Analyzing Spatial User Behavior in Computer Games using Geographic Information Systems

Anders Drachen
IT University of Copenhagen
Rued Langgaards Vej 7
2300 Copenhagen S, Denmark
+0045 72185038
drachen@itu.dk

Alessandro Canossa
IO Interactive/Danish Design School
Kalvebod Brygge 4
1354 Copenhagen K, Denmark
+0045 30233284
alessandroc@ioi.dk

ABSTRACT
An important aspect of the production of digital games is user-oriented testing. A central problem facing practitioners is however the increasing complexity of user-game interaction in modern games, which places challenges on the evaluation of interaction using traditional user-oriented approaches. Gameplay metrics are instrumentation data which detail user behavior within the virtual environment of digital games, forming accurate and detailed datasets about user behavior that supplement existing user-testing methods such as playtesting and usability testing. In this paper existing work on gameplay metrics is reviewed, and spatial analysis of gameplay metrics introduced as a new approach in the toolbox of user-experience testing and –research. Furthermore, Geographic Information Systems (GIS) are introduced as a tool for performing spatial analysis. A case study is presented with *Tomb Raider: Underworld*, showcasing the merger of GIS with gameplay metrics analysis and its application to game testing and –design.

Keywords: game development, user behavior, metric, gameplay metric, geographic information system

1. INTRODUCTION
User-oriented game testing [12] during production and post-launch has been performed for decades and forms a key aspect of game production [e.g. 16,17,23,25,29-31]. Unlike technical testing (e.g. bug hunting), user-oriented game testing has traditionally been performed using informal methods such as surveys and interviews [23]. Within the last decade, and notably the past few years, a variety of new structured methodologies have been adapted from Human-Computer Interaction research (HCI) and applied to user-oriented game testing or research [16,18], e.g. different forms of playtesting and usability-testing [7,23,26,27].

The methods currently employed for user research in the game industry – e.g. playtesting, usability testing, ethnographic methods and surveys - have different strengths and weaknesses [e.g. 21,26]. Methods such as these are useful for capturing player feedback and subjective experience, as well as for acquiring in-depth information on problems regarding gameplay or design. However, they are limited in that test managers can only hand-code so much information; and analysis of e.g. screen capture to a high level of detail is time consuming and not a good solution for the quick and effective game testing process typically required in the industry. A relatively newly developed potential solution to these shortcomings is presented in the automated collection and analysis of **gameplay metrics** data, i.e. instrumentation data about the user-game interaction, which supplement other user-oriented methods by providing very detailed quantitative data on the player behavior. Because these data are not recorded manually, they reduce error and save time.

Gameplay metrics form objective data on the interaction between players and games, and potentially any action the player takes while playing can be measured, including which buttons that are pressed, the movement of player-characters within the game world, or which weapons that are used to eliminate opponents. As an analysis tool, gameplay metrics supplement existing methods of games-based user research, e.g. *usability testing* (measuring ease of operation of the game) and *playability testing* (exploring if players have a good experience playing the game) by offering insights into how people are actually playing the games under examination.

Despite the relation to instrumentation data in software development, gameplay metrics have only in the past few years begun to see use in game production. There are therefore relative few methods developed for integrating gameplay metrics in game testing and game design/development, and little knowledge as to what metrics that should be tracked, when and how this varies across games [18,33]. Part of the reason is that techniques for the analysis of instrumentation data as they are applied in general HCI do not map directly to the field of game development and –research. First of all, the purpose of games is not usability, but entertainment [26]. Secondly, games take place in virtual environments (whether 2D or 3D) that simulates fictional worlds to a greater or lesser degree. This provides a very different interface and user-actionable environment than in productivity applications. Thirdly, game production can vary substantially from traditional software development processes with an emphasis on rapid iterative development, agile methodologies and needs for extremely rapid testing sessions involving human participants [23,26]. Jointly, these factors mean that the use of gameplay metrics remains largely unexplored, the literature and existing knowledge in the area minimal, and existing techniques rely on basic statistical analysis methods such as bar charts of a single variable showing e.g. level completion times [18,33]. It

MindTrek 2009, September 30th-October 2nd 2009, Tampere, FINLAND.

should be mentioned, that game metrics outside of the gameplay area, notably in terms of monitoring the economies of Massively Multiplayer Online Games (MMOGs) such as *World of Warcraft* and *Age of Conan,* or tracking regional sales for marketing purposes, form a stable part of the analysis work done by game companies. However, these metrics are not directly related to gameplay. Importantly, the analysis of gameplay metrics in relation to the spatial behavior of the player within the actual gaming environment (often 3D virtual worlds), remains largely unexplored. This is however an area of key interest to all games because the experience of the user (player) is directly related to the experience of navigating through, and interacting with, the game world [33]. Currently, the behavior of the user and the resulting interaction experience is primarily measured using usability and playability testing methods, and as mentioned above, these are not well suited towards providing highly detailed accounts of the user behavior, but more focused on evaluating the usability and experience of interacting with the game software [26,27]. The recording of gameplay metrics operating in the spatial environment of game – **spatial gameplay metrics** - worlds can however provide detailed information about the behavior of the person playing the game. For example, logging the position of the player within the game every second, and recording every action taken by the player, e.g. firing a weapon, picking up objects and interacting with other players.

In summary, what is needed to advance the use of spatial gameplay metrics – indeed any use of gameplay metrics – is flexible tools for analysis that are useful across game productions and game genres, that work both in an industry-context and research-context, and the development of specific analysis methods for different kinds of game metrics.

