/3464327.3464374

[5] Ferran Altarriba Bertran, Jared Duval, Katherine Isbister, Danielle Wilde, Elena Márquez Segura, Oscar Garcia Pañella, and Laia Badal León. 2019. Chasing Play Potentials in Food Culture to Inspire Technology Design. In *Extended Abstracts of the Annual Symposium on Computer-Human Interaction in Play Companion Extended Abstracts (CHI PLAY '19 Extended Abstracts)*. Association for Computing Machinery, New York, NY, USA, 829–834. https://doi.org/10.1145/3341215.3349586

[6] Ferran Altarriba Bertran, Jared Duval, Elena Márquez Segura, Laia Turmo Vidal, Yoram Chisik, Marina Juanet Casulleras, Oscar Garcia Pañella, Katherine Isbister, and Danielle Wilde. 2020. Chasing Play Potentials in Food Culture: Learning from Traditions to Inspire Future Human-Food Interaction Design. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. ACM, Eindhoven Netherlands, 979–991. https://doi.org/10.1145/3357236.3395575

[7] Ferran Altarriba Bertran, Núria Jordán Sánchez, Jordi Márquez Puig, Maria Llop Cirera, Ernest Forts Plana, Ester Montero Collell, Oriol Martín Capdeferro, Oğuz 'Oz' Buruk, and Juho Hamari. 2023. The WildTech Experience: A Playful Installation for Walking Through the Outcomes of a One-Month Backpacking Study. In *Designing Interactive Systems Conference*. ACM, Pittsburgh PA USA, 246–250. https://doi.org/10.1145/3563703.3596650

[8] Ferran Altarriba Bertran, Jordi Márquez Puig, Maria Llop Cirera, Eva Forest Illas, Joan Planas Bertran, Ernest Forts Plana, Oğuz 'Oz' Buruk, Çağlar Genç, Mattia Thibault, and Juho Hamari. 2023. Designing and Using the Wild Probes Toolkit (v1) to Co-Design From-the-Wild. In *Proceedings of the 2023 ACM Designing Interactive Systems Conference*. ACM, Pittsburgh PA USA, 765–778. https://doi.org/10.1145/3563657.3596102

[9] Ferran Altarriba Bertran, Elena Márquez Segura, Jared Duval, and Katherine Isbister. 2019. Chasing Play Potentials: Towards an Increasingly Situated and Emergent Approach to Everyday Play Design. In *Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19)*. Association for Computing Machinery, New York, NY, USA, 1265–1277. https://doi.org/10.1145/3322276.3322325

[10] Ferran Altarriba Bertran, Elena Márquez Segura, Jared Duval, and Katherine Isbister. 2019. Designing for Play That Permeates Everyday Life: Towards New Methods for Situated Play Design. In *Proceedings of the Halfway to the Future Symposium 2019 (HTTF 2019)*. Association for Computing Machinery, New York, NY, USA, 4. https://doi.org/10.1145/3363384.3363400

[11] Ferran Altarriba Bertran, Laia Turmo Vidal, Ella Dagan, Jared Duval, Elena Márquez Segura, and Katherine Isbister. 2020. Chasing Play with Instagram: How Can We Capture Mundane Play Potentials to Inspire Interaction Design?. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–8. https://doi.org/10.1145/3334480.3382913

[12] Jon Back, Caspar Heeffer, Susan Paget, Andreas Rau, Eva Lotta Sallnäs Pysander, and Annika Waern. 2016. Designing for Children's Outdoor Play. In *Proceedings of the 2016 ACM Conference on Designing Interactive Systems*. ACM, Brisbane QLD Australia, 28–38. https://doi.org/10.1145/2901790.2901875

[13] Jon Back, Elena Márquez Segura, and Annika Waern. 2017. Designing for Transformative Play. *ACM Trans. Comput.-Hum. Interact.* 24, 3 (April 2017), 18:1–18:28. https://doi.org/10.1145/3057921

[14] Ian Bogost. 2015. Why gamification is bullshit. *The gameful world: Approaches, issues, applications* 65 (2015), 65 – 79. tex.ids= bogost_why_2015 publisher: The MIT press London, UK.

