

Proyecto NeuroGame: Neuro-Evaluación Multimodal de Videojuegos

Antonio Loscertales¹, Eva Cerezo¹, Marco Creatura², José Luis González³, Javier Minguez¹,
Rosa Gil³, Sandra Baldassarri¹ and Yolanda López⁴

¹ Departamento de Informática e Ingeniería de Sistemas, Universidad de Zaragoza.
antonio.loscertales@gmail.com, {ecerezo, jminguez, sandra}@unizar.es

² BitBrain Technologies. CEEI, C/ María de Luna nº 11. Nave 4. 50018, Zaragoza.
marcocreatura@bitbrain.es

³ GRIHO, Universitat de Lleida.

joseluisgs@diei.udl.cat, rgil@griho.net

⁴ Departamento de Psicología y Sociología, Universidad de Zaragoza.
ylopez.iacs@gmail.com

ABSTRACT

Los videojuegos van tomando cada día un papel más importante en el sector del ocio y el entretenimiento, siendo su impacto económico cada vez mayor. Es por ello que la detección temprana, incluso en las fases de diseño, de la aceptación y el impacto del videojuego en los potenciales usuarios se ha vuelto una cuestión de suma importancia. Este trabajo propone una metodología para la evaluación hedónica de videojuegos basada en la medida en laboratorio de las respuestas emocionales y cognitivas de los usuarios. La metodología cubre no sólo la evaluación de la jugabilidad sino de los aspectos más relevantes del diseño que tengan impacto sobre la misma. La metodología se apoya en el uso de las tecnologías actualmente disponibles de detección emocional y cognitiva que permitan la evaluación multimodal a través de diversos canales (visuales, auditivos, fisiológicos). En el montaje experimental propuesto destaca el uso de los sensores fisiológicos que permitan la neuro-evaluación emocional de la experiencia de usuario. La metodología propuesta se aplica en el proyecto NeuroGame a la evaluación intensiva y detallada de tres videojuegos. Se presentan los resultados obtenidos de unas primeras pruebas con usuarios en laboratorio.

Categories and Subject Descriptors

H.1.2 [Models and Principles]: User/Machine Systems – *human factors, human information processing, software psychology.*

H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – *evaluation, methodology.*

General Terms

Measurement, Experimentation, Human Factors.

Keywords

Experiencia de Usuario, Evaluación emocional multimodal, Interfaces cerebro-computador, Videojuegos.

1. INTRODUCCIÓN

En el proceso de Interacción Persona-Ordenador, la evaluación de la Experiencia de Usuario (UX) se define como “la percepción de una persona y las respuestas subjetivas de ésta como resultado de la utilización y/o el uso de un producto, sistema o servicio” [ISO/IEC 9241-210, 2010. Clause 2.15], de forma que asume un rol fundamental en el diseño y desarrollo de productos interactivos. En este contexto, la UX y, en particular, la evaluación hedónica (¿qué sentimos al realizar la tarea?) completa las informaciones clásicas más orientadas a una evaluación pragmática (¿para qué sirve?). Este enfoque permite obtener una visión amplia direccionada a mejorar la aceptación del producto, la experiencia interactiva y el retorno económico en fase de desarrollo del producto [6] [11].

Es importante señalar que los estados emocionales usualmente asociados a la experiencia del usuario, influyen procesos cognitivos como la capacidad de atención y memorización, el rendimiento del usuario y, de esta forma, su valoración del producto [16] [2], teniendo así un rol fundamental en la fase de diseño de éste.

El presente artículo presenta NeuroGame, un proyecto multidisciplinar que tiene como objeto profundizar en la comprensión de la medición y el análisis de información asociada a la evaluación hedónica, en forma de respuestas emocionales y cognitivas, para obtener una perspectiva más detallada y precisa de la Experiencia de Usuario. En dicho proyecto participa el AffectiveLab de la Universidad de Zaragoza, el grupo GRIHO de la Universitat de Lleida, la empresa BitBrain Technologies y el Departamento de Psicología y Sociología de la Universidad de Zaragoza. El ámbito de aplicación del proyecto es el del ocio y entretenimiento y, más concretamente, los videojuegos.

El objetivo del artículo es doble. En primer lugar, se propone una metodología general de evaluación de videojuegos tanto en los aspectos de diseño como de jugabilidad, que se centra en la evaluación de la Experiencia de Usuario a través del análisis de las respuestas emocionales y cognitivas de los usuarios medidas en laboratorio. La metodología involucra una serie de tecnologías que permiten la evaluación a través de diversos canales. En segundo lugar, el artículo plantea, basándose en dicha metodología general, la evaluación hedónica de diversos aspectos de tres videojuegos concretos y presenta al final los resultados de las primeras pruebas llevadas a cabo con usuarios.

La estructura del artículo es la siguiente: en el apartado 2 se repasan los trabajos relacionados mientras que en el apartado 3 se detallan las tecnologías involucradas en la metodología presentada en el apartado 4. El apartado 5 detalla la aplicación de la metodología propuesta a la evaluación de la jugabilidad y el diseño de tres videojuegos, presentándose en el apartado 6 los resultados de las primeras pruebas llevadas a cabo. Por último, en la sección 7 se comentan las conclusiones y el trabajo futuro.

