

Analyzing the use of Gamification in a Tutorial for a Citizen Science Project

Agustín Kanner^a, Diego Torres^{a,b,c}, Julieta Lombardelli^a

^a LIFIA, Fac. de Informática, UNLP

^b Dep. Ciencia y Tecnología, UNQ

^c CICPBA, Prov. Buenos Aires

La Plata, Argentina

agustin.kanner@gmail.com,

diego.torres@lifia.info.unlp.edu.ar

julieta.lombardelli@lifia.info.unlp.edu.ar

Cintia Peri

Inst. Argentino de Radioastronomía - CONICET-CICPBA

Planetario Ciudad de La Plata, UNLP

La Plata, Argentina

cintiaperi@gmail.com

Abstract—Citizen science is a way of doing science in which people who do not necessarily have previous training to be part of a scientific project are invited. One of the challenges presented by citizen science is the training of volunteers in the tasks required by the research project. A valid approach is through gamification strategies. Gamification is the use of game elements in contexts that were not intended to be a game. In this article, we present a study of the use of gamification in tutorials for a citizen science project of astronomy photo classification. The evaluations showed that the aesthetics, the evaluation by points, the promotion with badges and the competition have been promising elements in this type of tutorials.

Index Terms—Citizen Science, Gamification, Evaluation

I. INTRODUCCIÓN

La ciencia ciudadana [1] es una forma de hacer ciencia en la que se invita a personas sin una formación necesaria previa a formar parte de un proyecto científico y realizando diferentes etapas de los mismos. En estos proyectos, cualquier persona sin necesidad de poseer una formación específica realiza tareas voluntariamente cediendo parte de su tiempo y en algunos casos recursos. A estas personas las llamamos científicos ciudadanos [2].

Uno de los desafíos que presenta la ciencia ciudadana es la formación de los voluntarios en las tareas que requiere el proyecto de investigación para que las mismas sean de calidad [3]. Las tareas que realizan los científicos ciudadanos pueden ser, de recolección de muestras, clasificación, análisis, o de transcripción. Por ejemplo, en un proyecto de clasificación de fotos, los científicos ciudadanos deben entrenarse para detectar las particularidades de las fotos que las hacen pertenecer a una categoría u a otra.

En todas ellas, deben incluirse una formación que permita que los voluntarios realicen la tarea en forma adecuada [4], [5] y que a la vez sea dinámica y entretenida para que luego de recibir el entrenamiento, los voluntarios deseen continuar con la actividad en sí.

La formación de los científicos ciudadanos se da en diferentes formas como pueden ser los encuentros presenciales de formación o manuales estáticos y multimediales. Por ejemplo,

el proyecto E-bird¹ y su capítulo argentino Aves Argentinas² [6], tutoriales multimedia [7], y acercamientos a la actividad guiados por juegos [8], [9].

Una forma de abordar una experiencia de aprendizaje es a través de estrategias de *gamification* [10]. *Gamification* es la utilización de elementos basados en el juego, como las mecánicas de juego, la estética y la forma de pensar en el juego para atraer a las personas, motivar la acción, promover el aprendizaje y resolver problemas. Es decir que la *gamification* no es un juego completo, sino una parte del concepto juego. Existen experiencias entre las que podemos destacar "I want to be a captain! I want to be a captain!" (Eveleigh et. al.) [11] donde se discute el uso de una narrativa y elementos de *gamification* dentro de un proyecto de ciencia ciudadana. También existen experiencias en el proyecto Eye-Wire (www.eyewire.org) y FoldIt (www.fold.it) en los cuales a través de *gamification* se deben ir coloreando imágenes de neuronas o completando un puzzle en modelos 3D. En todos los casos el uso de la gamificación se utiliza para motivar la participación.

En este artículo, presentamos un estudio de la utilización de *gamification* en tutoriales para un proyecto de ciencia ciudadana de clasificación de fotos de astronomía. En este proyecto, los ciudadanos deben detectar la presencia de unos arcos de color en fotos tomadas a estrellas particulares llamadas *runaway Stars* o Estrellas Fugitivas. La dificultad de esta clasificación radica en que los arcos de las fotos pueden ser de diferentes colores, o la ubicación puede ser diferente entre una estrella y otra. Además, las estrellas en las imágenes no poseen la forma característica que se espera en el imaginario de la sociedad sino que simplemente se ven como un punto en el centro de la foto.

El estudio se desprende de la implementación de tres prototipos evaluados en diferentes oportunidades. Cada prototipo presenta características diferentes sobre la utilización de elementos de *gamification*. En las evaluaciones analizamos

¹<https://ebird.org> accedido el 10 de mayo de 2018.

²<http://www.avesargentinas.org.ar> accedido el 10 de mayo de 2018.

qué elementos de *gamification* fueron más valorados por los usuarios, en qué medida han aprendido lo que el tutorial presentaba, el nivel de satisfacción y usabilidad. Las evaluaciones mostraron que las estéticas, la valoración por puntos, la premiación con insignias y la competencia han sido elementos promisorios en este tipo de tutoriales.

El artículo se encuentra organizado de la siguiente forma. En la siguiente sección se describe el problema general del proyecto astronómico y la relevancia de las estrellas *runaway*. La Sección III describe los desafíos relacionados con la clasificación de imágenes en el dominio específico de este tipo de estrellas. Luego, la Sección IV presenta el enfoque de *gamification* para abordar los desafíos. La forma de aplicar el enfoque se describe en la Sección IV-A y el prototipo de RunawayStars se describe en la Sección V. Una evaluación sobre la interacción con los prototipos se presenta en la Sección VI. Finalmente las conclusiones y trabajo futuro se describen en la Sección VII respectivamente.