[15] Elizabeth Boyle, Thomas M Connolly, and Thomas Hainey. 2011. The Role of Psychology in Understanding the Impact of Computer Games. *Entertainment Computing* 2, 2 (2011), 69–74.

[16] Elizabeth A Boyle, Ewan W MacArthur, Thomas M Connolly, Thomas Hainey, Madalina Manea, Anne Kärki, and Peter Van Rosmalen. 2014. A Narrative Literature Review of Games, Animations and Simulations to Teach Research Methods and Statistics. *Computers & Education* 74 (2014), 1–14.

[17] Stuart Brown. 2009. *Play: How It Shapes the Brain, Opens the Imagination, and Invigorates the Soul.* Avery/Penguin Group USA, New York, NY, USA, New York, NY.

[18] Richard Buchanan. 1992. Wicked Problems in Design Thinking. *Design Issues* 8, 2 (1992), 5–21. https://doi.org/10.2307/1511637 jstor:1511637

[19] Seth Cooper. 2014. Massively Multiplayer Research: Gamification and (Citizen) Science. In *The Gameful World* (1 ed.). MIT Press, Boston, 688.

[20] Sebastian Deterding, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. From game design elements to gamefulness: defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. ACM, Tampere Finland, 9–15. https://doi.org/10.1145/2181037.2181040

[21] Sean J. Dorr, James W. Rock, Vicente M. Diaz, and Daniel F. Keefe. 2024. Toward Indigenous Presence: A Radical Relationality Approach for Designing Mixed-Reality Indigenous Data Experiences. *IEEE Computer Graphics and Applications* 44, 6 (Nov. 2024), 61–69. https://doi.org/10.1109/mcg.2024.3456868

[22] Anders Drachen, Pejman Mirza-Babaei, and Lennart E Nacke. 2018. *Games User Research*. Oxford University Press.

[23] Jared Duval, Ferran Altarriba Bertran, Siying Chen, Melissa Chu, Divya Subramonian, Austin Wang, Geoffrey Xiang, Sri Kurniawan, and Katherine Isbister. 2021. Chasing Play on TikTok from Populations with Disabilities to Inspire Playful and Inclusive Technology Design. In *Proceedings of the 2021 Conference on Conference on Human Factors in Computing Systems (CHI '21), May 8–13*. ACM, Online (originally Yokohama, Japan), 15. https://doi.org/10.1145/3411764.3445303

[24] Jared Duval, Shelby Hagemann, Tochukwu Arinze Ikwunne, Dayra Quinonez, and Morgan Vigil-Hayes. 2024. Co-Designing Location-based Games for Broadband Data Collection. In *Designing Interactive Systems Conference*. ACM, IT University of Copenhagen Denmark, 2057–2072. https://doi.org/10.1145/3643834.3661502

[25] Jared Duval, Zachary Rubin, Elena Márquez Segura, Natalie Friedman, Milla Zlatanov, Louise Yang, and Sri Kurniawan. 2018. SpokeIt: Building a Mobile Speech Therapy Experience. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM, Barcelona Spain, 1–12. https://doi.org/10.1145/3229434.3229484

[26] Jared Duval, Laia Turmo Vidal, Elena Márquez Segura, Yinchu Li, and Annika Waern. 2023. Reimagining Machine Learning's Role in Assistive Technology by Co-Designing Exergames with Children Using a Participatory Machine Learning Design Probe. In *The 25th International ACM SIGACCESS Conference on Computers and Accessibility*. ACM, New York NY USA, 1–16. https://doi.org/10.1145/3597638.3608421

[27] Jared Scott Duval. 2022. *Playful Health Technology: A Participatory, Research through Design Approach to Applications for Wellness*. Ph. D. Dissertation. UC Santa Cruz.

[28] Federal Communications Commission. 2024. FCC Speed Test App. https://www.fcc.gov/BroadbandData/speed-test-app-faq.

[29] Olivier Friard and Marco Gamba. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7, 11 (2016), 1325–1330. publisher: Wiley Online Library.