2. TRABAJOS RELACIONADOS

El contexto de los videojuegos y su relación con las emociones es una combinación que en los últimos años ha atraído relativa atención en forma de varios estudios. Concretamente, en el estudio de Ravaja et al. [18] se analizan los patrones de las respuestas emocionales de los usuarios ante varios videojuegos. Los resultados sugieren la utilidad del análisis de los patrones de las respuestas emocionales en el diseño de videojuegos. Otros estudios [14] proponen y estudian la eficacia de las mediciones fisiológicas como evaluación de la UX en el campo de la tecnología de entretenimiento. Los resultados muestran la relación entre las medidas fisiológicas registradas y los cuestionarios proporcionados a los participantes. Este tipo de investigaciones suponen un primer avance hacia la evaluación completa de la Experiencia del Usuario. Recientemente se ha propuesto y llevado a cabo [6] una evaluación de la Experiencia del Usuario respecto a la jugabilidad de un videojuego con técnicas de análisis hedónico. Las conclusiones muestran que factores como la respuesta emocional y la influencia sociocultural enriquecen los métodos tradicionales de evaluación, aconsejando su uso desde etapas más tempranas del desarrollo.

La definición, el modelado y la representación de las emociones sigue siendo objeto de estudio en el campo de la psicología y la computación afectiva [17]. Existen dos enfoques fundamentalmente [7]: el enfoque categórico o discreto y el enfoque dimensional. El enfoque categórico se basa en la clasificación de emociones en categorías [5], por lo que supone una representación discreta de las emociones incapaz de reflejar el amplio rango de emociones complejas que un ser humano es capaz de expresar. El segundo enfoque es dimensional, y establece que las emociones conforman un espacio continuo multidimensional en el que unas y otras no son independientes sino que se relacionan de forma sistemática, siendo el modelo tridimensional (valencia, activación y control) [19] uno de los más extendidos, por lo que ha sido el elegido en NeuroGame y será detallado más adelante.

En cuanto a las respuestas emocionales en sí, estas pueden darse a través de diversos canales. El canal tradicional de detección emocional han sido las imágenes faciales, pero en los últimos años [7] se han ido incorporado otros canales como el sonoro (voz) y, muy especialmente, los fisiológicos (ver Tabla 1). Esto es debido a las crecientes evidencias de que los estados emocionales están relacionados con cambios fisiológicos específicos. En esta línea, se han llevado a cabo estudios experimentales para analizar las respuestas emocionales frente a diferentes tipologías de estímulos: sonidos especiales, música, imágenes, vídeos de películas, videoclips, etc. La Tabla 2 muestra los trabajos más relevantes en esa línea.

A continuación se concreta el despliegue tecnológico propuesto por el proyecto NeuroGame.

3. TECNOLOGÍAS INVOLUCRADAS

En el proyecto NeuroGame se propone el uso de un despliegue sensorial (ver Tabla 1) que permita recoger información tanto

fisiológica como no fisiológica con el objeto de obtener una visión fiable y completa del estado emocional del usuario. Los sensores fisiológicos se han dividido en aquellos que miden directamente la actividad del Sistema Nervioso Central (SNC) y del Sistema Nervioso Periférico (SNP). En cuanto a los no fisiológicos se persigue básicamente el análisis mediante videocámara de la respuesta facial emocional, complementada por un *EyeTracker* y un micrófono.

Tabla 1. Sensores usados en la literatura para la detección del estado emocional del usuario (en negrita los involucrados en el proyecto NeuroGame).

Sensores Fisiológicos	No fisiológicos
Originados en el Sistema Nervioso Centralizado (SNC): Electroencefalograma (EEG)	Videocámaras (VCF + VCC)
Originados en el Sistema Nervioso Periférico (SNP): Electrocardiograma (ECG) Electromiograma (EMG) Volumen de pulso sanguíneo (BVP) Conductividad de la piel (SC) Respiración (RSP) Saturación de Oxígeno (SpO2) Temperatura de la piel (ST) Electrooculograma (EOG)	Micrófonos (M) EyeTracker (ET)

3.1 Sensores fisiológicos

Las señales fisiológicas de origen eléctrico aportan medidas directas e indirectas sobre el estado emocional y además, aquellas que registran directamente la actividad del SNC, pueden aportar medidas sobre funciones cognitivas. A continuación se describe la funcionalidad de cada uno de los involucrados en el proyecto:

- **Electroencefalograma (EEG)**: mide la actividad eléctrica del cerebro. Se realiza colocando un gorro con electrodos puestos en contacto con el cuero cabelludo. Los potenciales registrados son originados en el SNC.
- **Sensor de Volumen de Pulso Sanguíneo (BVP)**: mide de forma no invasiva los cambios relativos de volumen sanguíneo. Esta medición indica la cantidad de sangre que circula actualmente a través de las venas, lo cual permite calcular la vasoconstricción, la dilatación vascular y la frecuencia cardíaca.
- **Sensor de Conductividad de la Piel (SC)**: conocido también como sensor de la actividad eléctrica de la dermis, sirve para medir la capacidad de la piel para la transmisión de corrientes eléctricas, que varían cuando hay sudoración y cambios en el organismo.
- **Sensor de Respiración (RSP)**: mide los cambios de respiración. Con la señal de este sensor se puede calcular parámetros fisiológicos respiratorios como la frecuencia respiratoria, la amplitud relativa de la respiración y la profundidad y rapidez de la ésta.
- **Sensor de Saturación de Oxígeno (SpO2)**: permite monitorizar el nivel de oxigenación de la hemoglobina.
- **Sensor de Temperatura de la Piel (ST)**: sirve para medir la temperatura de la superficie de la piel.

Es importante señalar que no todos los sensores se usan en todas las fases de la evaluación. Esto es debido a que algunos de ellos

Tabla 2. Trabajos relevantes donde se estudian, mediante diferentes sensores de tipo fisiológico, las respuestas emocionales de los usuarios ante diferentes tipos de estímulos.

