

## Case for Support

### Objectives

The intersection of video gaming and player health is an emergent and critical area of inquiry within the broader context of digital media research<sup>1</sup>. The UK Government's [Video Games Research Framework](#) acknowledges the cultural significance and pervasive nature of video games, emphasizing the need for a nuanced understanding of their psychological impacts. This framework calls for robust, evidence-based research to inform policy and practice, recognizing the potential of video games to contribute to various aspects of society, including education, well-being, and economic growth. In response to this call, our proposed project, GamePlay.Science, will entail the creation and validation of an innovative open source internet platform that will enable ethical and secure collection of large-scale gameplay data. This platform is designed to overcome the gaps identified by the government's framework, particularly the need for transparently and ethically sourced data that can be used to examine the complex relationships between gaming and player health outcomes. The overarching ambition of GamePlay.Science is to develop and validate a secure and user-centric system for video game players to donate their gameplay data to science and to provide a repository of rich, pseudonymised large-scale datasets for researchers. By aligning with the UK Government's framework, which advocates for collaborative efforts between academia, industry, and policymakers, GamePlay.Science will facilitate impactful research that can drive innovation in game design, promote mental health, and inform responsible regulatory approaches.

The confluence of technological advancements, the proliferation but opacity of digital data, and the pressing need for transparent research methodologies signifies that the time is ripe for an independent open source game research platform like GamePlay.Science. The current landscape, underscored by the ESRC's Digital Footprints Accelerator Scheme, calls for innovative projects that not only push the boundaries of methodological development in the social sciences but also address the challenges of data representativeness, bias, and ethical use. GamePlay.Science, with its commitment to open source principles, stands at the vanguard of this movement, poised to leverage the vast, yet untapped, reservoirs of digital footprint data (DFD) generated by the gaming industry. By providing a platform that facilitates the ethical collection, analysis, and sharing of gaming data, GamePlay.Science is set to catalyse a transformational shift in our understanding of the interplay between digital engagement and societal outcomes. This aligns perfectly with the strategic objectives of the ESRC to foster a robust, ethical, and innovative research environment that harnesses DFD for the public good. To this end, the project will achieve the following five objectives:

1. **Validation of GamePlay.Science:** An internet-based platform that allows video game players to donate their play data through account linking and/or EU/UK GDPR Data Requests.
2. **Open Source Development and Sharing:** The platform's source code and documentation will be developed and maintained, in accordance with best practices in free and open source software development, on the project's GitHub repository for public use and adaptation.
3. **Methodological Report:** A comprehensive report detailing the platform's infrastructure, required team skills, data manipulation, and usage will be submitted to a suitable academic outlet (e.g. *Behavioral Research Methods*). This document will also include validation studies that ensure the data and methodology are accurately able to record play behaviour data.
4. **Workshop at SIPS 2024:** A workshop will be hosted to encourage the platform's adoption, where researchers can learn about deployment, get help in using it, and provide feedback.
5. **Stakeholder Meetings:** Engagements with UK games industry representatives (e.g., UKIE), UK policymakers (e.g., DCMS), and stakeholders in Europe (e.g., VideoGames Europe) to foster collaborative conversations and share project insights.

### Background

Independent access to video game play data is crucial for researchers to ensure that studies on the impact of gaming are comprehensive, unbiased, and serve the public interest rather than corporate agendas. When technology companies control data access, there is a potential for conflicts of interest, as these entities may have a vested interest in promoting certain narratives or suppressing unfavourable findings<sup>2</sup>. Open access to data enables researchers to conduct transparent studies that can critically assess the effects of video games on mental health, social behaviour, and cognitive development without corporate filters. This autonomy is essential for advancing scientific knowledge,

## Case for Support

informing evidence-based policy, and fostering a balanced understanding of the benefits and risks associated with video game play.

We started work on GamePlay.Science in late 2022 to address this need by providing a complementary means of collecting digital footprint data from games that leverages infrastructure provided by the games industry but does not require direct involvement from games companies. This contrasts with the status quo means of collecting such data, which involves direct collaborative partnerships between academia and the games industry. Such partnerships are powerful and can benefit from industry's insider knowledge of player engagement dynamics and practical applications. However, industry has historically been hesitant to collaborate openly with independent scientific researchers, and potential collaborations create the potential for conflicts of interest, require complex and lengthy negotiations, and are not equally accessible across universities and research groups. We believe that direct collaborations need to be supplemented by an alternative open source pathway through which diverse researchers can freely, quickly, and independently collect data on video game play.