[30] Zoraida Frias, William Lehr, and Volker Stocker. [n. d.]. Building an Ecosystem for Mobile Broadband Measurement: Methods and Policy Challenges. 49, 5 ([n. d.]), 102905. https://doi.org/10.1016/j.telpol.2025.102905

[31] Zoraida Frias, William Lehr, Volker Stocker, and Luis Mendo. 2023. Measuring Mobile Broadband Challenges and Implications for Policymaking. *32nd European Conference of the International Telecommunications Society (ITS)* 32, 32 (2023).

[32] Batya Friedman and David Hendry. 2019. *Value Sensitive Design: Shaping Technology with Moral Imagination*. The MIT Press, Cambridge, Massachusetts.

[33] Kathrin Gerling, Kieran Hicks, Michael Kalyn, Adam Evans, and Conor Linehan. 2016. Designing Movement-based Play With Young People Using Powered Wheelchairs. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 4447–4458. https://doi.org/10.1145/2858036.2858070

[34] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. 2008. Understanding Individual Human Mobility Patterns. *Nature* 453, 7196 (2008), 779–782.

[35] Roland Graf, Pallavi Benawri, Amy E. Whitesall, Dashiell Carichner, Zixuan Li, Michael Nebeling, and Hun Seok Kim. 2019. iGYM: An Interactive Floor Projection System for Inclusive Exergame Environments. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play*. ACM, Barcelona Spain, 31–43. https://doi.org/10.1145/3311350.3347161

[36] Andrea Hamm, Yuya Shibuya, Stefan Ullrich, and Teresa Cerratto Cerratto Pargman. 2021. What Makes Civic Tech Initiatives To Last Over Time? Dissecting Two Global Cases. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–17. https://doi.org/10.1145/3411764.3445667

[37] Benjamin Hanussek. 2024. Ubisoft's Notre-Dame: Digital Gaming for Material Heritage's Sake. *Loading* 16, 26 (May 2024), 1–15. https://doi.org/10.7202/1111257ar

[38] Lobna Hassan and Juho Hamari. 2020. Gameful civic engagement: A review of the literature on gamification of e-participation. *Government Information Quarterly* 37, 3 (July 2020), 101461. https://doi.org/10.1016/j.giq.2020.101461