Ref.	Estímulos	Sensores	Resultados	Emociones involucradas
Maaoui et Pruski, 2010 [13]	Imágenes	SC, BVP, EMG, ST, RSP	Precisión del 92% en el reconoc. de las seis emociones.	Diversión, alegría, asco, miedo, tristeza, neutro.
Cutbert et al, 2000 [4]	Imágenes	EEG, EMG, ECG, SC	Cambios en parámetros durante estímulos.	Agrado, desagrado y neutro.
Wagner et al, 2005 [20]	Canciones	SC, EMG, ECG, RSP	Precisión en media del 80% en el reconoc. Emocional.	Alegría, ira, placer, tristeza.
Liu et al, 2010 [12]	Sonidos	EEG	Nuevo algoritmo para reconoc. y visualiz. de emociones en tiempo real.	Miedo, frustración, alegría, complacencia, tristeza, satisfacción.
Nasoz et al, 2003 [15]	Fragmentos de películas	SC, ST, ECG	Precisión en media del 76% en el reconoc. emocional.	Tristeza, ira, miedo, sorpresa, frustración, diversión.
Costa et al, 2006 [3]	Fragmentos de películas	EEG	Aumento del índice de sincron. durante estímulos emocionales.	Alegría, tristeza.

tienen un impacto invasivo sobre el sujeto de evaluación por lo que no es posible su uso en algunas fases (e.g. sensores colocados en los dedos no permiten al usuario jugar de forma natural).

3.2 Sensores no fisiológicos

Los sensores que miden actividad fisiológica se complementan con otros sensores, los cuales aportan información emocional y de atención visual. Los sensores asociados son:

- **Videocámara frontal (VCF):** pequeña videocámara situada sobre la pantalla, enfrente del usuario, con objeto de grabar las expresiones faciales del mismo. Trabaja a una resolución de 752x480 píxeles con una frecuencia de 60 FPS (fotogramas por segundo).
- **Videocámara de campo (VCC):** se utiliza para monitorizar toda la actividad durante el experimento. Su objetivo es conservar un registro de todo lo ocurrido durante la evaluación. Se empleará una resolución de 1280x720 píxeles y una frecuencia de 30 FPS.
- **Micrófono (M):** se utilizará el micrófono para registrar las palabras o sonidos que el sujeto de evaluación pueda realizar durante el experimento.
- **EyeTracker (ET):** sirve para evaluar el punto donde se fija la mirada. Aunque existen de diversas características, en el proyecto se dispone de una tecnología consistente en unas gafas con una cámara integrada en la parte superior del ojo derecho del usuario.

A diferencia de los sensores fisiológicos, las videocámaras necesitan para su correcta recogida de datos unas determinadas condiciones de iluminación, posición y frecuencia:

- Iluminación: debe alcanzarse un nivel suficiente para que el rostro del usuario sea capturado adecuadamente por la videocámara frontal (VCF) de forma que el seguimiento de las características faciales necesarias para la clasificación emocional sea factible [9].
- Posición: es muy importante, pues de ella depende la calidad de los datos que serán evaluados. De acuerdo con [10], se realizaron varias pruebas previamente con las videocámaras del experimento, y finalmente se concluyó que el sujeto

debería colocarse a 1.2 metros de la cámara frontal de modo que su cara quedara centrada horizontalmente y ligeramente por encima del eje horizontal que divide la imagen en dos.

- Frecuencia: En el experimento se ha utilizado una videocámara frontal especializada en tecnología para visión artificial y aplicaciones de visión por computador. Se utilizará aquella configuración que proporcione la mayor frecuencia posible para garantizar un correcto seguimiento facial.

4. METODOLOGÍA

A continuación se propone una metodología general de evaluación multimodal y emocional de videojuegos. En este apartado, se definirá el objeto y los sujetos de estudio, las dimensiones emocionales involucradas y el proceso de evaluación a seguir.

4.1 Objeto de estudio

Uno o varios videojuegos elegidos por sus similitudes en cuanto género, historia y protagonistas.

4.2 Sujetos de estudio

El tamaño muestral del experimento será de unos 20 sujetos balanceando el número de hombres/mujeres en la medida de lo posible. Se reclutarán personas atendiendo a los siguientes criterios de inclusión y exclusión:

Criterios de inclusión: 1.- Edad: 18-34 años. 2.- Ausencia de tratamiento farmacológico o psicológico actual o en el mes previo al estudio. 3.- Correcta comprensión de español hablado y escrito 4.- Aceptación de consentimiento informado de su participación en el estudio.

Criterios de exclusión: 1.- Trastorno psiquiátrico o físico que impidiese el correcto desarrollo del entrenamiento y/o evaluación. 2.- Imposibilidad de acudir o de realizar las sesiones de evaluación. 3.- No aceptación de consentimiento informado para su participación en el estudio. 4.- No haber jugado a ninguno de los videojuegos seleccionados para el experimento.

Para asegurar que el sujeto de evaluación cumple tanto los criterios de inclusión como los de exclusión, deberá rellenar un test antes de ser aceptado como sujeto del experimento. En el test se incluirán todas las preguntas necesarias para asegurar que el candidato cumple con el perfil necesario, además de aquellos

aspectos de origen sociocultural, útiles para la relación entre los resultados y el perfil de usuario, así como un consentimiento informado donde se explique claramente la naturaleza de las pruebas y el rol de los sujetos en ella, que deberá ser firmado por duplicado por el sujeto y un miembro del equipo evaluador.

4.3 Dimensiones emocionales

Se opta por un enfoque dimensional basado en el modelo de emociones de Russell [19]. En este modelo la variabilidad de una emoción se representa mediante tres dimensiones: activación, valencia y control o dominancia. La dimensión de activación se refiere al grado de excitación o apatía de la emoción y va desde emociones apáticas como el aburrimiento hasta emociones de máxima activación como la ira o la sorpresa. La dimensión de valencia se refiere a la forma negativa o positiva de la emoción, oscilando desde sentimientos desagradables hasta sentimientos agradables como la felicidad. Por último, la dimensión de control se refiere al grado de dominancia o control de la situación (i.e. si la situación supera al sujeto que está sintiendo la emoción o por el contrario es éste quien controla la situación).

