

La usabilidad y la escala diferencial de emociones en aplicaciones para Android. Un estudio de caso

AUTORES: Jorge Iván Pincay Ponce¹
Jorge Sergio Herrera Tapia²
Wilian Richart Delgado Muentes³

DIRECCIÓN PARA CORRESPONDENCIA: jorge.pincay@uleam.edu.ec

Fecha de recepción: 23-02-2021

Fecha de aceptación: 12-04-2021

RESUMEN

El objetivo principal de esta investigación fue evaluar los resultados de las escalas de emociones reportadas por las aplicaciones Emotion Detector y Image & Emotion Recognizer, seleccionadas de entre varias existentes para dispositivos Android, así como su usabilidad mediante la reconocida Escala de Usabilidad de un Sistema (SUS). La emoción es una parte integral de la existencia humana, que desempeña un papel importante en la vida cotidiana, por lo que el campo ha sido investigado significativamente en los últimos años, surgiendo diversas escalas discretas en variados contextos. Los resultados obtenidos en esta comparativa muestran cercanías en las escalas diferenciales de emociones reportadas, así como de la usabilidad que tienen las aplicaciones seleccionadas.

PALABRAS CLAVE: computación afectiva; SUS; Escala de usabilidad del sistema; escala diferencial de emociones.

Usability and the differential scale of emotions in Android applications

ABSTRACT

The main objective of this research was to evaluate the results of the emotional scales reported by the Emotion Detector and Image & Emotion Recognizer applications, selected among several existing for Android devices, as well as their usability through the recognized System Usability Scale (SUS). Emotion is an integral part of human existence, which plays an important role in everyday life, so that the field has been researched significantly in recent years, arising discrete scales in various contexts. The results obtained in this comparison show closeness in the differential scales of reported emotions, as well as the usability of the selected applications.

KEYWORDS: Affective computing; SUS; System usability scale; differential scale of emotions.

¹ Ingeniero en Sistemas. Alumno del programa de Doctorado en Informática de la Universidad Nacional de la Plata. La Plata, Buenos Aires, Argentina. Docente titular en la carrera de Ingeniería en Sistemas de la Universidad Laica Eloy Alfaro de Manabí. Manta, Manabí, Ecuador. E-mail: jorge.pincayp@info.unlp.edu.ar.

² Ingeniero en Sistemas. Docente titular en la carrera de Ingeniería en Sistemas de la Universidad Laica Eloy Alfaro de Manabí. Manta, Manabí, Ecuador. E-mail: jorge.herrera@uleam.edu.ec

³ Universidad Laica Eloy Alfaro de Manabí. Manta, Manabí, Ecuador. E-mail: wilian@uleam.edu.ec

INTRODUCCIÓN

La emoción es una parte integral de la existencia humana, que desempeña un papel importante en la vida cotidiana (Izard, 1972). La investigación en el campo de las emociones ha aumentado significativamente en los últimos años, en ese contexto han surgido diferentes trabajos, como la denominada Escala diferencial de emociones (DES), que es un instrumento estandarizado que divide de manera confiable la descripción individual de la experiencia de la emoción en categorías validadas y discretas, en ese momento específico en el que los usuarios responden a un instrumento (All about UX, 2018; Izard, 1972).

Esta investigación tuvo dos objetivos, el primero fue evaluar a dos aplicaciones destinadas al reconocimiento de emociones, seleccionadas de entre varias existentes para dispositivos Android, el aspecto que se evaluó son los resultados que estas reportaban respecto a una emoción que los usuarios intentaban expresar a través de una foto de su rostro, capturada empleando la cámara desde las mismas aplicaciones.

El segundo objetivo fue medir la usabilidad de ambas aplicaciones mediante la Escala de Usabilidad de un Sistema (SUS). La escala en sí consiste en 10 preguntas, cada una de las cuales puede ser puntuada de 1 a 5, donde 1 significa Total desacuerdo y 5 significa Total acuerdo (Brooke, 1996).

DESARROLLO

Un total de 10 personas, estudiantes universitarios, participaron voluntariamente en el estudio sin obtener ninguna retribución por su colaboración. La edad media era de 19,8 años, con una desviación estándar de 0,92 y un rango de edad entre 18 y 21 años. Cinco eran hombres y cinco eran mujeres.

