
**Herramienta de soporte para la evaluación de
videojuegos**
**Tool for supporting the evaluation of
videogames**



Trabajo de Fin de Grado
Curso 2022–2023

Autor

Jorge Bello Martín, Eva Lucas Leiro

Director

Guillermo Jiménez Díaz

Colaborador

Colaborador 1

Colaborador 2

Grado en Desarrollo de Videojuegos

Facultad de Informática

Universidad Complutense de Madrid

Herramienta de soporte para la evaluación de videojuegos

Tool for supporting the evaluation of videogames

Trabajo de Fin de Grado en **Desarrollo de Videojuegos**

Autor

Jorge Bello Martín, Eva Lucas Leiro

Director

Guillermo Jiménez Díaz

Colaborador

Colaborador 1

Colaborador 2

Convocatoria: *Junio 2023*

Grado en Desarrollo de Videojuegos

Facultad de Informática

Universidad Complutense de Madrid

22 de octubre de 2023

Dedicatoria

*A Pedro Pablo y Marco Antonio, por crear
TeXiS e iluminar nuestro camino*

Agradecimientos

A Guillermo, por el tiempo empleado en hacer estas plantillas. A Adrián, Enrique y Nacho, por sus comentarios para mejorar lo que hicimos. Y a Narciso, a quien no le ha hecho falta el Anillo Único para coordinarnos a todos.

Resumen

Herramienta de soporte para la evaluación de videojuegos

Un resumen en castellano de media página, incluyendo el título en castellano. A continuación, se escribirá una lista de no más de 10 palabras clave.

Palabras clave

Máximo 10 palabras clave separadas por comas

Abstract

Tool for suporting the evaluation of videogames

An abstract in English, half a page long, including the title in English. Below, a list with no more than 10 keywords.

Keywords

10 keywords max., separated by commas.

Índice

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	1
1.3. Plan de trabajo	1
2. Estado de la Cuestión	3
3. Descripción del Trabajo	5
4. Conclusiones y Trabajo Futuro	11
Introduction	13
Conclusions and Future Work	15
Contribuciones Personales	17
Bibliografía	19
A. Título del Apéndice A	21
B. Título del Apéndice B	23

Índice de figuras

Índice de tablas

Introducción

1.1. Motivación

1.2. Objetivos

El objetivo de la investigación es el desarrollo de una herramienta distribuida como un paquete para Unity que facilite a los desarrolladores la implantación de métodos de evaluación cuantitativos y cualitativos para la fase de producción y pre-producción de un videojuego. Nuestra herramienta pretende facilitar la realización de pruebas con usuarios para los desarrolladores que utilicen Unity y no dispongan de los medios o el conocimiento para realizarlas correctamente siguiendo buenas prácticas de GUR.

1.3. Plan de trabajo

Capítulo 2

Estado de la Cuestión

Capítulo 3

Descripción del Trabajo

Con el objetivo de permitir a un desarrollador de videojuegos integrar una prueba de GUR en su proyecto, la herramienta pondrá a disposición diferentes pruebas predefinidas entre las que podrá escoger a conveniencia, teniendo que preguntarse previamente algunas cuestiones: ¿Qué tipo de videojuego se está desarrollando? ¿Qué aspecto se pretende evaluar en el mismo? Una vez el desarrollador tiene claro qué tipo de prueba quiere realizar, comienza el trabajo de nuestra herramienta para poder implementarla en su videojuego.

En este apartado desarrollaremos cómo se utiliza la herramienta, qué aspectos se han tenido en cuenta a la hora de diseñar estas pruebas de usuario y qué resultados se pretenden conseguir. Para ello, utilizaremos un ejemplo concreto dentro del abanico de propuestas que disponemos en la herramienta: utilizar la herramienta para evaluar la dificultad de un videojuego o segmento del mismo donde existe una mecánica de penalización. El objetivo principal es averiguar cuál es el nivel de dificultad percibido por el usuario y cómo se ajusta al nivel de dificultad deseado por el desarrollador, partiendo de la certeza de que esta mecánica de penalización sirve como reflejo o consecuencia de los niveles de dificultad del videojuego, y por lo tanto, existe una relación directa.