Durante el experimento, se les pedirá a los usuarios que realicen pequeños test o valoraciones donde deberán indicar las emociones experimentadas. Para ello, se utilizará la escala SAM [1] (ver Figura 1). En ella, los usuarios deben etiquetarse en las tres dimensiones de interés (activación, valencia y control o dominancia) en una escala numerada del 1 (mínimo) al 9 (máximo). Además, la escala SAM viene acompañada de dibujos que van acordes a la valoración numérica, facilitando la comprensión y etiquetado en los sujetos de evaluación.

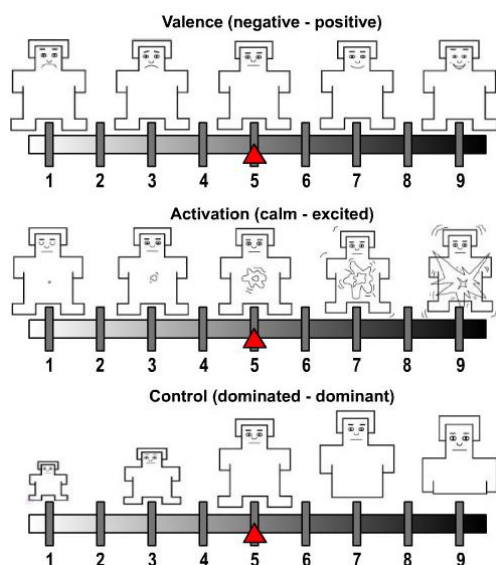


Figura 1. Escala SAM (Self-Assessment Manikin) a usar para la autoevaluación emocional.

4.4 Procedimiento

La evaluación se estructurará en una o varias sesiones. Todas las sesiones de evaluación constarán de dos partes bien diferenciadas: calibración y evaluación. En ambas se presentarán al usuario estímulos de una o varias modalidades (imágenes estáticas, videos, melodías,...). El objetivo de la calibración será estimular a la persona en las dos condiciones extremas dentro de cada una de las tres dimensiones emocionales (activación, valencia y control o dominancia). De este modo, se construirán las referencias para la evaluación posterior de la respuesta emocional del usuario. La

evaluación consistirá en la presentación de estímulos diana y de ruido (otros estímulos no relacionados con el tema de interés), obteniendo las medidas sensoriales de respuesta para su posterior análisis. En una evaluación se considerarán una o varias categorías del videojuego (por ejemplo, los personajes o la ambientación) que se considere interesante evaluar. Para cada una de las categorías se presentarán una serie de muestras. Una muestra se compondrá de un descanso, un estímulo (ruido o diana) y una valoración por parte del usuario. En la Tabla 3 se presenta un glosario de términos con objeto de ayudar a la comprensión del proceso.

Tabla 3. Glosario de términos relacionados con la evaluación

Término	Definición
Sesión	Unidad de tiempo de trabajo con el sujeto.
Modalidad	Tipo de estímulo (imagen, melodía, vídeo o interacción).
Calibración	Parte de la sesión dedicada a tomar medidas con el objeto de usarlas como estados de referencia. Una calibración está asociada a una modalidad y se descompone en muestras.
Evaluación	Parte de la sesión dedicada a tomar medidas con objeto de evaluarlas. La evaluación puede ser de una o varias categorías.
Categoría	Tipología de elementos de un videojuego a evaluar en una sesión. Una categoría está asociada a una modalidad y se descompone en muestras.
Muestra	Conjunto de acciones a realizar cuando se presenta un estímulo. Una muestra está siempre compuesta de descanso, estímulo y valoración.

4.5 Sesiones

En este apartado se detallan las partes principales de las que se compone una sesión. Como se ha mencionado anteriormente, se denomina sesión a una unidad de tiempo de trabajo con el sujeto. Debido a la duración que supone analizar todas aquellas categorías de interés en un experimento, es muy probable que éste deba partirse en varios días, realizando una sesión por cada uno de ellos. En la Figura 2 se presenta el esquema general de una sesión. A continuación se detallan cada una de las tres partes principales de una sesión: Configuración Inicial, Calibración y Evaluación.

4.5.1 Configuración Inicial

La configuración inicial es imprescindible al comienzo de cada sesión ya que en ella se realizará la puesta a punto de todos los sensores del experimento. Durante esta parte, además de la comprobación del funcionamiento y la configuración de sensores, se realizará también la toma del *baseline* en la que todos los sensores tomarán mediciones en las que el sujeto permanecerá en situación de reposo total. El propósito de la toma de esta medida es la adquisición de un punto de referencia en el que el usuario se encuentra en el estado neutro (sin emociones). No será necesario mostrar al sujeto de evaluación ningún estímulo para esta etapa.

4.5.2 Calibración

La calibración es un paso fundamental para lograr los objetivos, dado que las medidas calibradas se utilizarán para encontrar las relaciones entre las emociones y las señales fisiológicas para el usuario concreto que esté llevando a cabo la evaluación. La información de calibración será útil durante el procesado de datos de modo que puedan conocerse tanto los estados neutros como los límites en las dimensiones emocionales evaluadas para el sujeto