Respecto a las dos aplicaciones utilizadas, ambas disponibles para Android, una de ellas fue Emotion Detector, que es una aplicación de reconocimiento de emociones que toma una expresión facial en una foto como una entrada, y devuelve la valoración a través de un conjunto de emociones para cada cara en una imagen (ver Imagen 1), utilizando técnicas de software alojadas en los servidores de su compañía creadora HMA Labs (HMA Labs, 2017).

Las emociones detectadas por Emotion Detector son la ira, el desprecio, asco, miedo, felicidad, neutral, tristeza y sorpresa; que según HMA Labs, se entienden en distintas culturas y universalmente se pueden determinar con determinadas expresiones faciales (HMA Labs, 2017).

La segunda aplicación utilizada fue Image & Emotion Recognizer, que utiliza la herramientas Microsoft para reconocimiento de imágenes y de emociones (Hazra, 2018). Esta aplicación, a diferencia de la primera, solo muestra como resultado la emoción que tiene la valoración máxima de entre un conjunto de emociones (ver Imagen 1). El conjunto de emociones que reconoce Image & Emotion Recognizer es el mismo que Emotion Detector.

Los participantes evaluaron su emoción individualmente en un salón de clases, previamente el investigador les informaba de las instrucciones del procedimiento y de que el objetivo del estudio era obtener más conocimiento sobre los resultados de las emociones calculadas sobre una misma foto, por dos programas distintos. Los participantes debían relajarse un minuto antes de tomarse la foto, también se les indicaba que no existían respuestas correctas o incorrectas, y que solo debían expresar una emoción y que los programas tratarían de detectarla.

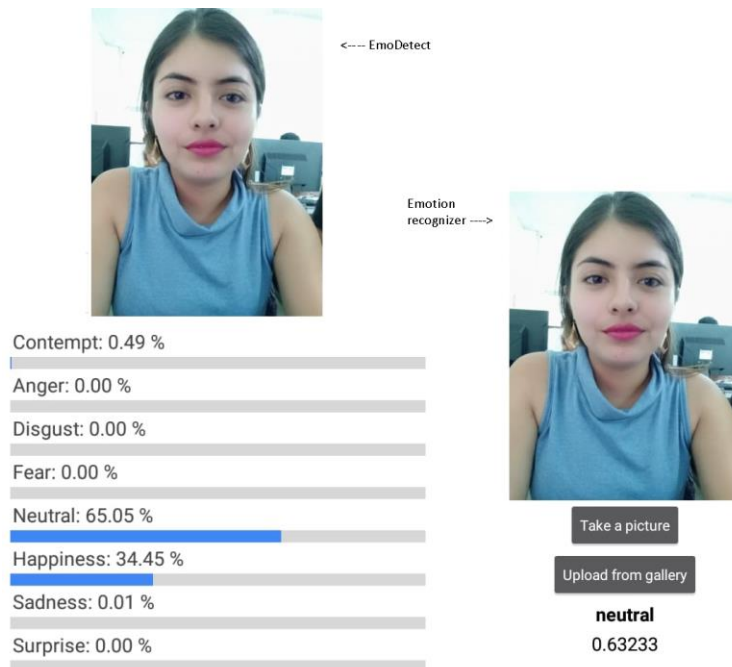


Imagen 1: A la izquierda se ejecuta el reconocimiento de emociones con la aplicación EmoDetect y a la derecha con la aplicación Emotion Recognizer

Una vez que los usuarios utilizaron Emotion Detector, se les pidió que llenen un formulario basado en el Sistema de Escalas de Usabilidad o SUS por sus siglas en inglés (System Usability Scale), que es una herramienta metodológica muy similar a la Escala de Likert y que se usa para medir la usabilidad de una aplicación (Brooke, 1996; UX Español, 2017). El mismo procedimiento se aplicó luego de que utilizaron Image y Emotion Recognizer. Como en total participaron 10 usuarios, se fue alternando la prueba de las aplicaciones entre cada usuario participante.

RESULTADOS

La tabla siguiente muestra los diez conjuntos de resultados para cada participante, tanto con la aplicación Emotion Detector y la aplicación Emotion Recognizer. Para facilitar la lectura se ha agregado la última columna de la tabla, donde se puede determinar fácilmente en qué casos los resultados de Emotion Detector fueron mayores, iguales o menores respecto a Emotion Recognizer.