II. RUNAWAY STARS: EL PROYECTO CIENTÍFICO

Las estrellas de alta velocidad, conocidas como *runaway stars* por su nombre en inglés, son estrellas que se mueven con una velocidad espacial alta respecto a la velocidad de rotación media de la galaxia [12]. En algunas ocasiones estas estrellas generan una estructura a su alrededor conocida como *bow shock*.

Los *bow shocks* obtienen su nombre por poseer una forma de arco, muy sencilla de identificar, similar a la forma que tiene una onda en el agua adelante de un barco moviéndose por ella. En los catálogos E-BOSS 1 y 2 [13], [14] se generó por primera vez una búsqueda sistemática de *bow shocks* sobre unas 630 *runaway stars*. Se presentaron en total 73 objetos y se analizaron los resultados en base a las muestras utilizadas. El trabajo se realizó mediante inspección visual de imágenes astronómicas infrarrojas de dominio público. Para llegar a ese número, se analizaron casi 3000 imágenes.

Ahora bien, las distintas estructuras que se generan alrededor de estrellas *runaway* pueden variar. No son todos *bow shocks*. Mas aún, incluso puede no formarse nada. Esto depende de una gran cantidad de parámetros de la estrella y del medio por el que se mueve. La combinación de estos parámetros hará también que varíe la forma y brillo de la estructura que se forma en el entorno de estas estrellas. Por ejemplo, puede verse un arco bien distinguible en la dirección de movimiento de la estrella, o el arco puede poseer poco brillo en la imagen, o aunque exista un *bow shock* no puede verse o no se ve como un arco, o que directamente no haya *bow shock* para la estrella seleccionada.

La búsqueda visual de *bow shocks* o estructuras similares, alrededor de estrellas *runaway*, es simple y puede realizarla cualquiera, con un mínimo entrenamiento visual. El problema es, la enorme cantidad de trabajo que representa poder aportar un mínimo resultado a la la estadística general en el tópico. Por este motivo un proyecto de Ciencia Ciudadana resulta vital. ‘Runaway stars’, tiene como objetivo encontrar *bow shocks*

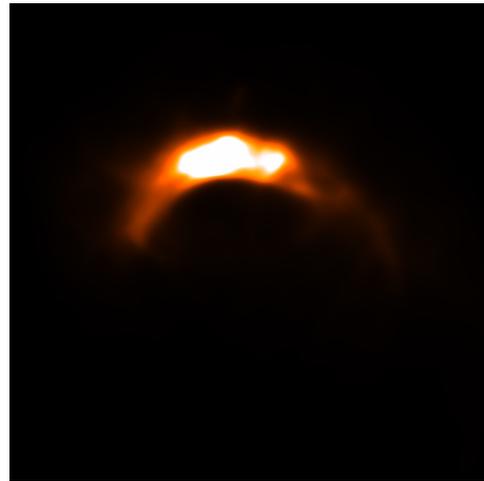


Fig. 1: Estrella con un bow shock arriba

estelares alrededor de estrellas de alta velocidad, en imágenes infrarrojas.

III. DESAFÍOS DE COMPRESIÓN

La implementación de un proyecto de ciencia ciudadana para la clasificación de imágenes de *runaway stars* genera una serie de desafíos relacionados con la formación de los científicos ciudadanos. A continuación se describen en forma general algunos de los desafíos detectados:

Contexto y vocabulario: Debido a la naturaleza del problema, es necesario comprender los conceptos astronómicos relacionados a las *runaway stars* y los *bow shocks*. En un proyecto de ciencia ciudadana en línea el participante no tiene a su disposición un astrónomo constantemente para evacuar sus dudas, es por esto que el material se le debe ser suministrado lo más claramente posible.

Visualización de una estrella: las imágenes que se utilizan para la clasificación presentan un segmento cuadrado del espacio con la estrella elegida en el centro del cuadrado (Fig. 1). Esta estrella en ocasiones se visualiza con un punto y en otros casos no llega a visualizarse. Esta forma de visualización es muy diferente al imaginario de una estrella de cinco puntas que poseen en general los ciudadanos.

Formas de los bow shocks: Un *bow shock* prototípicamente tiene la forma de un arco de color al costado de la estrella en el centro de la imagen, sin embargo este arco puede no estar completamente definido y el mismo puede ser de cualquier color.

Varias imágenes representan la misma estrella: Debido al proceso por el cual se generan las imágenes, es posible obtener diferentes imágenes de la misma estrella. Estas imágenes varían en intensidad de la luz y por lo tanto, la intensidad del *bow shock*, así mismo es posible representar las imágenes con distintas paletas de colores, de donde algunas resaltarán el *bow shock* y más que otras. Debido a que no existe una combinación que funcione para todos los casos, y tampoco es posible predecir la mejor combinación para cada uno, se le muestra a los científicos ciudadanos diferentes imágenes

de una misma estrella cada una generada con una alternativa diferente. La dificultad radica en que aunque se visualicen diferente, la estrella siempre es la misma. Además, si alguna de las imágenes de una estrella presenta un *bow shock* entonces la estrella es de las buscadas aunque en otras imágenes no se visualice. Este pensamiento requiere un gran nivel de abstracción por parte de los ciudadanos, el cual debe ser enseñado en del tutorial.

Es posible equivocarse: Lo importante es la participación y los errores se mitigan entre la masa de participantes. En los proyectos de clasificación, la tendencia de la mayoría de las clasificaciones es la que se tiene cuenta como valor final. Es decir, si la mayoría de las personas indica que una foto pertenece a la categoría A, entonces se la considera en esa categoría.

IV. ENFOQUE: USAR *gamification*

Una forma de abordar la experiencia de aprendizaje es a través de estrategias de *gamification*. *Gamification*, es la utilización de elementos basados en el juego [10], como son las mecánicas de juego, la estética y la forma de pensar en el juego para atraer a las personas, motivar la acción, promover el aprendizaje y resolver problemas. Es decir que la *gamification* no es un juego completo, sino una parte del concepto de juego.