[39] Matthias Hirth, Tobias Hossfeld, Marco Mellia, Christian Schwartz, and Frank Lehrieder. 2015. Crowdsourced Network Measurements: Benefits and Best Practices. *Computer Networks* 90 (2015), 85–98.
[40] Kristina Höök. 2018. *Designing with the Body: Somaesthetic Interaction Design.* The MIT Press, Cambridge, MA.
[41] Kristina Höök and Jonas Löwgren. 2012. Strong Concepts: Intermediate-level Knowledge in Interaction Design Research. *ACM Transactions on Computer-Human Interaction* 19, 3 (Oct. 2012), 1–18. https://doi.org/10.1145/2362364.2362371
[42] International Telecommunications Union. 2024. ITU Facts & Figures.
[43] Katherine Isbister. 2016. *How Games Move Us: Emotion by Design.* MIT Press, Cambridge, MA.
[44] Katherine Isbister, Elena Márquez Segura, and Edward F. Melcer. 2018. Social Affordances at Play: Game Design Toward Socio-Technical Innovation. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM, Montreal QC Canada, 1–10. https://doi.org/10.1145/3173574.3173946
[45] Sara Louise Kindon, Rachel Pain, and Mike Kesby. 2007. *Participatory Action Research, Approaches and Methods: Connecting People, Participation and Place.* Number 22 in Routledge Studies in Human Geography. Routledge, London.
[46] Jan HG Klabbers. 2009. *The Magic Circle: Principles of Gaming & Simulation.* Brill Sense.
[47] Chris Klimas. 2025. Twine: An Open-Source Tool for Telling Interactive, Nonlinear Stories. Interactive Fiction Technology Foundation.
[48] Samuli Laato, Daniel Fernández Galeote, Ferran Altarriba Bertran, Konstantinos Papangelis, and Juho Hamari. 2023. How Location-Based Games Incentivize Moving About: A Study in the Context of Nature-Going. *Proceedings of the ACM on Human-Computer Interaction* 7, CHI PLAY (Sept. 2023), 642–664. https://doi.org/10.1145/3611044
[49] Sybille Lammes and Clancy Wilmott. 2018. The Map as Playground: Location-based Games as Cartographical Practices. *Convergence* 24, 6 (2018), 648–665.
[50] J. Richard Landis and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (March 1977), 159. https://doi.org/10.2307/2529310 jstor:2529310
[51] Ho Ming Lau, Johannes H. Smit, Theresa M. Fleming, and Heleen Riper. 2017. Serious Games for Mental Health: Are They Accessible, Feasible, and Effective? A Systematic Review and Meta-analysis. *Frontiers in Psychiatry* 7 (Jan. 2017). https://doi.org/10.3389/fpsyt.2016.00209
[52] Zhijing Li, Ana Nika, Xinyi Zhang, Yanzi Zhu, Yuanshun Yao, Ben Y. Zhao, and Haitao Zheng. 2017. Identifying Value in Crowdsourced Wireless Signal Measurements. In *Proceedings of the 26th International Conference on World Wide Web.* International World Wide Web Conferences Steering Committee, Perth Australia, 607–616. https://doi.org/10.1145/3038912.3052563
[53] Sherlock A Licorish, Jade Li George, Helen E Owen, and Ben Daniel. 2017. "Go Kahoot!" Enriching Classroom Engagement, Motivation and Learning Experience with Games. In *Proceedings of the 25th International Conference on Computers in Education.* Asia-Pacific Society for Computers in Education / Asia-Pacific Society for Computers in Education, New Zealand, 755–764.
[54] Tarun Mangla, Udit Paul, Arpit Gupta, Nicole P. Marwell, and Nick Feamster. 2022. Internet Inequity in Chicago: Adoption, Affordability, and Availability. *SSRN Electronic Journal* (2022). https://doi.org/10.2139/ssrn.4182994
[55] Tarun Mangla, Esther Showalter, Vivek Adarsh, Kipp Jones, Morgan Vigil-Hayes, Elizabeth Belding, and Ellen Zegura. 2022. A Tale of Three Datasets: Characterizing Mobile Broadband Access in the US. *Commun. ACM* 65, 3 (2022), 67–74.
[56] Elena Márquez Segura, Laia Turmo Vidal, and Asreen Rostami. 2016. Bodystorming for Movement-Based Interaction Design. *Human Technology* 12, 2 (Nov. 2016), 193–251. https://doi.org/10.17011/ht/urn.201611174655
[57] Elena Márquez Segura, Laia Turmo Vidal, Asreen Rostami, and Annika Waern. 2016. Embodied Sketching. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM, San Jose California USA, 6014–6027. https://doi.org/10.1145/2858036.2858486
[58] Elena Márquez Segura, Laia Turmo Vidal, Annika Waern, Jared Duval, Luis Parrilla Bel, and Ferran Altarriba Bertran. 2021. Physical Warm-up Games: Exploring the Potential of Play and Technology Design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–14.
[59] Elena Márquez Segura, Annika Waern, Luis Márquez Segura, and David López Recio. 2016. Playification: The PhySeEar Case. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play.* ACM, 376–388.
[60] Tuuli Mattelmäki. 2008. *Design Probes* ([2nd ed.] ed.). University of Art and Design, Helsinki.
[61] Michael C Medlock, Dennis Wixon, Mick McGee, and Dan Welsh. 2005. The Rapid Iterative Test and Evaluation Method: Better Products in Less Time. In *Cost-Justifying Usability.* Elsevier, 489–517.
[62] James Miller. 2020. Measuring Broadband America's FCC Speed Test App for Android and iOS: Crowdsourcing Mobile Broadband Performance.
[63] Tony Morelli, Lauren Lieberman, John Foley, and Eelke Folmer. 2014. An Exergame to Improve Balance in Children Who Are Blind. In *FDG.* Society for the Advancement of the Science of Digital Games, Ft. Lauderdale, FL, 4.
[64] Syed Tauhidun Nabi, Zhuowei Wen, Brooke Ritter, and Shaddi Hasan. 2024. Red is Sus: Automated Identification of Low-Quality Service Availability Claims in the US National Broadband Map. In *Proceedings of the 2024 ACM on Internet*