Según la Tabla 1, las aplicaciones asemejan en gran medida en cuanto a resultados reportados, de hecho, en 7 de 10 ocasiones los resultados fueron iguales, en 3 de 10 casos la diferencia fue de una milésima y sólo en una ocasión el resultado difirió en 2 centésimas. Los resultados se valoraron en un rango de 0 a 1, siendo 1 el máximo valor posible de una emoción.

Tal cual se indicó en la sección de materiales y métodos, posterior a la evaluación de las emociones, se aplicó el Sistema de Escalas de Usabilidad (SUS) para medir la usabilidad de ambas aplicaciones.

Tabla 1: Resultados de las escalas de emociones reportadas para cada participante, con la aplicación Emotion Detector y Emotion Recognizer

		Emotion Detector						Emotion Recognizer				
		Contempt	Anger	Disgust	Fear	Neutral	Happiness	Sadness	Surprise	Happiness	Neutral	
	Gén	Desprecio	Enfado	Disgusto	Temor	Neutral	Felicidad	Tristeza	Sorpresa	Felicidad	Neutral	
P1	H*	0,0001	-	0,0001	0,0001	0,0087	0,9785	0,0125	0,0001	0,9784		>
P2	H	0,0356	0,0037	0,0211	0,0003	0,7426	0,1363	0,0595	0,0010	0	0,7426	=
P3	H	0,0253	0,0401	0,0008	0,0001	0,9097	0,0006	0,0230	0,0005	0	0,9097	=
P4	H	0,0003	-	-	-	0,9992	0,0005	-	-	0,9992		=
P5	H	0,0010	-	-	-	0,2405	0,7585	-	-	0,7585		=
P6	M**	0,0019	-	-	-	0,9934	0,0044	0,0002	-	0	0,9935	<
P7	M	0,0002	-	-	-	0,7345	0,2652	-	-	0	0,7345	=
P8	M	0,0008	-	0,0002	-	0,0310	0,9677	0,0001	0,0001	0,9677	0	=
P9	M	0,0049	-	-	-	0,6505	0,3445	0,0001	-	0	0,6323	>
P10	M	-	-	-	-	-	1,0000	-	-	1,0000		=

* Hombre, ** Mujer

Las 10 preguntas, cada una de las cuales puede ser puntuada de 1 a 5, donde 1 significa Total desacuerdo y 5 significa Total acuerdo, son las siguientes:

1. Creo que usaría esta aplicación frecuentemente
2. Encuentro esta aplicación innecesariamente compleja
3. Creo que la aplicación fue fácil de usar
4. Creo que necesitaría ayuda de una persona con conocimientos técnicos para usar esta aplicación
5. Las funciones de esta aplicación están bien integradas
6. Creo que la aplicación es muy inconsistente
7. Imagino que la mayoría de la gente aprendería a usar esta aplicación en forma muy rápida
8. Encuentro que la aplicación es muy difícil de usar
9. Me siento confiado al usar esta aplicación
10. Necesité aprender muchas cosas antes de ser capaz de usar esta aplicación

En el cuestionario SUS, los resultados se obtienen en escalas de 0 a 100, y se debe evitar modificar las preguntas o su orden, por qué para las preguntas impares (1, 3, 5, 7 y 9), el autor sugiere que estas tomen el valor asignado por el usuario y se le reste 1; en tanto que para las preguntas pares (2, 4, 6, 8, 10), el valor será de 5 menos el valor asignado por los usuarios. El número final obtenido se lo multiplica por 2,5 (Brooke, 1996; UX Español, 2017).

Por ejemplo, si los resultados obtenidos en las preguntas aplicadas a un usuario fueron: 3, 4, 3, 5, 3, 2, 1, 2, 4, 5. Asignamos los nuevos valores según el algoritmo de SUS, y nuestros nuevos

valores serán: $((3-1)+(5-4)+(3-1)+(5-5)+(3-1)+(5-2)+(1-1)+(5-2)+(4-1)+(5-5))*2,5$. Lo que nos resultaría: 40. Dado que el máximo teórico es de 100 puntos, este resultado nos indica que la calidad está por debajo de lo medianamente esperado y debería trabajarse en mejorar dichas aplicaciones.