In this paper, a step is taken in this direction. The purpose is twofold: 1) To review the existing knowledge on gameplay metrics in research and development [e.g. 18,31], focusing on how gameplay metrics supplement existing user-oriented research- and testing methods in game production; 2) To provide a case study of spatial analysis of gameplay metrics from the major commercial game title *Tomb Raider: Underworld*, showing the potential and benefits towards both game development and research studies of player behavior and user experience; 3) To showcase **Geographic Information Systems** (GIS) [9,22] as a tool for performing spatial analysis on gameplay metrics, adaptable to any game context, informing user-oriented game testing and –design. A GIS is a computerized data management system used to capture, store, manage, retrieve, analyze, query, interpret and display spatial information in the form of e.g. maps, reports and charts [9,22]. Neither GIS nor spatial analysis of gameplay metrics have formed the subject of previous publications and form the key innovations of this paper.

2. STATE-OF-THE-ART

In the context of game development, the term “metric” denotes a standard unit of measure, e.g. a second or an hour [1,12,21] (in general HCI literature the term is also used to define a calculated number post-analysis, e.g. total number of errors made by users during testing). Metrics are generally organized in systems of measurement, utilized for quantitatively measuring and evaluating processes, events etc. Systems of metrics are generally designed to a specific subject area. Within game development, metrics form measures of engine performance, sales, project progress or user interaction with the game software, the latter category being of interest here. Gameplay metrics can relate to all forms of actions performed by the player in-game, including movement and behavior in a virtual environment, use of character skills and abilities, interaction with objects and other players, etc. In general, gameplay metrical data analysis is useful to compare the intent of the designers with the actual behavior of the players and to assist developers with quantifying their vision into elements that can be measured.

Gameplay metrics can take different forms, from logging of keystrokes to recording specific types of player behavior, e.g. firing a weapon, completing a level etc. Some metrics will be particular to a specific game (e.g. kill methods in *Hitman: Blood Money*), others relevant to an entire game genre (e.g. tracking PC movement as a function of time in FPS´). Some gameplay metrics can be recorded on a continual basis, e.g. movement in the virtual world, or be recorded using specific frequencies, such as the location of the player avatar/-s (virtual representation/-s of the player) every three seconds. Metrics can also register triggered events, e.g. every time an avatar jumps or shoots a weapon, finalizes a quest or completes a level.

Gameplay metrics are inherently objective and quantitative data, which can be collected in large numbers. Metrics allow for incredible detail – data showing for example where a player was positioned in the game environment while firing a specific weapon, with the camera angled in a certain way, and the result of the attack. The level of detail depends on the metrics system in question. Generally, these are custom pieces of software developed in-house, and map data to specific events or points in a gameplaying session. It is common to aggregate the data for easy visualization or for creating overviews or e.g. economies of massively multiplayer online games (MMOGs). In comparison, player-based feedback has much less resolution and is inherently biased due to individual preferences. Gameplay metrics are unbiased, and there is a clear practical outcome to metrics analysis. E.g., in balancing the time players spend performing different in-game actions [10]. Gameplay metrics can be analyzed in different ways. The traditional approach is purely statistical, based on e.g. aggregate counts of variables of user behavior used in time-spent reports, completion times or similar [e.g. 10,18,23].

Gameplay metrics are related to User-Initiated Events (UIEs), which refer to events that occur when a user directly interacts with the game software [18]. UIEs deal specifically with events initiated by the user; however, gameplay metrics also include events taken by the game software, for example the behavior of computer-controlled agents. Using automated tracking of UIEs to better understand user behavior originates within the HCI field, where they have been used for over two decades [e.g. 13,15,28,35]. Logging for example the movement of players in the virtual environment enables analysts to define which sections of the game world that are experienced and which that are not. The major challenge with gameplay metrics analysis is choosing which variables that should be tracked – computer games can form complex systems and the number of potential gameplay metrics is considerable. Furthermore, gameplay metrics datasets are usually very large, and the process transforming, ordering, cleaning, analyzing and visualizing the data can be challenging. Despite these challenges, there are several examples of the successful use of gameplay metrics in production [10,18,24,32]. Within the past few years, gameplay metrics have gained increasing attention as a means for obtaining detailed records of the player-game interaction, being used e.g. by *Microsoft, Bioware, Nintendo, EA*, and other companies [e.g. 23,31].

It is necessary to differentiate between game metrics and game heuristics [11,19]. Heuristics are design principles upon which games can be build. Game metrics are instrumentation data derived from game engines about e.g. the user´s hardware configuration or interaction with the game or game interface.

Gameplay metrics analysis allows user-research professionals to examine player behavior at multiple levels of resolution [18,24,26]. A second core strength of gameplay metrics analysis is the ability to recreate the playing experience of the player in detail. Being able to model the navigation of players through a game environment is of interest to game development for a number of reasons, not the least because it allows designers to observe how their games are being played. Traditional methods of user-testing computer games can locate problems with e.g. gameplay with a fair degree of precision, e.g. reporting that a specific encounter is too difficult. When integrating gameplay metrics, the second-by-second behavior of the players can be modeled, enabling more detailed explanations for the observed behaviors.

A distinct advantage of game metrics is that these can be collected off-site, i.e. from client installs -players who have installed the game on their PC/console. Metrics data from client installs are typically referred to as live data within the industry, and the approach is a regular feature of modern online multi-player games, where the ability to provide basic statistics of game sessions is generally appreciated by the players (e.g. level completion time averages). The two most substantial advantages of collecting live data are: 1) The data are unbiased by environmental effects. This is a common problem of laboratory-based research, which remove players from their normal playing environments thus potentially impacting on the way they behave and interact with the games being tested [21,34]; 2) Long-term monitoring in natural environments. Gameplay metrics can be streamed from live users over weeks or months, providing a long-term perspective on trends of use. This form of data collection is also more detailed than what are possible using personal logging methods [21], and is not subjected to the same errors that personal logging over extended periods is.