En esta prueba de usuario partimos de una premisa necesaria que debe existir en el videojuego a evaluar: contiene una mecánica que se interpreta como una penalización al usuario. Definimos esta penalización como un evento del videojuego en el cual el jugador no consigue superar el obstáculo propuesto, reflejándose en una parada de la progresión (normalmente un reinicio), o, de manera más general, en una etiqueta que indicará que el jugador no ha superado el objetivo. En los videojuegos, se suele ver este evento como el clásico game over, pero se puede extender a cualquier tipo de castigo hacia el jugador: golpes recibidos en un videojuego de acción, resolución errónea en un juego de rompecabezas, etc. Pretendemos que sea lo suficientemente ambiguo como para que se pueda aplicar en muchos videojuegos, pero al mismo tiempo, lo suficientemente concreto para poder realizar una prueba de usuario más pormenorizada.

El interés en esta prueba de usuario se debe a que la mecánica de penalización está ligada directamente con las curvas de interés en el diseño de un videojuego. Se trata de una variable delicada, ya que la ausencia de veces en la que ocurre

puede llevar a que el jugador sienta menos desafío y por ende, menos interés; o por el contrario la repetición excesiva de la misma puede llevar a frustraciones. Según cómo sea el videojuego que se está desarrollando, puede existir más o menos interés en que se sucedan cierta cantidad de castigos en una partida, y por ello esta prueba empieza a ganar interés desde un punto de vista generalizado, en el que ofrezcamos las herramientas para observar ese número de penalizaciones (además de otros posibles datos relacionados con ella) y sacar conclusiones a raíz de estos datos recogidos.

La relación que existe entre la dificultad percibida y las penalizaciones es la siguiente: muy fácil pocas penalizaciones; muy difícil muchas penalizaciones. Debemos entender que lo que estamos evaluando es la dificultad, y el número de penalizaciones es una métrica en la que nos podemos apoyar. Esto quiere decir que el problema de la poca o excesiva dificultad se ve reflejado en las penalizaciones, pero no es generado por las mismas. Las penalizaciones son una consecuencia del nivel de dificultad, no un causante. Además, pueden existir otros factores que no dependan de estas penalizaciones que reflejen extremos no deseados en el nivel de dificultad.

Cabe destacar que la dificultad que se evaluará será la percepción general de la dificultad del videojuego o segmento a probar, no la dificultad de realizar una mecánica concreta o del diseño de niveles. Realizar una prueba para evaluar la dificultad general sin concretar en las particularidades del videojuego a tratar puede parecer contraproducente, pero a la hora de generalizar una implementación de esta prueba de usuario (con el objetivo principal de que sirva para el mayor número de desarrollos posibles) se debe abarcar desde un punto de vista más amplio, que no lo limite a ciertos tipos de videojuegos. Por esta razón, la definición que ofrecemos de una penalización o castigo acaba siendo la más acertada para poder empezar a realizar esta prueba de usuario. A partir de la realización de esta prueba, se podrán obtener ciertos resultados con los que poder realizar pruebas sucesivas intentando atacar de manera más concreta estos aspectos que puedan haber causado problemas específicos en cuanto a la dificultad.

En la prueba diseñada se proponen una serie de técnicas que combinarán métricas y cuestionarios que serán recogidos en diferentes momentos de la partida. Ambos apartados están diseñados para que en conjunto puedan ofrecer ciertas respuestas sobre las acciones del jugador que realiza la prueba. El hecho de que exista el cuestionario además de los datos objetivos y numéricos captados en el sistema de telemetría sobre las penalizaciones podrán hacer que el desarrollador entienda los diferentes perfiles de jugadores según cómo perciben la dificultad y cómo se han relacionado con la mecánica de penalización.