A continuación, se muestran los resultados obtenidos para cada aplicación, los cuales son muy favorables, pues Emotion Detector alcanza un promedio de 89,8/100 y una desviación estándar de 4,78; en tanto que Emotion Recognizer promedia 91,5/100 y una desviación estándar de 2,68. Estos resultados podrían tener relación con los resultados más detallados que muestra Emotion Detector.

Tabla 2: Escala SUS de cada participante para la aplicación Emotion Detector

Usuario	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	SUS Score
p1	5	2	5	1	5	1	4	1	1	2	82,5
p2	5	2	5	2	5	1	4	1	2	2	82,5
p3	5	2	5	1	5	1	3	1	4	3	85,0
p4	5	1	5	1	5	1	3	1	4	2	90,0
p5	5	2	5	1	5	1	3	1	5	2	90,0
p6	5	1	5	1	5	1	4	1	5	2	95,0
p7	5	2	5	1	5	1	4	1	5	2	92,5
p8	5	2	5	1	5	1	4	1	5	2	92,5
p9	5	2	5	1	5	1	4	1	5	2	92,5
p10	5	1	5	1	5	1	4	1	5	2	95,0
Promedio											89,8

Tabla 3: Escala SUS de cada participante para la aplicación Emotion Recognizer

Usuario	q1	q2	q3	q4	q5	q6	q7	q8	q9	q10	SUS Score
p1	4	2	5	2	5	1	4	1	5	2	87,5
p2	4	2	5	1	5	1	4	1	5	1	92,5
p3	5	2	5	2	5	1	4	1	4	2	87,5
p4	4	2	5	1	5	1	4	1	5	1	92,5
p5	5	2	5	2	5	1	5	1	5	1	95,0
p6	4	2	5	1	5	1	5	1	4	2	90,0
p7	4	2	5	1	5	1	5	1	5	2	92,5
p8	4	2	5	1	5	1	5	1	5	2	92,5
p9	4	1	5	1	5	1	3	1	4	1	90,0
p10	5	2	5	1	5	1	5	1	5	2	95,0
Promedio											91,5

La siguiente es una ilustración de cómo obtener e interpretar los resultados:

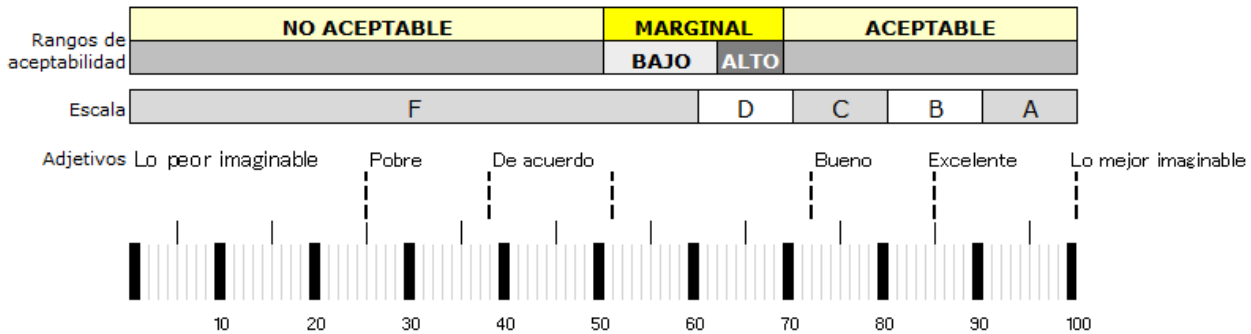


Imagen 2: Clasificación de los puntajes de SUS. Fuente: John Brooke, 2013 (Brooke, 2013)

SUS es una escala muy importante en la industria del software, con referencias en más de 6500 artículos y publicaciones (Brooke, 2018), pese a que no es un tests de diagnóstico, sino que su utilidad está en la clasificación de la facilidad de uso del sitio o aplicación móvil como es el caso de este estudio (usability.gov, 2013); además, proporciona una idea de si los usuarios recomendarían la aplicación a otros usuarios (Brooke, 2013). En el presente caso, y de acuerdo con la