2.1 Challenges of gameplay metrics analysis

As with any other user-research method, the collection and analysis of gameplay metrics is not without its challenges. First of all, an infrastructure is needed to track and capture the data, and this includes substantial storage for large, commercial productions. Secondly, finding the right metrics to track in order to answer the pertinent questions can be challenging. Thirdly, gameplay metrics inform what players are doing, not always why. Gameplay metrics provide information only regarding actions undertaken in-game by players, it is usually not possible to assess reasons and motivations behind the action, unless additional user data are captured [18]: Gameplay metrics do not inform whether the player is male or female, or what the player thinks of the game experience. In short, gameplay metrics cannot provide any contextual data. For example, whether the player is having fun or not, is male or female or whether another player is watching the screen – a metrics tracking tool can only record information from the specific game software. When an analysis of a set of metrics data point to a specific player behavior, it is often necessary to combine the analysis with traditional game user-research methods such as playtesting, video capture, usability testing or similar.

2.2 Spatial analysis in virtual environments

It is useful to separate between two overall types of gameplay metrics: **Non-spatial gameplay metrics** are data that do not contain any spatial information, e.g. recording the number of time a player interacts with an NPC but not where. **Spatial gameplay metrics** are data that come with some sort of spatial reference information attached, e.g. a specific set of X,Y,Z-coordinates, a subsector of a map or similar. For example, the coordinates where a player has interacted with NPCs in a test session of *Neverwinter Nights 2*. It is entirely possible to analyze spatial metrics without using the associated spatial information. Analysis of spatial metrics requires software that is capable of mapping the data onto maps of the game world, zone or level.

Spatial gameplay metrics are a known resource within game development and –testing, however only utilized to a limited degree due to a general lack of knowledge about how to do so. Therefore, spatial gameplay metrics are usually only used to produce visualizations of single variables, and by far the most well-known example of such visualizations is the **heatmap** [24]. Heatmaps are density/location-based aggregated visualization of – traditionally - the kill locations of players, combat units or similar. The most commonly known examples are based on FPS´s such as *Unreal Tournament* and *Half-Life 2*. However, in principle density-based visualizations can represent any gameplay metric that can be mapped to a specific coordinate set (X,Y,Z) on a map of the game world/level. For example, the locations where players activate a specific player-character power or skill, fires a weapon, interacts with an object etc. In-house developed custom tools are normally used for creating heatmaps. Heatmaps represent the first step of working with gameplay metrics, namely the **visualization** of the data, and generally only a single variable at a time. Visualizing data is one thing, however, it is another to actually perform analysis in the spatial domain, and this is only rarely done. One of the few published examples of such work is [14], who developed an application (*Lithium*) for visualizing some basic spatial information from *Return to Castle Wolfenstein: Enemy Territory*. However, the application developed is largely focused on visualization rather than analysis, and is restricted to the game it was built for. Furthermore, it can only handle the mapping of variables that have been coded into the program. It is therefore not flexible across games or in terms of adding new variables, and cannot adapt to changing demands without the addition of additional code and specialized knowledge. Additionally, it does not permit analysis across layers of variables.

A second example is Börner & Penumarthy [3] who developed an approach towards visualizing the evolution of virtual communities in multi-user online VEs. They utilized the education-oriented *ActiveWorlds* platform, recording spatial and temporally referenced user interactions, e.g. navigation, object manipulation as well as chatting between users. The logged information was used to visualize e.g. navigation data on top of 2D maps of the VE. Chat data are visualized as 3D data hills, with peaks indicating locations with a high amount of chatting.

A third example stems from *Microsoft Game User Research Group*, who have developed a setup referred to as *TRUE* [18], which permits the mapping of basic point-based information and linking this to attitudinal data collected from surveys during gameplay. The ability to perform actual spatial analysis is however limited – at least, such a capacity has not been described in any of the publications or presentations from the group, which have been the forerunners of integrating gameplay metrics directly into the user-oriented research and testing of game development. Looking outside the games domain, there have in recent years been a number of attempts aimed at visualizing web navigation and interaction between users and software for purposes such as improving usability of the specific applications [e.g. 2,4]. These mainly focus on websites, however a few projects have considered virtual environments (VEs), and the focus is on visualizing information rather than analysis. Chittaro & Ieronutti [5] & Chittaro et al. [6] focus on using post-visit data, i.e. logged records of user movement in VEs. They developed a tool (VU-Flow) engineered towards tracking the movement and orientation of agents operating in VEs, and developed a series of visualizations and basic analyses – e.g. paths follows and areas of maximal and minimal user visitation.

In summary, the work with instrumentation data outside of the games field forms sources of inspiration for the work with games-specific data, however, the amount of development of metrics-based methods within the games area itself remains limited and the applicability in the context of game production and broader research is limited: Hoobler et al. [14] and Börner & Penumarthy [3] focus on specific online multiuser VEs. The strongest feature is the visualization and limited analysis of group behaviors and the handling of application-specific events, during runtime. They are inherently inflexible and not portable across environments. Similarly, while the VU-Flow tool of Chittaro et al. [6] is cross-VE compatible, it remain a specialized tool designed for the research purposes of navigational data only.