Comenzando por el apartado del cuestionario, dadas las características de la prueba definida, éste aparecerá una sola vez y será al final de la realización de la prueba. No necesitamos que el cuestionario se muestre durante la prueba ni que lo haga más de una vez, ya que consiste en un cuestionario acerca de la dificultad sobre todo el segmento a evaluar de manera general, no de secciones de la prueba en específico. El cuestionario nos proporcionará datos subjetivos numéricos acerca de la cuestión a tratar de una manera mucho más directa. Se trata de la forma más típica de recibir feedback directo de los probadores. Siguiendo diferentes recomendaciones en cuanto a cómo realizar las preguntas y los sesgos a evitar (libro GUR, 9.4), hemos

realizado el siguiente cuestionario:

![Esto sería una captura de pantalla del cuestionario desde Unity](https://prod-files-secure.s3.us-west-2.amazonaws.com/e73f2072-0f15-4bb3-a291-db8ac100dd4e/b8ce095a-0b47-4b5d-a4ea-1b0598ad20df/Untitled.png)

Esto sería una captura de pantalla del cuestionario desde Unity

Como podemos observar, encontramos en primer lugar unas preguntas relacionadas con el perfil del probador. Este tipo de preguntas nos sirve como desarrollador para poder encontrar una relación entre las cuestiones a tratar y los diferentes tipos de jugadores. Siempre será útil poder categorizar según edad y experiencia jugando a videojuegos tanto general como específica.

En segundo lugar, nos encontramos con una breve pregunta en la cual se deberá indicar según una escala la conformidad con afirmaciones relacionadas con los picos o caídas de dificultad, frustración y habilidad. Serán las respuestas más útiles para el desarrollador, puesto que pueden tener relación directa con los datos de telemetría que describiremos posteriormente.

Finalmente hemos implementado un cuadro de texto en el que el probador podrá hacer cualquier tipo de comentario que vea pertinente sobre la dificultad. Siendo una herramienta amplia, con la intención de que sirva para el mayor número de proyectos, este tipo de preguntas son muy útiles para abarcar cualquier dato específico a cada posible videojuego que utilice nuestra herramienta, sin necesidad de tener que implementar caso por caso diferentes tipos de preguntas o elementos relacionados con la telemetría que no pudiesen utilizarse de manera general.

En combinación al cuestionario, tendremos una serie de eventos que serán recogidos por el sistema de telemetría integrado en la herramienta. Los eventos serán los siguientes:

- Evento de inicio y fin de sesión
- Eventos de progresión: comienzo de partida, comienzo de nivel, pausa de una partida, reanudación de una partida, fin de nivel (SUCCESS, LOSE, MATCH, NONE), abandono de nivel.
- Evento de penalización
- Evento de posición de objetos

Estos eventos serán recogidos con sus marcas de tiempo y ordenados según ese valor. Algunos de ellos serán opcionales o necesitarán de la implementación concreta por cada desarrollador, ya que dependerán de las características de cada videojuego. Esto se explicará detalladamente más adelante.

Por último, una vez realizada esta prueba, se pretende mostrar un análisis de los resultados en los que se evaluarán las pruebas tanto individualmente como en su conjunto, con el objetivo de dar una respuesta a la pregunta inicial en este ejemplo. Los datos propuestos a mostrar en el análisis serían los siguientes:

****DATOS PARA EL CONJUNTO DE PRUEBAS:****

- Media de penalizaciones por partida
- Media de tiempo por partida
- Grafico de media de penalizaciones y de tiempo por cada una de las preguntas del cuestionario, además de con la experiencia.
- Grafica de tiempo vs penalizaciones

- Media de posición de objetos durante la sesión
- **DATOS PARA LAS PRUEBAS INDIVIDUALMENTE:****
- Número de penalizaciones del jugador en la partida
- Tiempo de partida
- Respuesta de los datos de su cuestionario.
- Posición de objeto durante la sesión.

Este tipo de relaciones entre datos y las respuestas de los cuestionarios son aplicables de forma automática y nos pueden permitir llegar a conclusiones de manera certera tanto como para las pruebas individuales como para el conjunto de todas las pruebas. Siguiendo la relación entre el número de penalizaciones y la dificultad previamente comentada, sumado a la directa percepción subjetiva de cada probador ante esta dificultad comparado con las métricas numéricas, llegar a conclusiones resultará mucho menos aventurado para los desarrolladores. Se proporciona un colchón de datos objetivos y subjetivos que se podrán relacionar siempre y cuando la propia prueba se haya cultivado en los marcos necesarios para que una se de por válida, con variables como el número de probadores general y el número de probadores según cada sub-grupo (en caso de querer sacar conclusiones según los sub-grupos) o la correcta realización de las pruebas dentro del tiempo máximo establecido.