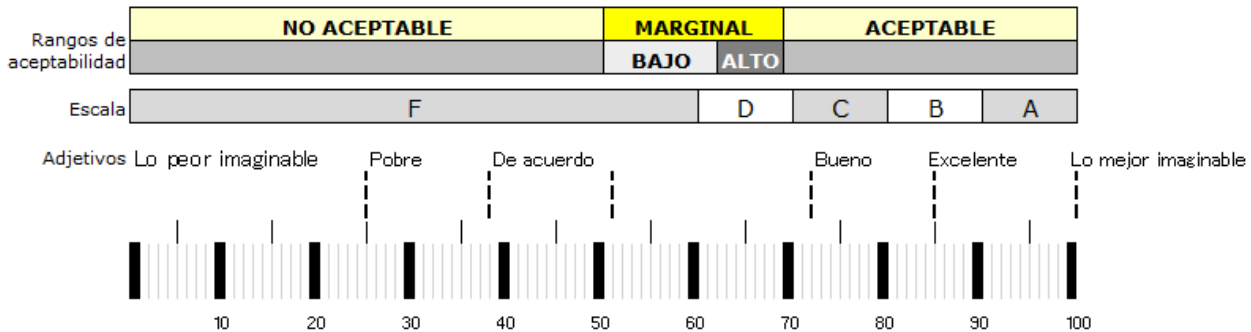


Imagen 2 (Brooke, 2013), los usuarios proporcionaron una valoración buena para Emotion Detector y excelente para Emotion Recognizer. Es de acotar que la diferencia es de apenas 7 décimas.

En lo referente a las escalas emocionales reportadas, sólo es posible realizar la comparativa a partir de los resultados de Emotion Recognizer, pues retorna solo el mayor valor del conjunto de emociones posibles, y ese valor se lo compara con el máximo valor del conjunto que reporta Emotion detector. Los resultados se asemejan tal como se indicó en la sección precedente, en cuanto a Emotion Recognizer sus desarrolladores indican que utilizan los servicios de reconocimiento de emociones de Microsoft, en tanto que HMA Labs no indica que algoritmos usa en Emotion Detector.

Por lo indicado, es de suponer que la diferencia de los resultados radica en los algoritmos que se emplean en las aplicaciones, concretamente para reconocer las emociones, pues en ambas aplicaciones se usó la misma imagen.

CONCLUSIONES

En esta investigación se empleó la escala SUS, misma que resulta bastante neutral respecto a la tecnología, lo que le permite tener vigencia en el tiempo sin rediseñar cuestionarios de valoración

continuamente. La escala reportó que las aplicaciones Emotion Detector y Emotion Recognizer, tienen a criterio de los usuarios considerados en el estudio, en promedio una buena y excelente facilidad de uso respectivamente.

Por otro lado, la escala de emociones diferenciales que reportan ambas aplicaciones se asemeja bastante, y, al ser utilizadas como entradas las mismas imágenes es de esperar que la diferencia esté basada en los algoritmos de reconocimiento de emociones que las aplicaciones incorporan.

REFERENCIAS BIBLIOGRÁFICAS

- All about UX. (2018). Differential Emotions Scale (DES). Retrieved October 5, 2018, from <https://tinyurl.com/y9xmq6df>
- Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194), 4–7.
- Brooke, J. (2013). SUS: a retrospective. *Journal of Usability Studies*, 8(2), 29–40.
- Brooke, J. (2018). Google Scholar - John Brooke. Retrieved October 10, 2018, from <https://tinyurl.com/ybhcadp3>
- Hazra, S. (2018). Image & Emotion Recognizer. Retrieved October 3, 2018, from <https://tinyurl.com/y72qmfor>
- HMA Labs. (2017). Emotion Detector. Retrieved October 2, 2018, from <https://tinyurl.com/y9hk7v7x>
- Izard, C. (1972). Preface. In C. E. B. T.-P. of E. IZARD (Ed.), *Patterns of Emotions* (pp. ix–x). Elsevier. <https://doi.org/10.1016/B978-0-12-377750-8.50003-9>
- usability.gov. (2013). System Usability Scale (SUS). Retrieved October 1, 2018, from <https://tinyurl.com/zmsq8uf>
- UX Español. (2017). Sistema de Escalas de Usabilidad: ¿qué es y para qué sirve? Retrieved October 4, 2018, from <https://tinyurl.com/y8ldmu9k>