El principal motivo en la implementación de *gamification* se enfoca en motivar y atraer a los posibles usuarios y facilitarles el proceso de aprendizaje o de entendimiento de la acción en cuestión.

Entre los elementos de *gamification* [15], [16] podemos enumerar :

- **Reglas**, que determinan la guía a través del cual los usuarios progresan.
- **Estructura de recompensas** que se puede visualizar en insignias o subir de niveles, como sistema de fácil lectura para analizar los logros obtenidos.
- **Tabla de posiciones**, que enumera una lista con los participantes que obtuvieron los mejores logros.
- **Puntaje**, implementándose como sistema mensurable de recompensa por logros, o por respuestas correctas como así también se puede utilizar para desbloquear contenidos.
- **Devolución en tiempo real**, es decir que mientras se actúa en el sistema propuesto, se perciba una devolución de los logros adquiridos
- **Desafíos**, para lograr atraer al usuario a intervenir en la experiencia, debe percibir cierto grado de dificultad, de intención de obstáculo a superar.
- **La progresión**, es decir el movimiento en etapas hacia estados más avanzados en el contenido, entendiendo que el usuario pueda percibir ese aprendizaje progresivo.
- **Objetivos claros**, que facilitan un propósito estimulando el acercamiento, participación y permanencia en la experiencia.
- **Status**, hace visible a otros individuos la posición, o las habilidades obtenidas.

Cabe destacar que en un determinado contenido es posible aplicar uno de estos elementos en forma individual o todos simultáneamente, puesto que finalmente los elementos se articulan en forma gradual para lograr la experiencia o el aprendizaje deseado.

A. *Gamification en Runaway Stars*

Uno de los principales objetivos de la Ciencia Ciudadana es conseguir la mayor cantidad de participación voluntaria y motivarlos a continuar colaborando en la experiencia propuesta. De aquí se desprenden una serie de desafíos específicos detallados a continuación.

El primer desafío tiene relación con el entrenamiento visual que se requiere para poder identificar correctamente las imágenes presentadas y sus características. De esta forma, se recurre a una característica que se implementa desde la *gamification* la cual puede ser concebida como Experiencia Estética [17].

El diseño integral de la *gamification* en esta experiencia esta constituido por dos aristas: el diseño de la estética y los elementos de *gamification*.

Con el diseño de la estética se busca asegurar una experiencia que integra los sentidos, en este caso específico a través del reconocimiento visual. Los recursos visuales se distinguen como:

- Aumento de contrastes en los colores de las zonas en donde se desea poner atención en el detalle
- Implementar un fondo o *backdrop* que hace alusión a la temática en la que se desea participe el ciudadano, para lograr un ambiente más inmersivo y motive la concentración en el proceso.
- Señalamiento por texto pero no invasivo, sobre los puntos de atención en las imágenes, para propiciar una devolución en tiempo real del aprendizaje sobre las mismas.

Por su parte, los elementos de *gamification* utilizados son:

- Asignación de puntos a cada respuesta correcta. De esta forma se puede otorgar a los participantes autonomía en identificar su nivel de aprendizaje al mismo tiempo que podemos evaluar si el diseño integral de la experiencia facilita el reconocimiento de las Runaway Stars.
- Insignias: Con este símbolo se hacen visibles los logros, lo que puede motivar a seguir participando en el entrenamiento.
- Tabla de posiciones: para que el ciudadano pueda comparar sus resultados con los de los participantes anteriores. A través de este recurso se propicia una oportunidad de interactuar socialmente y establecer discusiones con pares para motivar superarse.
- Niveles: El entrenamiento es progresivo, para asegurar que el participante reconozca activamente el progreso en su aprendizaje.
- Objetivos: Se determina en forma transparente y clara el objetivo principal que es aprender a reconocer ciertas características de una Runaway star.

V. LOS PROTOTIPOS DE RUNAWAY STARS

En *Runaway Stars* los voluntarios observan una estrella a la vez e indican si al rededor de la estrella observan o no un *bow shock*, nos referiremos a esta observación y análisis de una estrella como "pregunta". Los objetivos durante el desarrollo de este proyecto son buscar que los participantes respondan la mayor cantidad de preguntas posibles manteniendo un nivel aceptable de respuestas correctas.

En esta etapa de prototipo del proyecto, todas las imágenes de estrellas son conocidas por nosotros, es decir sabemos con exactitud en qué imágenes hay *bow shocks*

Para poder evaluar el enfoque planeado se han construido una serie de prototipos para el proyecto *Runaway Stars*. La interacción en todos los prototipos posee una estructura similar organizada en tres partes:

Tutorial: En esta etapa los participantes reciben la información general para comprender qué es un *Bow shock* y cómo detectarlos.

Entrenamiento: Durante esta etapa los participantes deberán responder una cantidad fija de preguntas, estas preguntas son iguales para todos los participantes, luego de cada pregunta respondida el participante conocerá si su respuesta fue correcta o incorrecta. Una vez que el participante haya realizado la cantidad preestablecida de preguntas en esta etapa se le indicará si la opción elegida es la correcta y una recompensa. Las recompensas pueden ser puntos o insignias. Dependiendo de la cantidad de respuestas correctas las insignias pueden ser de principiante, intermedio o experto. El usuario puede optar por repetir esta etapa o pasar a la siguiente.

Preguntas de control: En esta etapa el participante podrá responder la cantidad de preguntas que desee. A diferencia de la etapa anterior, cada usuario recibirá imágenes de forma aleatoria y no recibirá la respuesta, es decir, no sabrá si su respuesta fue correcta o no. Con esto se simula el caso real de la clasificación. Cuando el participante lo desee puede terminar de contestar preguntas y así finalizar su participación.