Measurement Conference (Madrid, Spain) *(IMC '24)*. Association for Computing Machinery, New York, NY, USA, 2–18. https://doi.org/10.1145/3646547.3688441

[65] Fiona Fui-Hoon Nah, Qing Zeng, Venkata Rajasekhar Telaprolu, Abhishek Padmanabhuni Ayyappa, and Brenda Eschenbrenner. 2014. Gamification of Education: A Review of Literature. In *HCI in Business*, David Hutchison, Takeo Kanade, Josef Kittler, Jon M. Kleinberg, Alfred Kobsa, Friedemann Mattern, John C. Mitchell, Moni Naor, Oscar Nierstrasz, C. Pandu Rangan, Bernhard Steffen, Demetri Terzopoulos, Doug Tygar, Gerhard Weikum, and Fiona Fui-Hoon Nah (Eds.). Vol. 8527. Springer International Publishing, Cham, 401–409. https://doi.org/10.1007/978-3-319-07293-7_39

[66] Donald A. Norman. 2004. *Emotional Design: Why We Love (or Hate) Everyday Things*. Basic Books, New York, NY.

[67] Beatriz Palacios Abad, Elizabeth Belding, Morgan Vigil-Hayes, and Ellen Zegura. 2022. Note: Towards Community-Empowered Network Data Action. In *Proceedings of the 5th ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (Compass '22)*. Association for Computing Machinery, Seattle, WA, USA, 585–588. https://doi.org/10.1145/3530190.3534836

[68] Beatriz Palacios Abad, Elizabeth Belding, Morgan Vigil-Hayes, and Ellen Zegura. 2024. Mending the Fabric: the Contentious, Collaborative Work of Repairing Broadband Maps. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW2, Article 464 (Nov. 2024), 24 pages. https://doi.org/10.1145/3687003

[69] Luca Pappalardo, Filippo Simini, Gianni Barlacchi, and Roberto Pellungrini. 2021. Scikit-mobility: a Python library for the analysis, generation and risk assessment of mobility data. https://doi.org/10.48550/arXiv.1907.07062 arXiv:1907.07062 [physics].

[70] Jeff Patton. 2008. Getting Software RITE. *IEEE Software* 25, 3 (2008), 20–21. https://doi.org/10.1109/MS.2008.62

[71] Udit Paul, Jiamo Liu, Mengyang Gu, Arpit Gupta, and Elizabeth Belding. [n. d.]. The Importance of Contextualization of Crowdsourced Active Speed Test Measurements. In *Proceedings of the 22nd ACM Internet Measurement Conference* (Nice France, 2022-10-25). ACM, 274–289. https://doi.org/10.1145/3517745.3561441

[72] Brian A Primack, Mary V Carroll, Megan McNamara, Mary Lou Klem, Brandy King, Michael Rich, Chun W Chan, and Smita Nayak. 2012. Role of Video Games in Improving Health-Related Outcomes: A Systematic Review. *American journal of preventive medicine* 42, 6 (2012), 630–638.

[73] Merten Reglitz. 2020. The Human Right to Free Internet Access. *Journal of Applied Philosophy* 37, 2 (May 2020), 314–331. https://doi.org/10.1111/japp.12395

[74] Marco Romano, Paloma Díaz, and Ignacio Aedo. 2022. Gamification-less: may gamification really foster civic participation? A controlled field experiment. *Journal of Ambient Intelligence and Humanized Computing* 13, 9 (Sept. 2022), 4451–4465. https://doi.org/10.1007/s12652-021-03322-6

[75] Johnny Saldaña. 2015. Developing a Codebook for a Case Participant. In *The Coding Manual for Qualitative Researchers*. SAGE Publications Ltd, 219–222.

[76] Katie Seaborn and Deborah I Fels. 2015. Gamification in Theory and Action: A Survey. *International Journal of human-computer studies* 74 (2015), 14–31.