Con este ejemplo descrito podemos ver un caso real de uso de la herramienta, pero aún debemos abordar el uso de la misma preguntándonos cómo podemos realizar correctamente estos casos y pruebas concretas y cómo se implementan en un proyecto.

*****COMO DISEÑADORES*****

Ahora explicar como esta diseñada nuestra herramienta como diseñadores sin entrar en detalles de la programacion, hacerlo desde el punto de vista del diseño de programas y usar diagramas o conceptos

Para ello, a continuación hablaremos sobre cómo está diseñada nuestra herramienta, desde el punto de vista del diseño y arquitectura de programas, aplicaciones y herramientas, qué hemos diseñado para cumplir los objetivos de la misma y cómo se pretende que se utilice.

Como se ha comentado anteriormente, nuestra investigación está dirigida para crear una herramienta que facilite a los desarrolladores la implantación de métodos de evaluación cuantitativos y cualitativos para la fase de producción y preproducción de un videojuego. Como tal, debemos cubrir no sólo los apartados de la implementación genérica y/o automática de un programa que contenga dichos métodos de evaluación, sino que debemos acercarnos a algo parecido a una herramienta de e-learning. Aunque no entremos tanto en detalle en aprender en profundidad sobre el GUR, sí que queremos conseguir que un desarrollador que no sepa nada sobre cómo realizar correctamente una prueba de usuario tenga una guía sobre qué hacer según lo que quiera investigar, y que pueda entender cómo está funcionando la herramienta. Debemos tener siempre en mente que esta herramienta está destinada a programadores y diseñadores de videojuegos, por lo que es necesario una buena comunicación usuario-herramienta.

Por lo tanto, a modo de acercamiento hacia esta visión más didáctica de cómo usar la herramienta y cómo funciona el GUR, la herramienta dispondrá de una

ventana de ayuda, en la cual existe una guía de uso de la misma: qué funciones puede realizar el desarrollador, cómo y cuando debe utilizarlas, guía de implementación en su proyecto y glosario acerca de los términos que se deban conocer previamente.

En segundo lugar, y como es lógico, la herramienta ofrecerá implementados los métodos de evaluación necesarios para poder realizar las pruebas. Aquí entrarán tanto los cuestionarios (métodos de evaluación cualitativos) y la telemetría (métodos de evaluación cuantitativos). Estos sistemas se explicarán con más detalle en el siguiente apartado, cuando los tratemos desde un punto de vista de programador. No obstante, volviendo a su diseño, los métodos de evaluación estarán implementados de manera independiente, es decir, se podrán combinar para formar las diferentes pruebas de usuario predefinidas.

Siguiendo la filosofía didáctica o de fácil uso/iniciación, dispondremos de diferentes pruebas de usuario predefinidas como la explicada en el apartado anterior. Estas pruebas han sido diseñadas previamente por nosotros, combinando cuestionarios y elementos del sistema de telemetría. Después de la investigación acerca del estado de la cuestión, conocemos ciertos patrones a seguir para obtener buenos resultados en pruebas de usuario que estén diseñadas para realizarse con equipos de desarrollo pequeños (o secciones del mismo dedicadas a las pruebas con usuario y QA). Por lo tanto, podemos ofrecer a los desarrolladores un conjunto de pruebas básicas comunes a la mayoría de videojuegos que puedan implementar en sus proyectos.

Para diseñar correctamente estas pruebas de usuario, seguimos diferentes metodologías. Estas metodologías van desde la correcta encapsulación de preguntas en un cuestionario, la duración de las pruebas o la relación entre los métodos cuantitativos con los cualitativos. Todas estas metodologías las hemos recogidos en una biblia de diseño en la que se define un workflow que seguir para crear una nueva prueba de usuario: definición del objetivo de la prueba y preguntas de investigación, definición del cuestionario, definición de qué recoger en el sistema de telemetría, y cómo implementar el análisis (WIP: relacionar con el apartado de trabajo futuro ->integrar la biblia de diseño para que los desarrolladores hagan sus propias pruebas de usuario). Una vez definida la prueba, se le pasa por un test de calidad en el que comparamos la prueba con otras pruebas predefinidas que sabemos que funcionan, o preguntas para saber si es una prueba de usuario útil para el objetivo de nuestra herramienta, que no podemos olvidar que es crear pruebas sencillas de entender y aplicar para la mayor cantidad de proyectos posibles.