Our team developed and have just tested an alpha version of GamePlay.Science, placing us in an ideal position for greatly expanding and validating an open source games research platform. Our pioneering research at the intersection of psychology and technology<sup>3,4</sup>, expertise in quantitative social science methodology and related open source software development<sup>5-7</sup>, and successful history of engagement with the games industry<sup>8,9</sup> ensures an informed approach that addresses the extant needs for scientific rigour and overcomes the intrinsic limitations of industry collaborations. Our pilot work (detailed below in the Current Status section) and team's proven track record of navigating the intricacies of industry collaboration, commitment to open science principles and methods, and ability to communicate complex findings in accessible reports for stakeholders at different levels positions us uniquely to develop a platform that not only meets the needs of researchers but also garners the support and trust of the gaming community. Such a platform, built on the principles of transparency and utility, would be an invaluable asset to scholars seeking to advance our understanding of the impacts of gaming on society.

### Method

The core contribution of this project is the expansion and validation of a new method of collecting behavioural data on video game play and associating it with other data. GamePlay.Science is intended as a pioneering web application that allows participants to link their video game accounts securely and confidentially to a research platform. Central to our initiative is the principle of independence; the data we gather is the property of the players, not the industry, ensuring that the industry has no direct influence over the data collection, analysis, or publication processes. As an open source application, it will be freely available for other researchers to use and adapt, and thus fosters a community-driven approach to research without commercial interests. GamePlay.Science is designed with transparency in mind, allowing for external auditing of its codebase to ensure the integrity of research results. To collect live objective play data (i.e., in near-real time), GamePlay.Science will interface with many of today's most popular gaming platforms and accounts, which are central to the modern gaming experience. These platforms include PCs, which are favoured for their versatility and extensive game libraries, and consoles like PlayStation and Xbox. Gamers create accounts on these platforms to access games, track achievements, socialize, and make digital purchases, including microtransactions. Discord, a communication platform integral to the gaming community, and Steam, a leading PC gaming platform, are also part of the ecosystem that GamePlay.Science taps into for data collection. These platforms record and make available, via application programming interfaces (APIs), data streams regarding users' gaming behaviours, which GamePlay.Science then leverages for scientific research.

## Case for Support

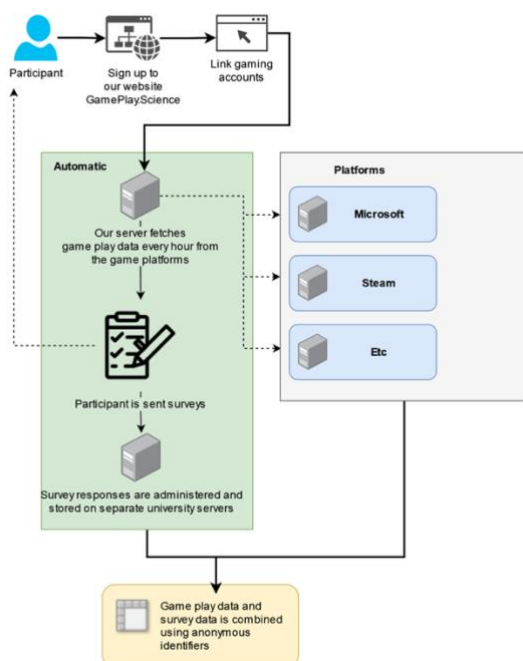


Figure 1. Schematic of how GamePlay.Science transparently combines participating players' data from multiple sources.

The system for live data capture is depicted in Figure 1. The platform ensures ethical and secure data access through OAuth, a widely adopted open standard protocol for online access control, that allows participants to grant data access without compromising their credentials. This process is part of a data donation model, where participants consent to sharing their gaming data, which is then linked to survey responses for research purposes. All live and historic data storage on GamePlay.Science will be minimal and pseudonymized, prioritizing participant privacy. Participants have full control over their data, with the ability to withdraw consent and request data deletion at any time. The data is collected, stored, and analysed following strict University of Oxford protocols for ethics and data management. After thorough pseudonymisation, the de-identified data, along with analysis code and materials, will be made publicly available under a CC-BY 4.0 license, reinforcing the platform's commitment to open science and data transparency.