Durante el desarrollo de *Runaway stars* hemos desarrollado diferentes versiones para evaluar las diferentes alternativas, cada versión posee un nombre con el propósito de identificarlas luego en la etapa de evaluación:

YesOrNo En esta primera versión, el participante debe observar una imagen correspondiente a un *runaway star* y responder si en la imagen observa o no un *bow shock*. Debido a que las estrellas pueden ser observadas utilizando diferentes patrones de colores se decidió hacer dos alternativas diferentes:

OnlyOne En la primer alternativa solo mostramos una imagen por pregunta, en el caso de que tengamos más de una imagen por estrella (es decir imágenes con diferentes patrones de colores) cada imagen es considerada como una pregunta diferente, actualmente poseemos cuatro imágenes por estrella, es decir, que habrá cuatro preguntas por estrella, una por cada imagen de la misma. En esta versión la etapa de entrenamiento consiste de siete preguntas, mientras que la etapa de control es libre, es decir, el usuario decide cuando finalizar.

Multiple En la segunda alternativa, por cada estrella se muestran cuatro imágenes generadas con diferentes reglas



Fig. 2: Prototipo *Multiple*

(Fig. 2. El participante debe indicar si visualiza o no un *bow shock* en al menos una imagen.

En esta versión la etapa de entrenamiento consiste de cinco preguntas, mientras que la etapa de control es libre, es decir, el usuario decide cuando finalizar. Luego de la etapa de entrenamiento el usuario visualizará los puntos o insignias que ha obtenido dependiendo de la cantidad de respuestas correctas.

Los conceptos necesarios para detectar *bow shocks* se indican mediante texto no invasivo sobre los puntos a observar en las imágenes, de esta forma se logra focalizar la atención del participante a las características de las imágenes que debe observar y se acostumbra al mismo a realizar las tareas que deberá afrontar en las futuras etapas.

El participante solo puede avanzar a la siguiente etapa si completa el tutorial exitosamente, es decir, si encuentra los *bow shocks* o la falta de ellos en las imágenes dadas.

Al igual que la alternativa anterior, luego de la etapa de entrenamiento el participante recibió puntos o insignias y la posibilidad de comparar su resultado con los resultados de los demás participantes.

Evolution: Esta versión se basó en el los prototipo *Multiple* al que se le aplicaron diferentes elementos relacionados a la experiencia del usuario:

- Los textos de ayuda en la etapa del tutorial se escribieron de forma más informal.
- Durante el tutorial se iluminan intermitentemente los *bow shocks* en las imágenes, con esto aumenta el contraste de color en la porción de la imagen en donde se encuentra el *bow shock*, con esto se busca atraer la atención del participante hacia los *bow shocks*.
- Se marca la estrella de la imagen con una imagen del estereotipo de una estrella. Con esto se busca que el participante observe con más naturalidad la estrella dentro de la imagen e inmediatamente busque un *bow hock* alrededor de la misma.
- Se hizo más enfática la pregunta señalando como objetivo si visualiza en la imagen un arco de color. De esta forma se busca asemejar al *bow shock* a una forma más familiar para el participante, haciendo que le sea más sencillo encontrarlos.
- Se aplicó un fondo que hace referencia a la astronomía, de esta forma se busca aumentar la inmersión del participante aumentando su nivel de concentración.

VI. EVALUACIÓN

Para evaluar la propuesta realizamos una serie de evaluaciones con el objetivo de responder las siguientes preguntas:

- 1) ¿Cuál fue el nivel de usabilidad del prototipo?
- 2) Los participantes que realizaron el tutorial, ¿han adquirido los conocimientos necesarios para la clasificación de imágenes?
- 3) ¿El nivel de confianza de los usuarios luego de finalizar el tutorial, se correspondió con el conocimiento real? En este caso se quiere verificar que ante un nivel de confianza alto, las respuestas sin asistencias sean correctas. /
- 4) ¿Qué elementos de *gamification* fueron los más relevantes para los usuarios?
- 5) Luego de finalizar el tutorial, ¿los ciudadanos continuaron con la clasificación de las imágenes?

Para poder responder las preguntas enunciadas anteriormente, realizamos una serie de evaluaciones con voluntarios en donde utilizaron diferentes versiones del prototipo. Luego, cada participante completó una serie de cuestionarios. La información que recabamos de cada muestra incluye:

- Cantidad de participantes.
- Preguntas correctas durante la fase de entrenamiento.
- Nivel de confianza de cada participante luego del entrenamiento (del uno al diez).
- Cantidad de preguntas contestadas luego del entrenamiento.
- Cantidad de preguntas contestadas correctamente luego del entrenamiento.
- Evaluación de usabilidad utilizando el cuestionario SUS [18].

A. Desarrollo de las evaluaciones

Se realizaron tres evaluaciones en los siguientes contextos:

Doctorado en Informática y Planetario de La Plata

La segunda prueba se realizó con alumnos del Doctorado en Informática de la Facultad de Informática de la UNLP y Asistentes al planetario de la ciudad de La Plata. Estas pruebas se realizaron dentro del edificio de la Facultad de Informática de la UNLP y en el hall del edificio del planetario.

Meet the devs Estas evaluaciones se realizaron en el contexto de la convención "Meet The Devs" realizada en el Centro Cultural San Martín de la ciudad de Buenos Aires. "Meet The devs" es una exposición de videojuegos argentinos, en la cual los participantes pueden ver y probar los videojuegos argentinos que se están desarrollando y conocer a sus creadores.