Estas pruebas de usuario predefinidas se encuentran integradas en la propia herramienta, por lo que una vez que el desarrollador acceda a la ventana de ayuda y decida qué prueba utilizar, deberá integrarla en su proyecto siguiendo las indicaciones pertinentes.

Entrando en detalle en el apartado de los métodos cualitativos, hemos integrado los cuestionarios de manera modular, utilizando las herramientas del motor en el que integremos la herramienta para crear los diferentes tipos de preguntas-respuestas (más detalles en el apartado de programación). Estas preguntas estarán paginadas a conveniencia de la prueba (según el tamaño del cuestionario o diferentes secciones que se encuentren en el mismo). Es el conjunto de estas preguntas-respuestas lo que consideraremos como un cuestionario.

Los diferentes módulos que se han decidido utilizar para crear cuestionarios son

los siguientes:

- Respuestas de elección según escalas (ejemplo: Indica, siguiendo la siguiente escala 0 (nada conforme) - 5 (totalmente conforme), tu grado de conformidad con la afirmación x/y/z). Se agrupan un conjunto de afirmaciones en las cuales el usuario responderá utilizando la escala

- Cuadros de texto (numéricos y alfanuméricos)
- Cuadros de selección individual o múltiple

Según la prueba que se esté utilizando, tendrán más o menos peso en la balanza estos cuestionarios. El contrapeso se trata de los métodos cuantitativos integrados en el sistema de telemetría. Los métodos en cuestion son los eventos que se captarán mediante el sistema de telemetría.

añadir aquí una definición o relación hacia el apartado donde se explique (estado de la cuestión?) qué es el sistema de telemetría y los eventos. si no, hacia el glosario

P

COMO PROGRAMADORES

Aquí explicar cómo está programada la herramienta que explicamos justo antes metiendo las cosas necesarias sobre unity y c y la distribución mediante la unity store

Capítulo 4

Conclusiones y Trabajo Futuro

Introduction

Introduction to the subject area. This chapter contains the translation of Chapter 1.

Conclusions and Future Work

Conclusions and future lines of work. This chapter contains the translation of Chapter 4.

Contribuciones Personales

En caso de trabajos no unipersonales, cada participante indicará en la memoria su contribución al proyecto con una extensión de al menos dos páginas por cada uno de los participantes.

En caso de trabajo unipersonal, elimina esta página en el fichero `TFGTeXiS.tex` (comenta o borra la línea `\include{Capitulos/ContribucionesPersonales}`).

Estudiante 1

Al menos dos páginas con las contribuciones del estudiante 1.

Estudiante 2

Al menos dos páginas con las contribuciones del estudiante 2. En caso de que haya más estudiantes, copia y pega una de estas secciones.

Bibliografía

*Y así, del mucho leer y del poco dormir, se
le secó el cerebro de manera que vino a
perder el juicio.*

(modificar en Cascaras\bibliografia.tex)

Miguel de Cervantes Saavedra

Apndice **A**

Título del Apéndice A

Los apéndices son secciones al final del documento en las que se agrega texto con el objetivo de ampliar los contenidos del documento principal.

Apndice	B
---------	----------

Título del Apéndice B

Se pueden añadir los apéndices que se consideren oportunos.

Este texto se puede encontrar en el fichero Cascaras/fin.tex. Si deseas eliminarlo, basta con comentar la línea correspondiente al final del fichero TFGTeXiS.tex.

*–¿Qué te parece desto, Sancho? – Dijo Don Quijote –
Bien podrán los encantadores quitarme la ventura,
pero el esfuerzo y el ánimo, será imposible.*

*Segunda parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

*–Buena está – dijo Sancho –; fírmela vuestra merced.
–No es menester firmarla – dijo Don Quijote–,
sino solamente poner mi rúbrica.*

*Primera parte del Ingenioso Caballero
Don Quijote de la Mancha
Miguel de Cervantes*