In a game production context, testing and data analysis needs to be carried out effectively, and therefore using a series of custom applications is not feasible. A GIS offers the advantage of being able to encompass virtually any spatial analysis – and if an analysis method is not already present in the myriad GIS systems or additional analysis modules available, it can be programmed in-house or via consultants using the built-in scripting languages. GIS offers a unified framework for spatial analysis, whether in 2D or 3D. Unlike most academia-developed systems, which are too inflexible for general analytic/testing work in the industry (or academia), GIS have been developed within a range of industries and is therefore engineered towards being flexible towards different uses. This makes it adaptable to spatial analysis in games research and –production. Importantly, GIS software is designed with the purpose of providing flexible exporting of visualizations and analysis results to a variety of stakeholders, e.g. marketing, design and management, which is a core requirement for adoption by the game industry.

2.3 Geographic Information Systems

In a GIS, a map links map features with attribute information. When working with a GIS, spatial gameplay metrics are generally referred to as **geospatial metrics**. This signifies that the data have both a **spatial** (coordinates or topological relationship signifying location) and a **thematic** component (the variables or attributes that are under study) that is being visualized/analyzed.

This correlates with the properties of spatial gameplay metrics, which contain spatial information as well as thematic (see above). The spatial information provides the type of feature to be mapped (for example a point, line or area), and the thematic information the attributes of the metric (for example, that a player character was killed at the given coordinate set). In a GIS, gameplay metrics can be plotted on e.g. level maps, and the data analyzed and visualized. Plotting for example the progression of one or more playtesters in the form of points registering the location (X,Y,Z) of the player every five seconds, with a color grading provides an instant visualization of progression which can subsequently be analyzed to evaluate whether the playtester is moving through the map as intended and if any problems are encountered and where [see also 32]. In this example, the map is treated as a traditional paper map, i.e. it is not interactive. In a GIS, it is possible to tag map features with detailed information in the form of attributes: For example, the level map could be divided into sub-sections and each section be named individually and contain various attributes such as degree of lighting, number of enemies, expected completion time etc. When mapping gameplay metrics on top of maps that themselves contain detailed feature information, a high degree of flexibility in the gameplay metrics analysis is gained, for example in calculating the number of kills occurring in "inside" type environments vs. "outside" type environments in an FPS – as well as specific numbers for each map feature. Conceptually, a GIS is reminiscent of the architecture of computer game systems that feature virtual environments linked with objects and entities whose attributes are stored in databases [9,22]. The difference is that a GIS is specifically engineered towards analysis of the featured data and their underlying attributes in a spatial environment (although analysis can also happen outside of the mapping environment, for example in building models or querying the underlying attribute databases). In practice, gameplay metrics are added in the form of layers on top of a level or zone map (Figure 1). GIS allows the placement of several layers and performing calculations along and across the associated multiple variables and their attributes.

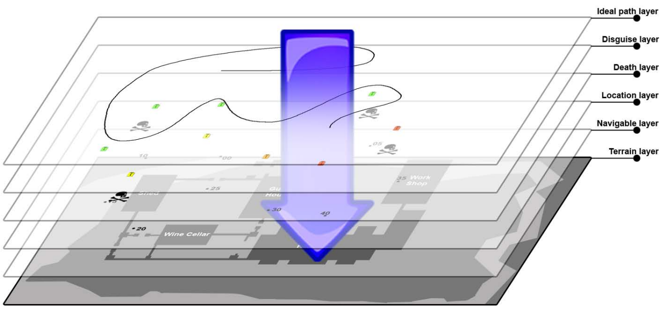

Figure 1: A GIS represent different data sets as layers on top of a game level map

2.4 The EIDOS Metrics Suite

The data for the case studies presented here are derived from the game *Tomb Raider: Underworld* (*Crystal Dynamics*, 2008). The data have been logged via the EIDOS Metrics Suite (**Figure 2**), an instrumentation system is designed to be able to interact with the existing user-oriented research and testing at the *EIDOS* studios and development houses; during production as well as in the live period, in addition to providing gameplay metrics data. It is constructed so as to facilitate the collection of data from any *EIDOS*-produced game, and to deliver these data directly to a variety of analysis software, e.g. the native GIS. The user research conducted at within EIDOS, e.g. at *IO Interactive* usually involves an array of methods, recordings and analyses [e.g. 16,25,33]. These varied data input are generally utilized together (e.g. attitudinal data, screen capture and gameplay metrics), as the various user research testing methods come with inherent strengths and weaknesses, requiring method triangulation to produce viable results. The *EIDOS* metrics suite has several features in common with instrumentation techniques used in the general HCI-field, e.g. that users interact with a system that automatically record application events of interest in message streams, which are send to a central server for analysis. The metrics suite was developed by and is maintained by the *EIDOS Online Development Team*. The team is responsible for maintaining the metrics suite and assist *EIDOS* developers in enabling and tracking metrics.

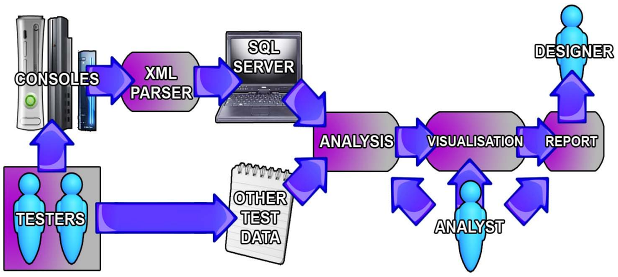

Figure 2: General framework for the gameplay metrics capture and analysis process at *IO Interactive*. Data are captured from testers (users) to a central *EIDOS* SQL server, from where they are drawn into different software analysis packages such as SPSS or ArcGIS for cleaning, evaluation and analysis. The data are then exported to visualization software), and a report developed for the target user (e.g. designer, QA).