Importantly, the live data that GamePlay.Science can reach does not have a historic scope: Players cannot contribute their play data from before they signed up to the platform. We therefore implement a functionality in

GamePlay.Science that allows collecting objective historic game play data: GamePlay.Science will include an interface that facilitates players' ability to exercise their right under the GDPR and UK GDPR to access a copy of the data held about them, and to freely donate that data for research. The team will develop an upload feature for player-participants to share these data files and extract the historic play data ethically and securely. In the first instance we add this feature for historic data from PlayStation, as it presents the lowest burden for users to exercise their rights under the law and the data come in a fixed machine-readable format that the team has experience with. In the future, this functionality has potential to be rolled out to any number of gaming and technology companies.

### Current Status

An unfunded alpha version of the platform has been in development since late 2022. So far, we have achieved basic functionality with two kinds of account integrations (Steam and Discord). We have conducted a small-scale pilot study with 55 participants, with initial results suggesting that the system can reliably capture gameplay data. This initial success demonstrates proof of concept and serves to de-risk the project by ensuring that a subset of gameplay data can be collected. No manuscripts related to the alpha version of the project have been published, and the codebase currently lacks documentation and is not yet publicly available.

However, significant further development is needed before the platform can be shown to facilitate reliable and transparent research across the games research ecosystem. The measured success of the pilot helped us to identify the next steps needed to develop the alpha version into a fully-fledged resource for the research community. Specifically, the next stage of the project looks to achieve the following five goals.

**Development Goal 1.** Add the ability for users to request and upload data returned through GDPR and UK GDPR subject data access requests. Collecting data using account integrations requires that gaming platforms maintain infrastructure in the form of APIs and OAuth protocols such that players can authorise GamePlay.Science to retrieve their data. However, not all platforms have such systems (e.g., PlayStation does not support account integration with unapproved third-party platforms), and nor can this typically be used to track historic gameplay data. To complement account integration, we will therefore add an alternative pathway based on GDPR requests. The platform will first provide instructions about how players can request their data from gaming companies. Once the players have received their data, the platform will provide an interface wherein players can securely upload the data they receive. Data will be automatically stripped of any unnecessary identifying information or data that is not relevant to the research project. The remaining

## Case for Support

data will then be presented for inspection, such that users are able to give informed consent about each piece of data they share.

**Development Goal 2:** Expand the list of platforms for which data can be collected with further account integration options. In the alpha version, GamePlay.Science can collect data from PC gaming data through Steam's OpenID system, and indirectly from console gaming through Discord. However, this has notable limitations: Steam users must use specific privacy settings to allow us access to their data, and Discord users must separately link their Discord account with various game services. To address this, we will supplement these account integration options with two other platforms whose infrastructure supports potential linking, but that are not yet available within GamePlay.Science: Xbox and Epic Games. The addition of these will enable GamePlay.Science to directly capture data from the roughly 3 million Xbox users in the UK<sup>10</sup> and from players of Epic Games titles, which include some of the world's most popular games such as *Fortnite*.

**Development Goal 3.** Establish protocols for harmonising data from multiple sources. The pilot study with the alpha version, which compared only Steam and Discord data, indicated that these sources can diverge from each other, and are not straightforward to match up. With the addition of a new source of data in the form of GDPR requests, these divergences will become more complicated. To address this, we will conduct research with users explicitly targeted at collecting the same gameplay data using multiple methods, and will use qualitative and quantitative data to assess how these sources can best be merged to product the most accurate picture of play.

**Development Goal 4.** Assess the quality of the data generated by the platform and the user experience of donating gameplay data through validation studies with users. At present, little is known about how players view the process of sharing gameplay data for scientific researchers. To better understand this, we will solicit written feedback from users to understand their concerns or challenges with regard to donating data, and will interview a subset of these in depth to better understand their perception of how logged data concords with their personal records of time spent gaming. In doing so, we will also demonstrate how GamePlay.Science data can be seamlessly linked with psychological measures to understand the effects of video game play.

Together, achieving these four development goals will convert the promising potential shown in the alpha version of GamePlay.Science into a mature suite of data collection options that can support a diverse group of researchers in their investigations into the effects of video games over time.

### Data Collection and Analysis

To assess and validate the platform, we will conduct a study of 400 video game players, each generating platform-wide data that is not currently available in any other data source. This study will accomplish several goals: it will validate the reliability of the platform at scales typical for social science research; evaluate data quality by comparing three sources of gameplay behaviour data (account integrations, GDPR requests, and self-reported player diaries); and assess user's experiences of the platform and identify factors that influence their willingness to donate data. It will also produce a highly valuable dataset of real-world video game play across one or more platforms, linked with intensive longitudinal data of players' survey responses over 2 weeks, which is suitable for both descriptive analysis of video game behaviour, and inference about its relationship with mental health.