[77] Elena Márquez Segura, Laia Turmo Vidal, Luis Parrilla Bel, and Annika Waern. 2019. Using Training Technology Probes in Bodystorming for Physical Training. In *Proceedings of the 6th International Conference on Movement and Computing*. ACM, Tempe AZ USA, 1–8. https://doi.org/10.1145/3347122.3347132

[78] Wamika Sharma, Weng Marc Lim, Satish Kumar, Aastha Verma, and Rajeev Kumra. 2024. Game on! A State-of-the-Art Overview of Doing Business with Gamification. *Technological Forecasting and Social Change* 198 (2024), 122988.

[79] Varshika Srinivasavaradhan, Owen Park, and Elizabeth Belding. 2024. Mapping Cellular Network Evolution and Infrastructure Criticality: A Nationwide Analysis. In *Proceedings of the TPRC2024 the Research Conference on Communications, Information and Internet Policy*.

[80] Jay Shiro Tashiro and David Dunlap. 2007. The Impact of Realism on Learning Engagement in Educational Games. In *Proceedings of the 2007 Conference on Future Play - Future Play '07*. ACM Press, Toronto, Canada, 113. https://doi.org/10.1145/1328202.1328223

[81] Laia Turmo Vidal and Jared Duval. 2024. Ambiguity as a Resource to Design for a Plurality of Bodies. In *Proceedings of the Halfway to the Future Symposium*. ACM, Santa Cruz CA USA, 1–9. https://doi.org/10.1145/3686169.3686176

[82] Laia Turmo Vidal and Elena Márquez Segura. 2018. Documenting the Elusive and Ephemeral in Embodied Design Ideation Activities. *Multimodal Technologies and Interaction* 2, 3 (June 2018), 35. https://doi.org/10.3390/mti2030035

[83] Michael Ulm, Peter Widhalm, and Norbert Brandle. 2015. Characterization of Mobile Phone Localization Errors with OpenCellID Data. In *2015 4th International Conference on Advanced Logistics and Transport (ICALT)*. IEEE, Valenciennes, France, 100–104. https://doi.org/10.1109/ICAdLT.2015.7136601

[84] Laia Turmo Vidal, Elena Márquez Segura, and Annika Waern. 2018. Sensory Bodystorming for Collocated Physical Training Design. In *Proceedings of the 10th Nordic Conference on Human-Computer Interaction (NordiCHI '18)*. ACM, New York, NY, USA, 247–259. https://doi.org/10.1145/3240167.3240224

[85] Morgan Vigil-Hayes, Italo Santos, and Dayra Quiñonez. 2024. Critical Analysis of the Inclusivity of Crowdsourced Internet Measurement Apps. In *The 52nd Research Conference on Communications, Information and Internet Policy (TPRC51)* (Washington, DC, USA). Available at SSRN 4911129.
[86] Annika Waern and Jon Back. 2017. Activity as the Ultimate Particular of Interaction Design. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 3390–3402. https://doi.org/10.1145/3025453.3025990
[87] Steffen P. Walz and Sebastian Deterding. 2015. *The Gameful World Approaches, Issues, Applications*. The MIT Press, Cambridge.
[88] Shaojung Sharon Wang and Chih-Ting Hsieh. 2020. Ubiquitous Pokémon Go: Human–Environment Relationships and the Location-Based Augmented Reality Game. *Environment and Behavior* 52, 7 (Aug. 2020), 695–725. https://doi.org/10.1177/0013916518817878
[89] Voravika Wattanasoontorn, Imma Boada, Rubén García, and Mateu Sbert. 2013. Serious Games for Health. *Entertainment Computing* 4, 4 (2013), 231–247.
[90] Saovapa Wichadee and Fasawang Pattanapichet. 2018. Enhancement of Performance and Motivation through Application of Digital Games in an English Language Class. *Teaching English with Technology* 18, 1 (2018), 77–92.
[91] Peter Williams. 2014. Collaboration in the Gameful World. In *The Gameful World* (1 ed.). MIT Press, Boston, 688.
[92] Helen Woolley and Alison Lowe. 2013. Exploring the Relationship between Design Approach and Play Value of Outdoor Play Spaces. *Landscape Research* 38, 1 (Feb. 2013), 53–74. https://doi.org/10.1080/01426397.2011.640432

Received 11 September 2025