3. TOMB RAIDER: A CASE STUDY

To showcase the potential of spatial analysis of gameplay metrics, three case studies are presented which also serve to highlight the functionality of GIS. At *IO Interactive*, the GIS software utilized is *ArcGIS* (produced by *Esri Corp*.). *ArcGIS* forms one of a handful of analysis packages utilized in conjunction with the game metrics collected from *EIDOS* games thanks to the work of the *EIDOS Online Development Team*. These include in-house tools developed for specific purposes, such as the easy visualization of key gameplay metrics such as kill locations and player movement trails. The overall research problem addressed by this paper is to evaluate the usefulness of spatial gameplay metrics analysis. The chosen case study. While the case studies are limited to the game *Tomb Raider: Underworld* [TRU], it focuses on analyzing gameplay features that are common to not only First-Person and Third-Person shooters but also Adventure Games as well as Massively Multi-Player Online Games (MMOGs) and similar genres where the player interacts with the game world via a central character.

The gameplay metrics data utilized are live data, i.e. data from players in their natural environment, captured using the EIDOS Metrics Suite. Due to the confidential nature of metrics data from commercial game titles, some facts about the underlying dataset cannot be revealed. For example, the precise dates that the data were generated, and the absolute numbers represented. The dataset consists of logs from roughly 28,000 players playing the game once in the fall 2008 after the game was released. One of key gameplay aspects to be tested via user-oriented methods in First-Person or Third-Person perspective games, where the players control a single character (or avatar), considers the level of challenge – are there any areas where the level of challenge is too high? Too low? Etc. one way to get an initial grasp of this key research question is to consider the locations and causes of player death in a game – in essence, areas where players die consistently and repeatedly are potentially imbalanced in terms of the challenge posed by the areas. This kind of design/research problem can be targeted from a non-spatial and a spatial angle. Using non-spatial gameplay metrics, analyses can only be performed on the causes of death (e.g. enemies, falling) and their relative frequencies (Figure 3). Such an analysis however does not inform whether there are specific areas that pose problems. Via spatial analysis it is possible to pinpoint exactly where problematic areas of the game world are located.

TRU is a 3D game focusing on the player controlling Lara Croft, an internationally rather well-known game character who forms a cross between an action heroine and Indiana Jones. The game is played in Third-Person perspective with a flexible camera system and forms an advanced platform game, where the player has to apply 3D-movement and jumping to solve a series of puzzles and navigate exotic environments such as Thailand and Jan Van Mayan Island. Apart from the ever-present risk of falling from the precarious heights that Lara Croft needs to navigate, the player regularly encounters different kinds of enemies, such as mercenaries or animals. Finally, the environment itself can pose a hazard to the player, such as fire.

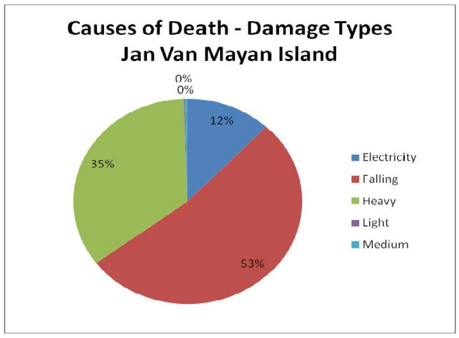

Figure 3: A pie chart showing the distribution of causes of death in the form of specific damage types, in the Jan Van Mayan Island level of *Tomb Raider: Underworld*. Live data from Xbox Live during a period in the fall 2008.

The game consists of seven main levels plus a prologue. Each game level is comprised of multiple "map units", for which the EIDOS Metrics Suite collects gameplay metrics data. One of these is the *Valaskjalf* map unit, occurring about two-thirds through the game. It is one of the more complex puzzle/trap locations in the game, featuring multiple different challenges to the players' skill. The map unit was subjected to a thorough **challenge analysis** during the early work with TRU gameplay metrics. In analyzing the patterns of death in the *Valaskjalf* map unit, the first step was to produce a heatmap based on locations of player death (X,Y,Z-coordinates) using *ArcGIS* (with the Spatial Analyst extension loaded (Figure 4). Heatmaps can be produced in different ways, e.g. using density functions or simply summing the number of deaths occurring within grid cells. The heatmap is excellent for informing about the lethality of different game areas. However, it is unspecific as to the nature of the deaths. In order to evaluate where different causes of death such as falling, different kinds of environmental dangers and computer-controlled enemies occurred (and if they occurred as intended by the game´s design!), a series of visualizations was produced using ArcGIS, showing the areas where players died of different causes (Figure 5).

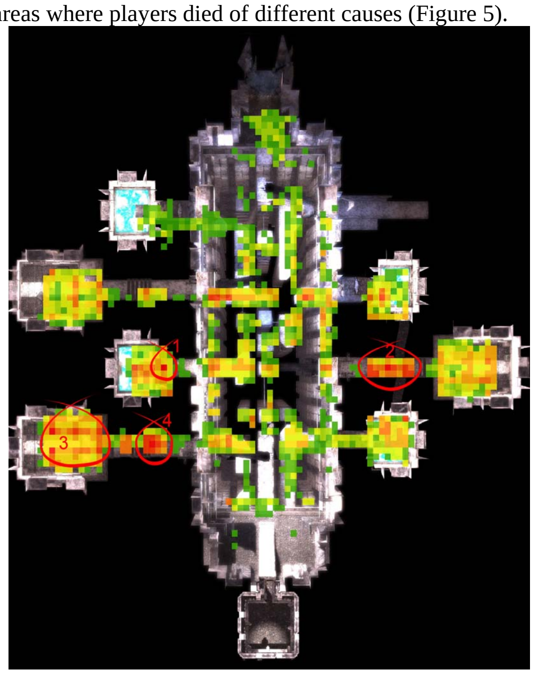

Figure 4: Grid-based heatmap of the locations of player death in the Valaskjalf map unit of *Tomb Raider: Underworld*. Scale ranges from light green (low numbers of death) to red (high numbers of death). Locations with no color have zero deaths. Dark red corresponds to 3050 deaths occurring within a single grid cell. Heatmap created in ArcGIS. Four of the most lethal areas are marked with red circles.