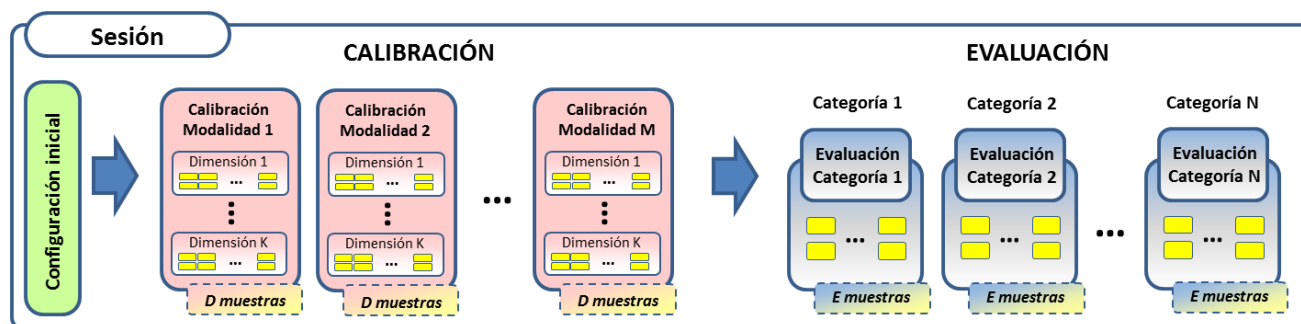


Figura 2. Esquema general de una sesión

calibrado (i.e. un sujeto de evaluación puede mostrar una leve sonrisa al ser inducido con un estímulo etiquetado en el extremo de una dimensión, mientras que otro puede mostrarse mucho más expresivo con el mismo estímulo). La calibración puede ser de 4 tipos distintos, atendiendo a las diferentes modalidades: imagen, melodía, vídeo e interacción. A su vez, para cada modalidad habrá que calibrar cada una de las dimensiones emocionales. Por último, por cada dimensión emocional a calibrar se presentarán al sujeto un número de muestras divididas en dos cantidades iguales: la primera mitad de las muestras contendrá estímulos etiquetados en el máximo de esa dimensión, y la segunda mitad etiquetados en el mínimo. El orden de la presentación de las muestras en la calibración de una modalidad para una determinada dimensión será aleatorio. Como se ha visto antes en el glosario las muestras de la calibración contienen, además del estímulo, un pequeño tiempo de descanso al inicio y una valoración del sujeto al final (donde éste realizará un pequeño test y una autoevaluación en las dimensiones que se esté calibrando). Es importante señalar que una calibración en una determinada modalidad debe hacerse siempre que posteriormente se vaya a evaluar categorías asociadas a dicha modalidad. Esto es, si en una sesión van a evaluarse personajes en forma de imágenes, deberá hacerse previa y necesariamente una calibración de imágenes.

4.5.2.1 Evaluación

La evaluación es aquella parte de la sesión dedicada a tomar medidas con objeto de su posterior análisis. Durante la parte de la evaluación de una sesión, se evalúan categorías (elementos del videojuego). Cada categoría tiene asociada una modalidad. Esto es, si se desea evaluar la categoría relacionada con la banda sonora del videojuego, necesariamente estará asociada a la modalidad melodías. Una categoría está compuesta de un número determinado de muestras. Las muestras, del mismo modo que en la calibración, están compuestas de un breve descanso, un estímulo y una parte de valoración. Los estímulos también estarán divididos a partes iguales, pero en el caso de la Evaluación será en estímulos diana (aquellos con contenidos de los videojuegos a evaluar) y estímulos ruido (aquellos no relacionados con los videojuegos de interés). El descanso será un breve periodo de tiempo durante el cual el sujeto deberá relajarse y la valoración es un pequeño test donde el sujeto responderá preguntas relacionadas con el estímulo y la categoría, autoevaluándose en las dimensiones calibradas.

5. EVALUACIÓN PROPUESTA

El objetivo final de NeuroGame es llevar a cabo una evaluación multimodal y hedónica comparativa del diseño y la jugabilidad de varios videojuegos, con especial atención a los aspectos socioculturales que tengan relevancia en el diseño del juego. En

este apartado se aplica la metodología general presentada en el apartado anterior al diseño de una evaluación concreta. La evaluación tendrá como objeto el análisis de tres videojuegos: *Castelvania: Lords of Shadow*, *Dante's Inferno* y *God of War III*, elegidos por sus similitudes en la historia y en los protagonistas, por estar disponibles en la misma plataforma de juego (Sony *Playstation 3 – PS3*) y por ser del mismo género (i.e. *hack and slash* – basados en combates con armas cuerpo a cuerpo).

La evaluación se llevará cabo en tres sesiones, atendiendo a tres grandes aspectos a evaluar en un videojuego. La primera sesión plantea una Evaluación del Diseño, donde se estudiarán las características del diseño sin interacción, analizando por separado las diferentes componentes estéticas que dan forma característica a cada videojuego (es decir, las imágenes de los personajes principales, las melodías, los escenarios y la historia). La segunda fase es una Evaluación de la Interacción, donde se estudia la interacción directa del usuario con el videojuego, evaluando la integración en un entorno dinámico de las características analizadas previamente durante la fase de juego. Por último, se realizará una Evaluación de la Comunicación, en la cual se valora la efectividad de los elementos utilizados en la campaña promocional del videojuego.

A continuación se explican las tres sesiones propuestas detallando las modalidades, categorías y el número de muestras que intervendrán en cada una de ellas. La Figura 3 muestra el proceso completo de la evaluación propuesta.

5.1 Evaluación del Diseño

El objetivo de esta evaluación es analizar la respuesta emocional y cognitiva producida por aquellas categorías de un videojuego relacionadas con la experiencia de usuario en la fase de juego. Tras realizar la configuración inicial y la calibración en las modalidades de imagen, melodía y vídeo, se procederá a evaluar las siguientes categorías: 1. Personajes, 2. Contexto, Entorno y Ambiente; 3. Música y Efectos Sonoros; 4. Historia y Narrativa. Los estímulos de las muestras (seis por categoría) le serán presentados al sujeto en las modalidades calibradas: imágenes para las categorías 1 y 2, melodías para la categoría 3 y vídeos para la categoría 4. En la evaluación de la primera categoría, le serán presentadas al usuario imágenes de los protagonistas de los videojuegos objeto de estudio. La segunda categoría se centra en el entorno del videojuego, donde le serán presentados al usuario paisajes y lugares ambientados en los videojuegos. La categoría Música y Efectos Sonoros se centrará en reproducir su música y banda sonora. Por último, la categoría Historia y Narrativa presentará al sujeto la sinopsis o fragmentos de los videojuegos en forma de vídeo.