Expo ludica La última evaluación se realizó en la exposición "Expo Ludica" en el Pasaje Dardo Rocha de la ciudad de La Plata. "Expo ludica" es una exposición de juegos en general incluyendo videojuegos, juegos de tablero, juegos de cartas, entre otros. Los participantes en esta exposición fueron entusiastas de los videojuegos y desarrolladores.

TABLE I: Configuración de las evaluaciones

	OnlyOne X	Multiple X	Evolution
Doctorado y Planetario			
Meet the dev			X
Expo Lúdica			X

B. Resultados

A continuación detallaremos los resultados para las diferentes evaluaciones. La Tabla I muestra la organización del uso de las versiones de los prototipos en los diferentes contextos.

C. Expo ciencias

En la exposición participaron 25 personas. La edad promedio de los participantes fue de 29,8 años, con un desvío estándar (DE) de 10,92, dichos participantes respondieron tres preguntas a modo de entrenamiento y siete preguntas a modo de control con los siguientes resultados:

de las 75 preguntas de entrenamiento, 72 fueron correctas (96 %). Además, se respondieron 175 preguntas de control de las cuales 162 fueron respondidas correctamente (96,57 %). Los Participantes indicaron un nivel de confianza promedio de 7,4 (DE 1.65), y la evaluación SUS obtuvo 86,56.

D. Doctorado en informática y Planetario de La Plata

los participantes se dividieron en 2 grupos, un grupo utilizó la versión del prototipo *OnlyOne* y la otra parte utilizó *Multiple*. Adicionalmente a las pruebas con el prototipo y el cuestionario SUS, los participantes respondieron el siguiente cuestionario:

- ¿Que te pareció el entrenamiento?
- ¿repetí el entrenamiento?
- ¿Que tipo de puntaje te dimos? (Puntos o insignias)
- ¿Notaste la tabla de lideres? ¿te interesó?
- ¿Como crees que te fue con las preguntas que no te dimos respuesta? (1-5)

1) *OnlyOne*: 13 participantes utilizaron el prototipo con una sola imagen por estrella, los mismos tenían una edad promedio de 35,15 años (DE 7,65). De los 13 participantes, 8 eran docentes universitarios participando del posgrado y 5 eran personas interesadas por la astronomía que se acercaron al planetario. Esta versión del prototipo contaba con siete preguntas de entrenamiento, el participante tenía la posibilidad de repetir la etapa de entrenamiento y una cantidad libre de preguntas a modo de control, es decir el participante decidía cuando finalizar la prueba. Como resultado obtuvimos que, de las 106 preguntas de entrenamiento realizadas 71 fueron respondidas correctamente (66,9 %). Luego de este entrenamiento los participantes indicaron un nivel de confianza promedio de 5,28 sobre 10 (DE 2,46).

En total se respondieron 289 preguntas de control de las cuales 201 fueron correctas es decir 69,55 %. Los participantes realizaron en promedio 18,27 preguntas de control, sin embargo con desvío estándar de 17,06; Esto se debe a que algunos participantes respondieron una gran cantidad de

preguntas (hasta 52 preguntas), otros en cambio decidieron responder una o dos preguntas solamente, e incluso 1 caso no respondió ninguna de las preguntas.

En esta versión del prototipo a los participantes se les asignaron puntos o insignias luego de completar el entrenamiento.

Notamos que a los participantes que se les asignaron insignias tuvieron un mejor feedback que los participantes que recibieron puntos, debido a que las insignias son más llamativas gráficamente que los puntos y pueden llegar a dar una mayor sensación de recompensa, sin embargo muchos participantes no notaron la tabla de líderes y, los que los notaron, no les interesó conocer su puntaje.

Con respecto a la evaluación SUS, la misma obtuvo un resultado de 77,18 sobre 100.

En cuanto al cuestionario realizado, de las 8 personas que recibieron insignias, 2 personas no las notaron, mientras que de las 5 personas que recibieron puntos ninguno los notó. Los participantes indicaron su percepción sobre como le fue luego del entrenamiento, los participantes tuvieron un promedio de 2.95 sobre 5 (Desvío estándar 0.68)

2) *Multiple*: Doce participantes utilizaron el prototipo con múltiples imágenes por estrella, los participantes tuvieron una edad promedio de 32,25 (DE 9,15). De los doce participantes, ocho eran docentes universitarios participando del posgrado y cuatro eran personas interesadas en astronomía que se acercaron al planetario.

De las 70 preguntas de entrenamiento, 48 fueron respondidas correctamente (68,57 %). Luego de este entrenamiento, los participantes indicaron un nivel de confianza promedio de 5,6 (DE 3,08). Luego del entrenamiento, los participantes realizaron en promedio 11,91 (DE 7,94). Esto indica que la cantidad de preguntas de control realizadas entre los participantes de versión del prototipo es más homogénea si la comparamos con la versión de una imagen por estrella. En total se respondieron 143 preguntas de control, de las cuales 94 preguntas fueron respondidas correctamente, es decir 65,73 %.

Al igual que la versión con una sola imagen, a los participantes se les consultó sobre los elementos de *gamification* incluidos en el prototipo. Los participantes se mostraron más entusiasmados con las insignias que con los puntos luego del entrenamiento. Al mismo tiempo, a los participantes no les interesó conocer su puntaje en comparación con el resto de los participantes. Con respecto a la evaluación SUS tuvo un valor de 58.75, lo cual indica que los usuarios se sintieron menos cómodos al utilizar la versión con muchas imágenes en comparación con la versión con una sola imagen.