Both above examples form visualizations of gameplay metrics data rather than analysis. Apart from the ability to rapidly create visualization of key data, ArcGIS facilitates spatial analysis. In the current case, the question posed calls for an evaluation of which areas of the *Valskjalf* map that represent areas of not only high lethality, but also where multiple different causes of death occur. Such areas represent sites of high challenge to the players and therefore form targets for evaluation about whether their challenge level is too high. Answering this kind of question requires spatial analysis. In the current case overlaying several layers, each containing the distribution of one cause of death, and performing a count across these layers. A total of eight causes of death were included (including death by **falling**, via contact with **enemies** and **environment** causes, such as traps) (Figure 6). The result of the overlay analysis shows that most areas of the map unit have one to two causes of death. Four areas (Figure 4) have been marked for closer analysis (Figure 7). For Area 1, a high number of deaths occur in one specific grid cell (about 5*5 meters), being caused by a low variety of causes, namely the attack of a Thrall (an AI-enemy, third row in Figure 7) combined with a tricky jump (death by falling, fourth row in Figure 7).

If the number of deaths occurring in this area is deemed to high (i.e. prevents or diminishes player enjoyment), the analysis suggests two ways of solving the problem, e.g. making the jump easier or eliminating the Thrall enemy.

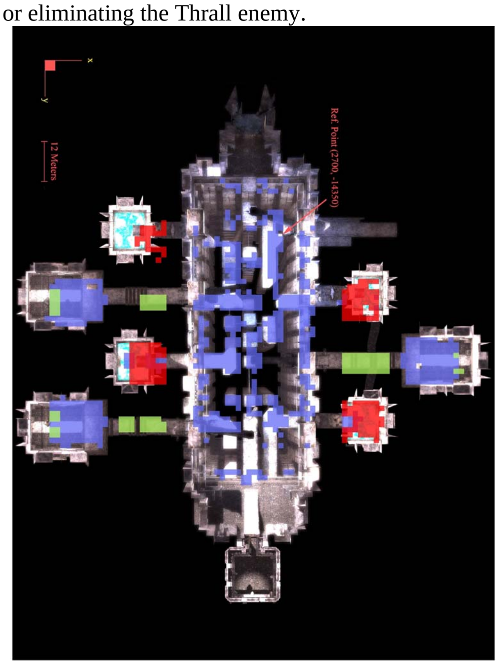

Figure 5: The Valaskjalf map unit level map has been overlain with three layers showing the extent of three separate causes of death: Falling (light blue), traps (green) and water volume [players drowning by being submerged in rising waters] (red).

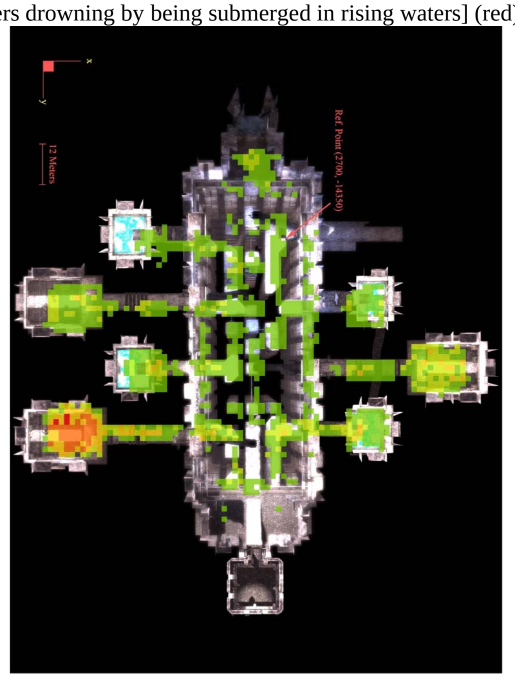

Figure 6: Overlay analysis using ArcGIS. The analysis shows the areas of the map where the highest number of different causes of death occur, on a scale from light green (one cause of death) to red (six causes of death). The area with the most causes of death is also one of the areas with the highest overall death count (Fig. 4).

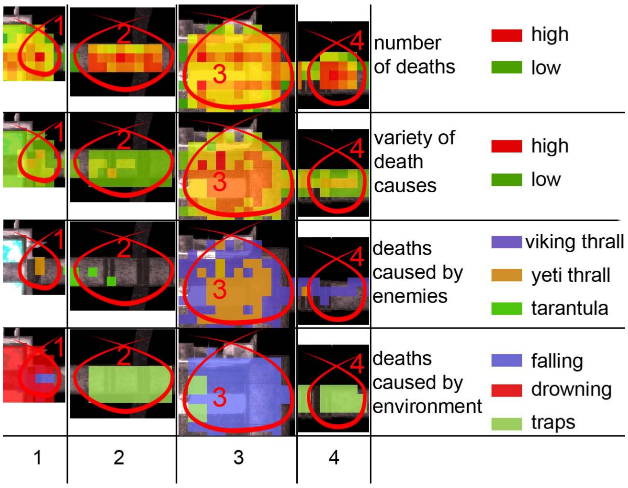

Figure 7: Detail of the overlay analysis with a breakdown of the four targeted areas with multiple causes of death. Four *ArcGIS*-derived layers included: Aggregated death count, aggregated causes of death (top two rows) and deaths specifically caused by enemies or environment effects (bottom two rows).