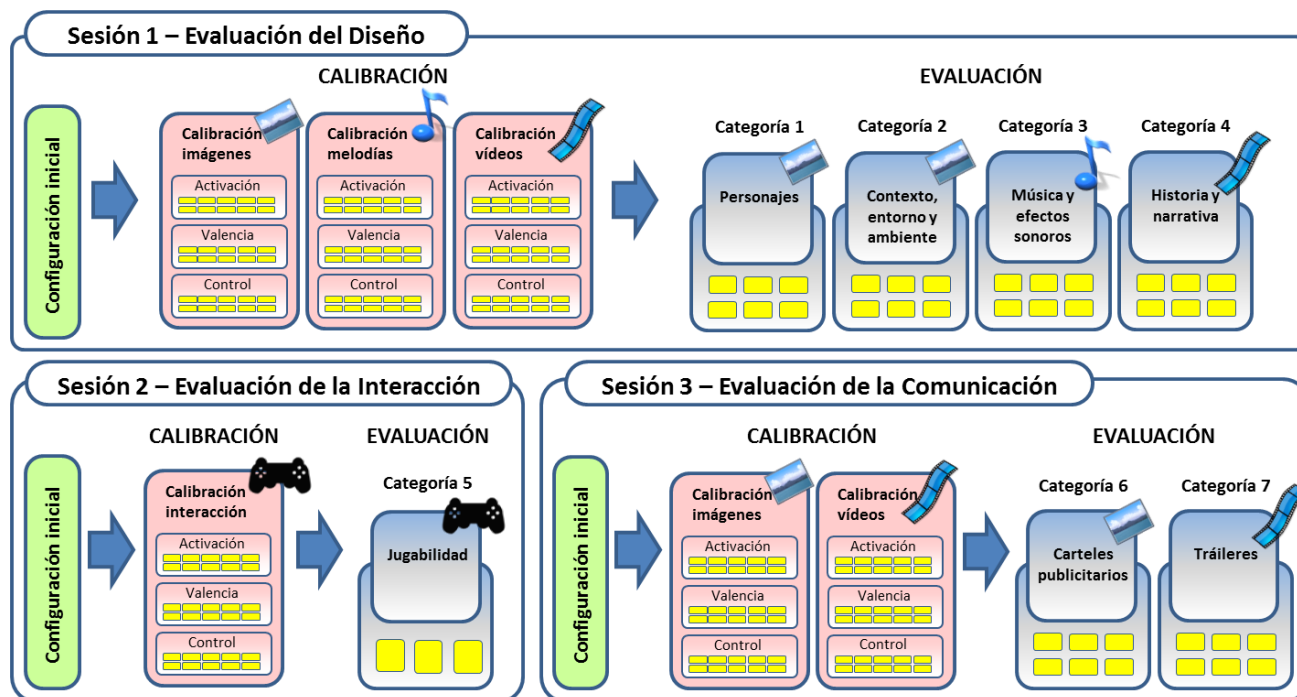


Figura 3. Esquema del proceso completo de la evaluación propuesta

5.2 Evaluación de la Interacción

El objetivo de esta fase es evaluar cómo las emociones se relacionan con la estrategia de juego ante los retos propuestos. Se analizará la dinámica de la respuesta emocional y cognitiva del sujeto durante su experiencia directa con el juego. El proceso será el siguiente: primero se realizará la configuración inicial. Dado que la única categoría a evaluar es la Jugabilidad, se calibrará en la modalidad Interacción. Una vez calibrado el sujeto, seguiremos con la evaluación.

La categoría a evaluar estará compuesta por 3 muestras. Cada una de ellas contiene un breve periodo de descanso, un estímulo en forma de una de las demos de videojuego y una parte de valoración en la que se pide al sujeto que evalúe la sección jugada. Durante la demo, el sujeto se enfrentará al videojuego, debiendo superarla con éxito (para ello, todas las demos se configuran en un nivel de dificultad adecuado de forma que puedan ser superadas en un tiempo razonable).

5.3 Evaluación de la Comunicación

El objetivo de la última sesión es evaluar la efectividad de los elementos utilizados en la campaña promocional. Para el experimento, el estudio se centra en las categorías de carteles publicitarios y tráileres de los diferentes videojuegos objeto de estudio. El proceso comenzará de manera habitual con la configuración inicial, seguida de la calibración en las modalidades de imágenes y vídeo. La evaluación de ambas categorías se desarrollará presentando al sujeto de evaluación 6 estímulos de imágenes en forma de carteles y pósters para la primera categoría, y 6 estímulos de vídeo que serán tráileres. Ambas categorías contendrán 3 estímulos diáfanos correspondientes a los tres videojuegos de interés y otros 3 estímulos ruido. En la parte de valoración de cada muestra, el sujeto de evaluación realizará un pequeño test post-juego para analizar las bondades a nivel de experiencia del juego y de relaciones con los elementos.

6. PRIMERAS PRUEBAS

Con objeto de detectar fallos en la metodología y, muy especialmente, de ver la viabilidad del montaje experimental propuesto se han realizado unas primeras pruebas con usuarios. Las pruebas se llevaron a cabo con tres sujetos no ajenos al proyecto, en una única jornada intensiva en un laboratorio de *BitBrain Technologies* (ver Figura 4). La realización de estas primeras pruebas fue muy satisfactoria, ya que se logró poner en marcha todo el set-up tecnológico con éxito y se comprobó el correcto funcionamiento de todo el software de sincronización y toma de datos de los sensores (ver Figura 5). Para no alargar las pruebas se calibró y evaluó sólo en una dimensión, la de activación. Se consiguieron realizar las partes más importantes (Evaluación del Diseño y Evaluación de la Interacción) obteniendo datos de los sensores para su posterior procesado. Además, los sujetos voluntarios proporcionaron un valioso *feedback* que engloba aspectos de protocolo, de jugabilidad, y de carácter ergonómico a través de los comentarios realizados durante la evaluación y de un informe que redactó cada uno de ellos sobre la experiencia.