During evaluation, the population of interest is UK adult video game players. We will recruit participants on Prolific, given previous success and high data quality, and the ability to screen participants on factors such as gender identity, marital status, education, and video game play status. While we do not aim for a representative sample, Prolific facilitates diverse sampling among gender identities and sociodemographic backgrounds. This is important because GamePlay.Science could in the future be used for studying more targeted populations (e.g., neurodiverse video game players).

To get our target sample of 400 players, we first recruit 1500 individuals who report playing video games (A). Based on data from our alpha-version pilot, about 30% of such Prolific participants play games on any of the four platforms supported by GamePlay.Science (Steam, Xbox, Epic Games, and PlayStation) and are willing to link their play data (B). The remaining valid participants will be onboarded with detailed information about how the data is collected, stored, and used (C).

## Case for Support

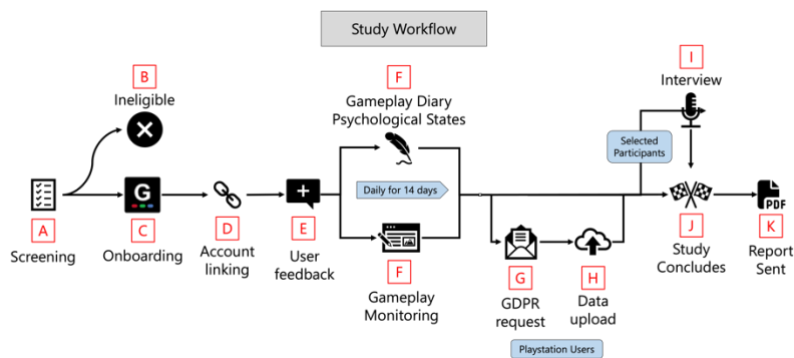


Figure 2. Overview of the participant recruitment and data collection procedure

Consenting participants will then complete the account linking process (D) and can provide feedback about the platform and their motivations and concerns about sharing gameplay data (E).

Next, we will invite the participants to a 2-week study, where they log what games they play and when into a provided diary, every evening for the following 14 days. Each day, participants will receive a reminder to complete the diary and a short

questionnaire with measures of their motivation and mental health (F).

At the end of the 14-day period, the subset of participants who are PlayStation users ( $n \approx 100$ ) will be provided with instructions for how to request a copy of their play data under GDPR law from Sony (G), who at present have the most streamlined portal for doing this. Participants can expect to receive their data from Sony in no more than 4 weeks. Upon receiving their data from the company, participants return to GamePlay.Science to upload, inspect, and share this data (H).

Finally, we will recruit 20 participants for 30-minute semi-structured interviews about their experiences using GamePlay.Science, their perceptions about the validity of the data collection methods, and their thoughts about potential divergences (I). We will sample these 20 participants based on the concordance between the three data types, such that some participants have close alignment between diaries, GamePlay.Science, and/or GDPR requests and others have greater divergence. After the study concludes and data has been processed, we will distribute the reports with insights about each player's gaming during the study (K).

The data generated during the study will first and foremost allow us to assess the platform's reliability at production scale and validity of resulting data. Regarding platform reliability, we will be able to identify any issues with account creation and deletion, user interface, functionality across operating systems and web browsers, and accessibility. Regarding data validity, comparing multiple data sources allows us to assess data convergence and make recommendations about best practices in selecting, implementing, and processing data from each. In this sense, the study serves to mitigate three of the primary risks we foresee for this research: technical failures or incompatibilities, divergences between data sources, and barriers for participants to donate data at scale.

We will analyse the qualitative data (i.e., the user feedback during account linking and transcripts of the interviews) using thematic analysis methods to evaluate how well the platform meets user needs and produces high-quality data. The quantitative data will further allow us to report naturalistic descriptions of digital play behaviour at a previously inaccessible level of detail. To date, seemingly basic questions such as how many games people play at one time, the frequency of alternation, and the stability of playtime over time remain unknown but are answerable using the data generated by GamePlay.Science. The data will also be openly available as part of the published article, enabling others to do more detailed analyses of gaming behaviour, motivation, and mental health.