El cuestionario realizado arrojó que, de las 4 personas que recibieron insignias solo 3 las notaron, mientras que de las 8 personas que recibieron puntos, la mitad de ellos no los notó, as mismo los participantes indicaron que, según ellos mismos, su percepción de sus resultados obtenidos luego del entrenamiento fue de 3 sobre 5 en promedio (desvío estándar 1.07)

3) *Meet the devs*: En la exposición "Meet the devs" participaron 7 personas, las mismas eran desarrolladores de

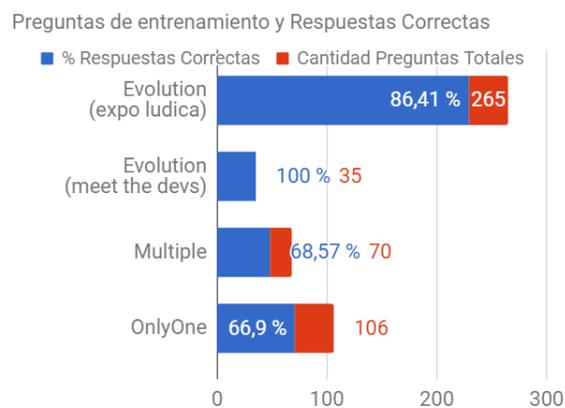


Fig. 3: Preguntas de entrenamiento

videojuegos. Los participantes respondieron cinco preguntas de entrenamiento y luego una cantidad indefinida de preguntas de control, es decir, los participantes podían terminar la prueba cuando ellos querían.

En total se respondieron 35 preguntas de entrenamiento, con 35 respuestas correctas, es decir, el 100 % de las preguntas fueron respondidas correctamente. Luego de este entrenamiento los participantes indicaron un nivel de confianza promedio de 9 con un desvío estándar de 1. Se respondieron 120 preguntas de control de las cuales 85 fueron respondidas correctamente, es decir 70,81 %, en promedio los participantes respondieron 8 preguntas con un desvío estándar de 4,14, de este promedio estamos excluyendo a un participante que respondió 72 preguntas ya que consideramos que no es representativo en relación al resto de los participantes.

A diferencia de las pruebas anteriores a los participantes les gustó la idea de comparar su puntaje con las del resto de los participantes. En cuanto a la evaluación SUS, la misma tuvo un valor de 78.5, lo cual muestra una gran mejoría con respecto de las evaluaciones anteriores.

4) *Expo Lúdica*: La última prueba se realizó dentro del la exposición "Expo Ludica" en la que participaron 53 personas.

Cada participante respondió 5 preguntas de entrenamiento. En total se realizaron 265 preguntas de entrenamiento, de las cuales 229 fueron respondidas correctamente, es decir un 86.41 % de respuestas correctas dentro del entrenamiento. Luego de esta etapa los participantes indicaron un nivel de confianza de 6,9 en promedio (DE 2.7). Se respondieron 486 preguntas de control, de las cuales 386 fueron correctas, es decir 79.42 %. En promedio, cada participante respondió 9,16 preguntas de control (DE 8,11).

Al igual que la prueba en "Meet The Devs", los participantes reaccionaron de forma positiva al ver su puntuación comparada con la de otros. La evaluación SUS obtuvo 78,75.

E. Comparación entre versiones

En esta sección analizamos en forma comparativa los resultados obtenidos en los diferentes contextos y versiones.

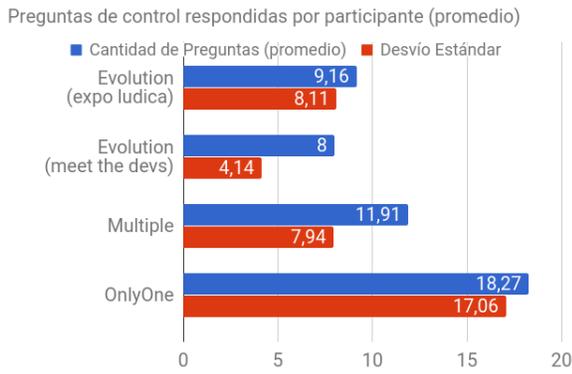


Fig. 4: Promedio de preguntas de control respondidas por participante

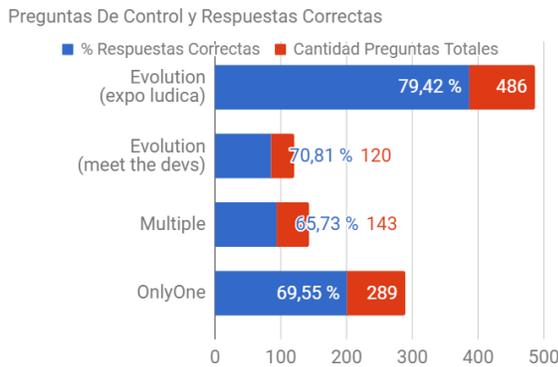
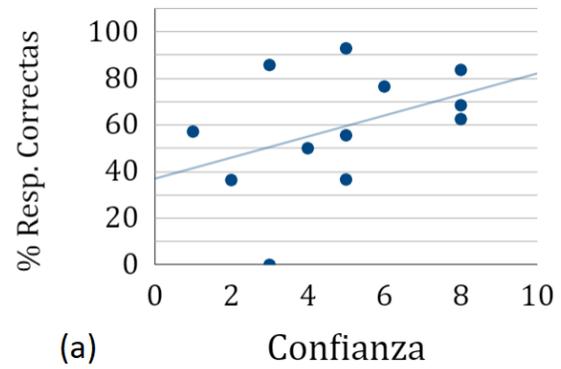