Figura 4. Vista del Laboratorio de pruebas durante la evaluación de la jugabilidad de las demos.

En los dos siguientes sub-apartados se recogen los problemas encontrados y las mejoras propuestas para solventarlos. Finalmente, se incluye un apartado que trata sobre los primeros análisis de los datos obtenidos durante las primeras pruebas.

6.1 Problemas encontrados

En las notas tomadas por los miembros del equipo evaluador y los informes de los usuarios se detectaron los siguientes puntos problemáticos:

- [P1] **Duración de las sesiones.** Si bien es cierto que en una primera estimación se ajustaron tiempos, durante el experimento se encontró con que la duración era excesiva para el sujeto de evaluación.
- [P2] **Sensores invasivos.** El *EyeTracker* usado, pese a ser una fuente de información potencialmente interesante, es una tecnología invasiva en cuanto a rasgos faciales se refiere. Se hicieron varias pruebas diferentes pero debido a su colocación sobre la cara y los reflejos producidos, en ninguna de ellas se consiguió que el seguimiento facial fuese el adecuado para la detección de la respuesta emocional facial del usuario (ver Figura 6). Por otra parte, el sensor de saturación de oxígeno ubicado en el dedo índice impedía al usuario sujetar el controlador correctamente.
- [P3] **Dificultad de autoevaluación emocional.** Los sujetos mencionan en sus informes la dificultad de evaluar su propio estado emocional, sin haber recibido algún tipo de entrenamiento con las escalas SAM. Además, la autoevaluación a lo largo del experimento depende en gran medida de la primera autoevaluación inicial realizada, normalmente llevada a cabo de manera imprecisa.
- [P4] **Transición entre estímulos.** Los sujetos expresan que resulta difícil pasar de un estímulo de sonido al siguiente (i.e. volver al estado neutro y prepararse para el siguiente estímulo).
- [P5] **Incomodidad.** En ocasiones, los sujetos se sentían abrumados al estar equipados con tanta tecnología y observadores en la sala.
- [P6] **Volumen de Datos.** Hay que tener en cuenta el despliegue tecnológico que supone una evaluación completa y la cantidad de datos que se están tomando durante la misma. En un determinado momento, hubo que volcar datos entre dispositivos de almacenamiento externo para no saturar los discos duros locales con toda la información que estaban recibiendo.



Figura 5. Sujeto de evaluación con la tecnología equipada

6.2 Mejoras necesarias

En este apartado se recogen aquellas recomendaciones y mejoras relacionadas directamente con los problemas encontrados en las primeras pruebas y descritos en el punto anterior. Estas deberán ser tenidas en cuenta para llevar a cabo la futura evaluación completa con usuarios.

- [S1] Reducción del número de estímulos tanto en calibración como en evaluación de categorías.
- [S2] Supresión del sensor *EyeTracker* (ET) por entrar en conflicto con el reconocedor facial y del sensor de saturación de oxígeno (SpO2).
- [S3] Sustitución de la escala SAM de autoevaluación emocional por una escala visual-analógica (EVA). Dicha escala es análoga al SAM, con la particularidad de que solo contiene indicaciones en los extremos de ésta, resultando menos confusa y más sencilla de utilizar.
- [S4] Ampliación del tiempo de descanso de las muestras para dejar al sujeto volver a un estado neutro emocional. En el caso de estímulos más largos como las demos, puede mostrarse al sujeto material audiovisual relajante.
- [S5] Presencia de un único evaluador en la sala que haga el papel de facilitador, siguiendo un protocolo de actuación perfectamente detallado, que aporte al usuario información constante sobre el proceso y el montaje experimental.
- [S6] Revisión de la necesidad de recopilar toda la información visual de forma continua; estudio de la reducción de la frecuencia de captura de las cámaras de vídeo y aplicación de técnicas de compresión de datos a los vídeos capturados.



Figura 6. Problemas de reflejos generados por el uso del *EyeTracker*.

6.3 Análisis inicial de los datos

Los datos obtenidos de las pruebas están siendo analizados con objeto de adecuar tanto los algoritmos de clasificación emocional facial [9] y fisiológica, como de fusión de los resultados de cada modalidad [8].

Los primeros análisis tanto de los datos faciales (ver Figura 7) como de los fisiológicos están siendo satisfactorios. Se ha podido comprobar la captura satisfactoria por los diferentes sensores de los diversos estados emocionales inducidos en los usuarios dispuestos durante la evaluación. Otro de los aspectos que cabe destacar es la reafirmación de la importancia de la utilización de sensores complementarios (tanto fisiológicos como no fisiológicos), que suplan los fallos en la clasificación emocional de una modalidad. Así, los sensores fisiológicos son capaces de recoger/detectar información emocional incluso cuando el usuario no la está mostrando explícitamente mediante sus expresiones

faciales. Y los faciales pueden ayudar a discernir entre estados emocionales que den lugar a respuestas fisiológicas similares.

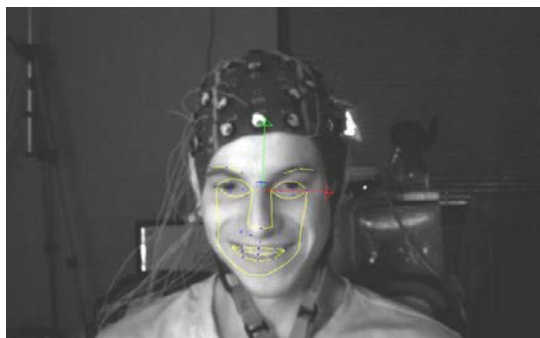


Figura 7. Primeros resultados del analizador facial. Usuario contento tras pasar uno de los niveles.