Area 2 (second column, Figure 7) also shows a high number of deaths and even though there are only two different causes, tarantulas (third row) and traps (fourth row), the distribution of tarantula kills on the *Valaskjalf* map is not spread enough to justify all the deaths displayed, meaning that most of the deaths are caused by the traps. This could suggest that the traps should be more lenient. The third area displays a high number of deaths, however it is motivated by a varied array of causes: Enemies, environment effects and falling – this is the climax of the level and clearly the toughest part to get through without dying. As with Area 1, a revision of the challenge level might be useful here. Area 4 displays very similar characteristics to Area 2 with similar implications in terms of the play experience.

The spatial analysis has thus identified potential trouble spots in the *Valaskjalf* map design, which subsequently can be analyzed in further detail, for example by comparing with user-satisfaction feedback from the level, to evaluate whether there is a problem or not. Additionally, this kind of spatial analysis provides valuable knowledge for future designs of Tomb Raider-style game levels.

ArcGIS permits different layers to be turned off and on flexibly, and even permits specific layers to be given different weights in the analysis – if e.g. players dying of electrocution is an unwanted occurrence in the game design, this can be given a greater weight and thus show up stronger in the analysis. Additionally, maps can be exported using the *ArcPublisher* extension as dynamic reports, which permit the user to add or remove layers dynamically, forming the perfect reporting tool for giving feedback to e.g. designers.

4. SIGNIFICANCE AND DISCUSSION

One of the major challenges to user research and –testing in commercial game development is the increasing complexity of interaction between players and game software, due in part to the sheer variety of interaction options. With the uptake in complexity, instrumentation data such as gameplay metrics form an increasingly more important tool for analysis of player-game interaction. The analysis of instrumentation data forms an important contribution to the user research and –testing performed during game development; as well as in the monitoring and continued evaluation of games during the live period. Furthermore, gameplay metrics form an excellent supplement to the existing user-oriented methods utilized within the game industry by providing detailed quantitative data about player behavior.

The *EIDOS* Metrics Suite, a system that permits the tracking, capturing, storage and reporting of game metrics, customized to the different games produced at the *EIDOS* studios; was developed by *IO Interactive* and is maintained by the *EIDOS Online Development Team*, as a response to the requirement of understanding the details of interaction between players and game, and to be able to monitor and analyze player behavior over extended periods of time. It is only recently that the game development industry has begun utilizing instrumentation data on a regular basis [18,31]. The use is often limited to aggregate counts of specific metrics – limited efforts have been directed at more complex analyses of user behavior (e.g. locating patterns of play [33]), as well as to the spatial behavior of player of computer games. The latter is crucial, because the vast majority of contemporary games are set in 3D virtual worlds. It therefore is imperative to be able to analyze and visualize gameplay metrics spatially. In essence, it is not enough to know that X% of player spent too long completing the level. In order to find out why, spatial analysis provides a venue for pinpointing exactly where players experience problems with progression. Visualizing and analyzing data in a Geographic Information System [9,22], which is specifically engineered towards handling data with a spatial component, provides a powerful tool and a plethora of new opportunities for user research experts in game development - as well as those working with other forms of virtual environments - to study how their users interact with and behave in these environments. Importantly, GIS´ are developed for multiple target industries and research fields and therefore flexible – they can handle a variety of needs as opposed to the systems developed by e.g. [3,5,6,14] which are focused on specific types of data/visualization.

While a modern GIS is perfectly capable of visualizing single variables in whatever format desired, the true strength and novelty of the approach lies in the capacity to perform analyses involving multiple variables. The case study presented here is based on the major commercial title *Tomb Raider: Underworld*, eight game in the *Tomb Raider*-series and one of the biggest franchises in the industry. While the case study is specific to TRU, it is focused on occurrences of player death, a phenomenon common in all games and the approach is directly applicable to any game or virtual environment (VE) where players control a single character: First a general look at areas of high death rate (e.g. via a traditional heatmap to locate areas of interest), followed by a closer examination of the variety of causes of death. This followed by evaluation of whether too many challenges are overlaid, providing assessment of the relatively playability of the areas. With analyses of this type, designers can make informed decisions as how to change the game design of any “trouble spots” in order to improve player’s experience. Death events is not the only occurrence that the approach can be applied to, other point-based variables can also form the basis for this type of analysis.

Spatial analysis of gameplay metrics has also been applied to Real-Time Strategy games, and forms the focus of future publications. The case study highlights the benefits of using spatial analysis of gameplay metrics data, not only as a testing tool but also in terms of user experience research – using gameplay metrics in combination with attitudinal data [18], it is possible to pinpoint exactly where players have good or bad experiences in the games, and which game features that cause them. Additionally, spatial analysis form sources of feedback to the players. Heatmap visualizations already form part of the community feedback from a variety of games such as *World in Conflict, Half-Life 2* and *Team Fortress 2*, however, actual spatial analysis provide the potential to expand on community feedback.

ACKNOWLEDGEMENTS

This would not have been possible without the stellar work of the EIDOS Online Development Team. Also sincere thanks to the many other colleagues at IO Interactive, EIDOS, the IT University of Copenhagen and the Danish Design School.