Fig. 5: Preguntas de Control

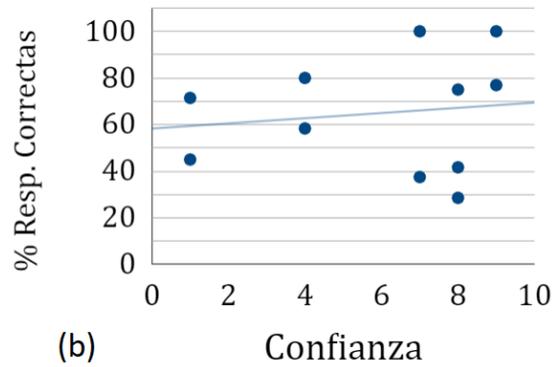
¿Cuál fue el nivel de usabilidad del prototipo? Utilizando las evaluaciones SUS, podemos evaluar el nivel de usabilidad de las distintas versiones del prototipo. Las versiones *OnlyOne* y *Multiple* contaban con un tutorial interactivo semejante a las tareas que el participante debía resolver, sin embargo el diseño de la estética marcó algunos problemas de usabilidad con resultados SUS de 77,18 para la versión *OnlyOne* y 58,75 para *Multiple*. Al mejorar la estética en la versión *Evolution*), el valor de SUS repercutió positivamente ya que obtuvimos 78,5 en la exposición "Meet The Devs" y 78,75 en "Expo Lúdica".

Los participantes que realizaron el tutorial, ¿han adquirido los conocimientos necesarios para la clasificación de imágenes? Observando la Figura 5 podemos observar que en la versión *Evolution* del prototipo se realizaron 486 preguntas de control, obteniendo un 79,42 % de respuestas correctas, podemos decir que los participantes obtuvieron los conocimientos necesarios para detectar *bow shocks*. En los otros prototipos los valores fueron menores en coincidencia con los elementos de mejoras aplicados.

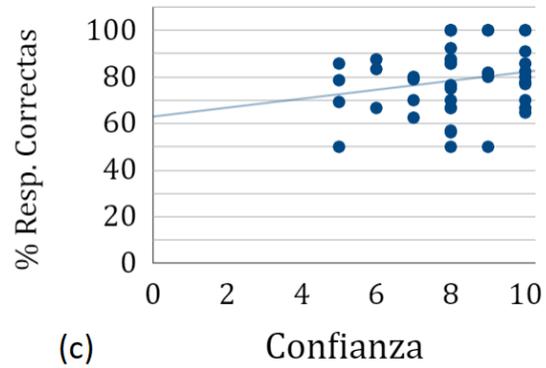
¿El nivel de confianza de los usuarios luego de finalizar el tutorial, se correspondió con el conocimiento real? La Figura 6 muestra una comparativa entre el nivel de confianza y las respuestas correctas de cada participante. Podemos observar a lo largo de las versiones se va generando un cumulo donde



(a)



(b)



(c)

Fig. 6: Relación entre confianza y porcentaje de respuestas correctas. (a) OnlyOne (b) Multiple (c) Evolution

coinciden la confianza con las respuestas correctas. Al igual que en las repuestas correctas, el prototipo *Evolution* es el que mejor desempeño muestra (Alternativa (c) en la Fig. 6).

¿Qué elementos de gamification fueron los más relevantes para los usuarios? Las insignias fueron más valoradas en las pruebas realizadas en doctorado y el planetario, donde podemos caracterizar usuarios con menos cultura de competencia. Por el contrario, en las exposiciones de juegos las tablas de puntuaciones fueron más valoradas. Esto ultimo coincide con un público más acostumbrado a la cultura de la competencia. El perfil de cada participante se mostró diferente en los contextos de evaluación.

Luego de finalizar el tutorial, ¿los ciudadanos continuaron con la clasificación de las imágenes? En términos generales los participantes han continuado con la clasificación.

El número de preguntas respondidas es en promedio de 9,16 (DE 8,11) en el caso del prototipo *Evolution* en la evaluación de "Expo lúdica" y con valores similares en la evaluación de "meet the devs". El resto de los prototipos lo ha superado en cantidad (*Multiple* con 11,91 (DE 7,94) y *OnlyOne* con 18,27 (DE 17,06)). El tipo de participantes de los dos primeros casos parece notar que las sesiones de competencia son intensas y cortas, con necesidad de nuevos desafíos, mientras que en los últimos dos casos las sesiones de participación son más largas. Ambos casos muestran un nivel de participan muy bueno, ya que cada jugador (en el peor de los casos) ayudó en la clasificación de 8 estrellas.

VII. CONCLUSIÓN Y TRABAJOS FUTUROS

En este artículo presentamos una serie de evaluaciones con el fin de analizar la utilización de estrategias de *gamification* en tutoriales para proyectos de ciencia ciudadana para la clasificación de fotos sobre un tipo de estrellas llamadas *Runaway stars*. El artículo presenta tres alternativas de prototipos donde en cada una se analizan diferentes elementos de la aplicación de lógicas de juego.

Las evaluaciones han mostrado que los elementos de gamificación deben ser aplicados de acuerdo al perfil de los usuarios, permitiendo competencia en aquellos usuarios acostumbrados a las lógicas de juegos como así también adecuarse a aquellos participantes menos competitivos, donde las insignias han sido mejor recibidas. Por otra parte, el nivel de correctitud y aprendizaje necesario para la detección de los elementos específicos del proyecto de ciencia ciudadana se han completado con mayor nivel de acierto con aquellos prototipos donde se han desarrollado más los elementos de *gamification*.

En este artículo presentamos un conjunto de desafíos que requieren los participantes de un proyecto de ciencia ciudadana para la clasificación de imágenes de elementos poco cotidianos, esto incluye la variedad de casos existentes, la variedad de representación de la información.

El estudio se desprende de la implementación de tres prototipos evaluados en diferentes oportunidades. Cada prototipo presenta características diferentes sobre la utilización de elementos de *gamification*. En las evaluaciones se analizaron cuales de los elementos de *gamification* fueron más valorados por los usuarios, en qué medida han aprendido lo que el tutorial presentaba, el nivel de satisfacción y usabilidad. Las evaluaciones mostraron que las estéticas, la valoración por puntos, la premiación con insignias y la competencia han sido elementos promisorios en este tipo de tutoriales.