7. CONCLUSIONES Y TRABAJO FUTURO

En este artículo se ha presentado NeuroGame, un proyecto que recoge la experiencia de un equipo multidisciplinar especializado por un lado, en la evaluación hedónica de videojuegos y por otro, en la detección facial y fisiológica del estado emocional de los usuarios. El objetivo inicial del proyecto y en el que se ha centrado este artículo, es el desarrollo de una metodología para la medición y análisis de información asociada a la evaluación hedónica de videojuegos, abarcando no sólo la evaluación de la interacción, sino del diseño y comunicación del videojuego. Una vez realizada una primera prueba con tres usuarios, y comprobada la viabilidad del montaje experimental y de la metodología general propuestos, el siguiente paso será mejorar el protocolo existente para afrontar las futuras evaluaciones con usuarios. El objeto final del proyecto es contribuir al enriquecimiento de los métodos tradicionales de evaluación, haciéndolos presentes en las etapas más tempranas del desarrollo de videojuegos.

8. AGRADECIMIENTOS

Este trabajo ha sido parcialmente financiado por el Gobierno Español a través del DGICYT contrato TIN2011-24660.

9. REFERENCIAS

- [1] Bradley, M.M. and Lang, P.J. 1994. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry* 25, 49-59.
- [2] Brave, S. and Nass, C. 2002. Emotion in human-computer interaction. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications (chap. 4)*, 81-96. J. Jacko & A. Sears (Eds.) Hillsdale, NJ: Lawrence Erlbaum Associates.
- [3] Costa, T., Rognoni, E. and Galati, D. 2006. EEG phase synchronization during emotional response to positive and negative film stimuli. *Neuroscience Letters* 406, 159-164.
- [4] Cuthbert B.N., Schupp H.T., Bradley M.M., Birbaumer N. and Langey P.J. 2000. Brain potentials in affective picture processing covariation with autonomic arousal and affective report. *Biological Psychology* 52, 95-111.
- [5] Ekman, P., Dalglish, T. and Power, M. 1999. *Handbook of Cognition and Emotion*. Wiley Online Library.
- [6] González Sánchez, J.L., Gil Iranzo, R.M. and Gutiérrez Vela, F. L. 2011. Enrichiendo la Evaluación en Videojuegos. *Actas del XII Congreso Internacional de Interacción Persona-Ordenador. Interacción 2011* (Lisboa, Portugal, 2-6 de Septiembre). Néstor Garay y Julio Abascal (Eds.), 273-282.
- [7] Gunes, H. and Pantic, M. 2010. Automatic, dimensional and continuous emotion recognition. *Int. Journal of Synthetic Emotions (IJSE)*, 1(1), 68-99.
- [8] Hupont, I., Ballano, S., Cerezo, E. and Baldassarri, S. 2011. Scalable multimodal fusion for continuous affect sensing. In *Proceedings of IEEE Workshop on Affective Computational Intelligence - WACI* (Paris, 11-15 April), 1-8.
- [9] Hupont, I., Cerezo, E. and Baldassarri, S. 2010. Sensing Facial Emotions in a Continuous 2D Affective Space. In *Proceedings of IEEE International Conference on Systems, Man and Cybernetics – SMC 2010*, 2045-2051.
- [10] Kim, Y.O., Paik, J.K., Heo, J., Koschan, A., Abidi, B. and Abidi, M. 2003. Automatic face region tracking for highly accurate face recognition in unconstrained environments. In *Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'03)*, 29-36.
- [11] Law, E.L-C., Roto, V., Hassenzahl, M., Vermeeren, A. and Kort, J. 2009. Understanding, scoping and defining user experience: a survey approach. In *Proceedings of CHI 2009* (Boston, USA, April 04-09), 719-728.
- [12] Liu Y., Sourina, O. and Nguyen, M. K. 2010. Real-time EEG-based Human Emotion Recognition and Visualization. In *Proc. 2010 Int. Conf. On Cyberworlds*. (Singapore, 20-22 Oct), 262-269.
- [13] Maaoui, C. and Pruski, A. 2010. Emotion Recognition through Physiological Signals for Human-Machine Communication. *Cutting Edge Robotics (chap. 20)*, Vedran Kordic (Ed.) 317-332.
- [14] Mandryk, R., Inkpen, K. and Calvert, T.W. 2006. *Using psychophysiological techniques to measure user experience with entertainment technologies*. *Behaviour & Information Technology*, 25(2), 141 – 158.
- [15] Nasoz, F., Alvarez, K., Lisetti, C.L. and Finkelstein, N. 2003. Emotion recognition from physiological signals for presence technologies. *International Journal of Cognition, Technology and Work, Special Issue on Presence*, 6(1), 1-32.
- [16] Norman, D. 2002. *Emotion and Design: Attractive things work better*. *Interactions Magazine*, 9(4), 36-42.
- [17] Picard, R.W. 1997. *Affective Computing*. MIT Press
- [18] Ravaja, N., Salminen, M., Holopainen, J., Saari, T., Laarni, J. and Järvinen, A. 2004. Emotional response patterns and sense of presence during video games: potential criterion variables for game design. In *Proceedings of the third Nordic conference on Human-Computer Interaction* (Tampere, Finland), 339-347.
- [19] Russell, J.A. and Mehrabian, A. 1977. Evidence for a three-factor theory of emotions. *Journal of Research in Personality* 11(3), 273-294.
- [20] Wagner, J., Kim J. and André, E. 2005. From physiological signals to emotions: implementing and comparing selected methods for feature extraction and classification. *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference*, 940-943