REFERENCES

1. Blythe, M. A., Overbeeke, K., Monk, A. F. Wright, P.C.: Funology – from usability to enjoyment. Springer (2004)
2. Brainard, J. Becker, B.: Case Study: E-Commerce Clickstream Visualization. In Proceedings of IEEE Symposium on Information Visualization (2001), 153-158
3. Börner, K., Penumarthy, S.: Social Diffusion Patterns in Three-Dimensional Virtual Worlds. Information Visualization, 2 (2003), 182-198
4. Cadez, I., Heckerman, D., Meek, C., Smyth, P., White, S.: Visualization of Navigation Patterns on a Web Site Using Model- Based Clustering. In Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (2000), 280-284
5. Chittaro, L. & Ieronutti, L.: A visual tool for tracing users´ behavior in virtual environments. In Proc. Working Conference on Advanced Visual Interfaces (2004), 40-47
6. Chittaro L., Ranon R., Ieronutti L.: VU-Flow: A Visualization Tool for Analyzing Navigation in Virtual Environments. IEEE Transactions on Visualization and Computer Graphics, 12,6 (2006), 1475-1485
7. Davis, J.; Steury, K., Pagulayan, R.: A survey method for assessing perceptions of a game: The consumer playtest in game design. Game Studies: The International Journal of Computer Game Research, 5 (2005)
8. De Kort, Y., Ijsselssteijn, W: People, places and play: Player experience in a socio-spatial context. ACM Computers in Entertainment 6, 2, article no. 18 (2007)
9. DeMers, M. N.: Fundamentals of Geographical Information Systems, Wiley & Sons (2008)
10. DeRosa, P.: Tracking Player Feedback to Improve Game Design. Gamasutra, Aug. 7 (2007)
11. Desurvire, H., Caplan, M., Toth, J. A.: Using heuristics to evaluate the playability of games. In CHI '04 extended abstracts on Human factors in computing systems (2004)
12. Dumas, J. S. 2003. User-Based Evaluations. In The HCI handbook. Lawrence Erlbaum Associates (2003), 1093-1117
13. Hilbert, D. M. and Redish, J. C.: A practical guide to usability testing. Intellect books (1999)
14. Hoobler, N., Humphreys, G., Agrawala, M.: Visualizing Competitive Behaviors in Multi-User Virtual Environments. In Proc. of the IEEE Visualization Conference (2004), 163-170
15. Hurst, A., Hudson, S. E., Mankoff, J.: Dynamic detection of novice vs. skilled use without a task model. In Proceedings of CHI 2007, ACM press (2007), 271-280
16. Isbister, K., Schaffer, N.: Game Usability: Advancing the Player Experience. Morgan Kaufman (2008)
17. Jørgensen, A. H.: Marrying HCI/Usability and Computer Games: A Preliminary Look. In Proceedings of NordiCHI, Tampere, Finland (2004)
18. Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B. C., Pagulayan, R. J., Wixon, D.: Tracking Real-Time User Experience (TRUE): A comprehensive instrumentation solution for complex systems. In Proceedings of CHI, Florence, Italy (2008), 443-451
19. Korhonen, H., Koivisto, E.M.I.: Playability Heuristics for Mobile Multi-Player Games. In Proceedings of DIMEA (2007)
20. Kort, J., Steen, M. G. D., de Poot, H., ter Hofte, H., Mulder, I.: Studying usage of complex applications. In Proceedings of Measuring Behavior. Noldus Information Technology (2005).
21. Kuniavsky, M.: Observing the User Experience: A Practitioner´s Guide to User Research. Morgan Kaufman (2003)
22. Longley, P.; Goodchild, M. F., Macquire, D. & Rhind, D.: Geographic Information Systems and Science. Wiley (2005)
23. Medlock, M.C., Wixon, D., Terrano, M., Romero, R.L., Fulton, B.: Using the RITE method to improve products: A definition and a case study. In Proceedings of UPA (2002)
24. Mellon, L.: Metrics in MMP Development and Operations. Presentation at the GDC, Los Angeles, USA (2004)
25. Nørgaard, M., Rau, J.: User Testing in the Combat Zone. In Proceedings of ACE, Toronto, Ontario (2007)
26. Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., Fuller, T.: User-centered design in games. In The HCI handbook. Lawrence Erlbaum Associates (2003), 883-906
27. Pagulayan R., Keeker, K.: Measuring Pleasure and Fun: Playtesting. In Wilson C. (Ed.): Handbook of Formal and Informal Interaction Design Methods. Morgan Kaufmann Publishers (2007)
28. Renaud, K. and Gray, P.: Making sense of low-level usage data to understand user activities. In Proceedings of SAICSIT 2004, ACM Press (2004), 115-124
29. Rouse, R. Game Design: Theory and Practice. Wordware publishing Inc. (2001)
30. Schultz, C. P.: Game Testing All in One. Course Technology PTR (2005)
31. Swain, C. 2008. Master Metrics: The Science Behind the Art of Game Design. Presentation at NLGD Conference, Utrecht, Holland (2008)
32. Thompson, C.: Halo 3: How Microsoft Labs Invented a New Science of Play. Wired Magazine 15, 9 (2007)
33. Tychsen, A., Canossa, A.: Defining Personas in Games Using Metrics. In proceedings of Future Play 2008, Toronto, Canada. ACM publishers (2008), 73-80
34. Tychsen, A., Newman, K., Brolund, T., Hitchens, M.: Cross-format analysis of the gaming experience in multi-player role playing games. In DIGRA 2007, Tokyo, Japan (2007), 49-58
35. Whiteside, J., Archer, N. P., Wixon, D., Good, M.: How Do People Really Use Text Editors? In Proceedings of´the SIGOA Conference, ACM Press (1982), 29-40
36. Williams, D.; Yee, N., Caplan, S. E.: Who plays, how much, and why? Debunking the stereotypical gamer profile. Journal of Computer-Mediated Communication 13 (2005), 993-1018