### Alignment

The GamePlay.Science platform can enable a diverse group of scholars in the UK and beyond to conduct research in line with existing government interests, represented by the Video Games Research Framework (VGRF) and the Smart Data Research (ne Digital Footprints) initiatives. For example, the VGRF calls for research into "how and why people interact with video games [...]" and how this var[ies] across different player demographics", as well as into "whether, and if so how and to what extent, different types and genres of games have positive and negative impacts for players." Both of these are topics that benefit substantially from the availability of high-resolution, objective platform-level data collected with GamePlay.Science, in comparison to previous data that reflects engagement with just one (of potentially many) games a user plays<sup>8,11</sup>. Similarly, we see close alignment between GamePlay.Science and the Health and Wellbeing pillar of the Smart Data Research programme: The data we will generate can provide insights that help us better understand



## Case for Support

drivers of the mental health crisis, better identify how games can be used as a potential intervention to counteract it, and ultimately help players live healthier lives.

## Outputs and Beneficiaries

Sharing the source code and documentation of *GamePlay.Science* on a collaborative platform such as GitHub, according to best practices in open source software development, will be a significant boon for stakeholders within and beyond academia. This empowers other researchers to scrutinize, validate, and extend the platform's functionality, fostering a culture of transparency and reproducibility that is the cornerstone of rigorous scientific inquiry. For academic researchers, open source code facilitates peer collaboration, allowing for the collective refinement of tools and methodologies, which can accelerate the pace of discovery in video games research. Outside academia, practitioners, independent researchers, and developers can adapt and integrate the platform for specific research needs or public interest projects, such as supporting clinical treatment of people struggling with possible disordered gaming by providing clinicians access to information about their gaming. This openness not only democratizes research by levelling the playing field for all interested parties but also encourages a cross-pollination of ideas that can lead to breakthroughs in understanding player behaviour and the broader implications of video gaming on society.

Engaging directly with representatives of the video games industry in the UK Interactive Entertainment (UKIE) and Europe (VideoGames Europe) as well as civil servants working on the implementation of the Video Games Research Framework as the project develops will have benefits for all three parties. For example, it will give UK- and Europe-based games firms an opportunity to understand and positively contribute to the evolving landscape of rigorous and independent research through dialogue and advancing their own processes to improve compatibility with open source tools like Gameplay.Science. Engaging with DCMS constitutes another avenue through which the project can contribute to ongoing consultation and implementation of the VGRF.

Finally, we will provide a benefit to users and participants in the form of automated reports through which players can receive feedback about their play. At present, players generally receive only limited, coarse information about their play history from platforms (e.g., total time played over the lifetime of a game). By contrast, we will use the data from `GamePlay.Science` to produce easily digestible personal reports with detailed summaries of players' gaming and the development of their play over the course of the study (see Figure 4). These reports can incentivize participation, and in the future can be a vital tool for players seeking to monitor their behaviour and play games in ways that support their wellbeing.

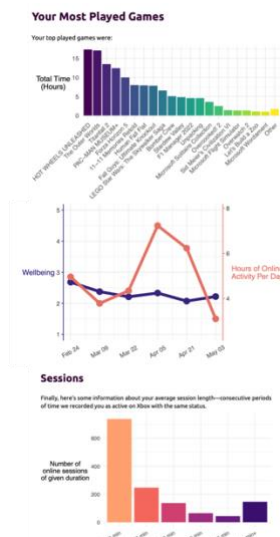


Figure 3. Example personal reports to be sent to participants

## Academic Outputs and Engagement

We engage the academic community regarding this project via two routes. First, we will host a workshop about GamePlay.Science at the 2024 Society for the Improvement of Psychological Science conference in Nairobi. In this workshop, we place emphasis on fostering engagement with early career researchers (ECR), who will have a unique opportunity to actively participate in shaping the future of the research tools we are developing. By providing a forum where ECRs can voice their needs and suggest features, the workshop ensures that the platform evolves in a direction that is informed by the fresh perspectives and diverse requirements of those at the early stages of their careers. The choice of Nairobi as the host city reflects a commitment to accessibility and inclusivity, offering a platform for researchers who do not have the chance to attend such events in North America or Europe and even less of a chance to engage and benefit from relationships with major game publishers who are based in these same regions.

Second, we will detail the development, validation, and use of `GamePlay.Science`, in a manuscript that we aim to submit for peer review at a methodologically oriented journal, such as *Behavior Research Methods*. This manuscript will also report on the platforms' technical details and critical analyses of qualitative and quantitative data accrued throughout the validation studies. Moreover, a comprehensive case study of *using* `GamePlay.Science` in this manuscript will introduce and complement the other outputs of this project (e.g., source code).