Como trabajo futuro se debe profundizar la aplicación del enfoque propuesto, para poder continuar mejorando en el entrenamiento de los participantes, dentro de los elementos a trabajar podemos incluir mejoras en las pistas sonoras, mejoras en la utilización de textos, y mejoras en la estética en general. Referido a la *gamification* se puede seguir trabajando en la adaptación de las formas de *gamification* de acuerdo a los perfiles de usuario y una proyección de los mismos a las etapas siguientes al tutorial.

Finalmente se debe trabajar para que, luego de finalizada la etapa de entrenamiento el usuario no solo pueda clasificar las imágenes, si no que entienda el fenómeno natural que ocurre en cada una de ellas, esto se debe lograr sin sacrificar demasiado la simpleza y facilidad de uso del entrenamiento.

REFERENCES

- [1] J. Silvertown, "A new dawn for citizen science," *Trends in Ecology & Evolution*, vol. 24, no. 9, pp. 467–471, 2009. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S016953470900175X>
- [2] M. Celasco, J. I. Yañez, R. Gamen, A. Fernández, A. Díaz, and D. Torres, "Galaxy {Conqueror}: {Astronomy}, citizen science and gamification," in *Proceedings - 2016 11th Latin American Conference on Learning Objects and Technology, LACLO 2016*, 2016.
- [3] J. F. Nerbonne and K. C. Nelson, "Volunteer Macroinvertebrate Monitoring: Tensions Among Group Goals, Data Quality, and Outcomes," *Environmental Management*, vol. 42, no. 3, pp. 470–479, 2008. [Online]. Available: <https://doi.org/10.1007/s00267-008-9103-9>
- [4] J. P. Cohn, "Citizen science: Can volunteers do real research?" *BioScience*, vol. 58, no. 3, pp. 192–197, 2008.
- [5] A. W. E. GALLOWAY, M. T. TUDOR, W. M. V. HAEGEN, and West, "The Reliability of Citizen Science: A Case Study of Oregon White Oak Stand Surveys," *Wildlife Society Bulletin*, vol. 34, no. 5, pp. 1425–1429, dec 2006. [Online]. Available: [https://doi.org/10.2193/0091-7648\(2006\)34\[1425:TROCSA\]2.0.CO](https://doi.org/10.2193/0091-7648(2006)34[1425:TROCSA]2.0.CO) <http://0.0.0.2>
- [6] F. Gorleri and F. González Taboas, "ebird: una plataforma integral de ciencia ciudadana para conocer y conservar nuestra avifauna," in *I Workshop de Ciencia Abierta y Ciudadana-Argentina (La Plata, 2017)*, 2017.
- [7] G. Newman, A. Crall, M. Laituri, J. Graham, T. Stohlgren, J. C. Moore, K. Kodrich, and K. A. Holfelder, "Teaching citizen science skills online: Implications for invasive species training programs," *Applied Environmental Education and Communication*, vol. 9, no. 4, pp. 276–286, 2010.
- [8] R. Tinati, M. Luczak-Roesch, E. Simperl, and W. Hall, "Because science is awesome": Studying participation in a citizen science game," in *{WebSci} 2016 - {Proceedings} of the 2016 {ACM} {Web} {Science} {Conference}*, 2016, pp. 45–54.
- [9] R. Tinati, M. LuczakRoesch, E. Simperl, and W. Hall, "An investigation of player motivations in eyewire, a gamified citizen science project," *Computers in Human Behavior*, vol. 73, pp. 527–540, 2017.
- [10] K. M. Kapp, *The gamification of learning and instruction fieldbook: Ideas into practice*. John Wiley & Sons, 2013.
- [11] A. Eveleigh, C. Jennett, S. Lynn, and A. L. Cox, "'I want to be a captain! I want to be a captain!': gamification in the old weather citizen science project," *Proceedings of the First International Conference on Gameful Design, Research, and Applications - Gamification '13*, pp. 79–82, 2013. [Online]. Available: <http://doi.acm.org/10.1145/2583008.2583019>
- [12] A. Blaauw, "On the origin of the o-and b-type stars with high velocities (the "run-away" stars), and some related problems," *Bulletin of the Astronomical Institutes of the Netherlands*, vol. 15, p. 265, 1961.
- [13] C. Peri, P. Benaglia, D. Brookes, I. Stevens, and N. Isequilla, "E-boss: an extensive stellar bow shock survey-i. methods and first catalogue," *Astronomy & Astrophysics*, vol. 538, p. A108, 2012.
- [14] C. Peri, P. Benaglia, and N. Isequilla, "E-boss: An extensive stellar bow shock survey-ii. catalogue second release," *Astronomy & Astrophysics*, vol. 578, p. A45, 2015.
- [15] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: Defining gamification," in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, ser. {MindTrek} '11. New York, NY, USA: ACM, 2011, pp. 9–15. [Online]. Available: <http://doi.acm.org/10.1145/2181037.2181040>
- [16] S. Deterding, A. Canossa, C. Hartevelde, S. Cooper, L. E. Nacke, and J. R. Whitson, "Gamifying research: Strategies, opportunities, challenges, ethics," in *Conference on Human Factors in Computing Systems - Proceedings*, vol. 18, 2015, pp. 2421–2424.
- [17] S. Ayoun, "Applying game design elements in the workplace," in *2015 International Conference on Information Systems*, 2015.
- [18] J. Brooke, "SUS: A quick and dirty usability scale," in *Usability evaluation in industry*, P. W. Jordan, B. Weerdmeester, A. Thomas, and I. L. McLelland, Eds. London: Taylor and Francis, 1996.