



ESPE
UNIVERSIDAD DE LAS FUERZAS ARMADAS
INNOVACIÓN PARA LA EXCELENCIA

**DEPARTAMENTO DE CIENCIAS DE LA
COMPUTACIÓN**

CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA

**TRABAJO DE TITULACIÓN PREVIO LA OBTENCIÓN DEL
TÍTULO DE INGENIERO EN SISTEMAS E INFORMÁTICA**

**TEMA: REVISIÓN SISTEMÁTICA DE LITERATURA SOBRE
EVALUACIÓN DE LA USABILIDAD DE APLICACIONES EN
DISPOSITIVOS MÓVILES MULTIPLATAFORMA PARA
ESTABLECER EL ESTADO DEL ARTE DE UN PROYECTO DE
INVESTIGACIÓN**

**AUTOR: VELASTEGUI HINOJOSA SANDRA
VERÓNICA**

DIRECTOR: ING. FONSECA RODRIGO, PhD

SANGOLQUÍ

2016

CERTIFICADO



DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA

CERTIFICACIÓN

Certifico que el trabajo de titulación, "Revisión sistemática de Literatura sobre Evaluación de la Usabilidad de Aplicaciones Móviles Multiplataforma para establecer el estado del arte de un proyecto de Investigación" realizado por la señorita Sandra Verónica Velastegui Hinojosa, ha sido revisado en su totalidad y analizado por el software anti-plagio, el mismo cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad de Fuerzas Armadas ESPE, por lo tanto me permito acreditarlo y autorizar a la señorita Sandra Verónica Velastegui Hinojosa para que sustente públicamente.

Quito, 02 de diciembre del 2016

Ing. Rodrigo Fonseca, PhD.

DIRECTOR

AUTORÍA DE RESPONSABILIDAD



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

AUTORÍA DE RESPONSABILIDAD

Yo, Sandra Verónica Velastegui Hinojosa, con cédula de identidad N° 1723959027, declaro que este trabajo de titulación "Revisión sistemática de Literatura sobre Evaluación de la Usabilidad de Aplicaciones Móviles Multiplataforma para establecer el estado del arte de un proyecto de Investigación", ha sido desarrollado considerando los métodos de investigación existentes, así como también se ha respetado los derechos intelectuales de terceros considerándose en las citas bibliográficas.

Consecuentemente declaro que este trabajo es de mi autoría, en virtud de ello me declaro responsable del contenido, veracidad y alcance de la investigación mencionada.

Quito, 02 de diciembre del 2016

Sandra Verónica Velastegui Hinojosa

C.C 1723959027

AUTORIZACIÓN



**DEPARTAMENTO DE CIENCIAS DE LA COMPUTACIÓN
CARRERA DE INGENIERÍA EN SISTEMAS E INFORMÁTICA**

AUTORIZACIÓN

Yo, Sandra Verónica Velastegui Hinojosa, autorizo a la Universidad de las Fuerzas Armadas ESPE publicar en la biblioteca Virtual de la institución el presente trabajo de titulación "Revisión sistemática de Literatura sobre Evaluación de la Usabilidad de Aplicaciones Móviles Multiplataforma para establecer el estado del arte de un proyecto de investigación", cuyo contenido, ideas y criterios son de mi autoría y responsabilidad.

Quito, 02 de diciembre del 2016

Sandra Verónica Velastegui Hinojosa

C.C. 1723959027

DEDICATORIA

La presente tesis se la dedico a Dios, quien supo guiarme y darme fuerzas para salir adelante y no desmayar frente a las adversidades que se presentaban en el camino de mi formación profesional.

A mis adorados padres, que son pilar fundamental en mi vida, por darme su apoyo incondicional. A mi madre Rosario Hinojosa, por su amor, sus consejos, y valores que me ha inculcado y han hecho de mí una persona de bien. A mi padre Julio Velastegui, por ser mi ejemplo de constancia y perseverancia, su amor me ha ayudado a salir siempre adelante.

A mis amados hermanos Julio y Gabriela, por siempre estar pendientes de mí, y por estar presentes en cada momento difícil y de alegría en mi vida. A mis sobrinas Valentina, Samantha y Agatha, quienes en mi vida son una fuente de motivación, inspiración y felicidad.

AGRADECIMIENTO

Quiero agradecer primero a Dios, por darme la vida y una familia maravillosa que siempre han estado a mi lado apoyándome, en cada reto que me he propuesto.

A mis amados padres por su apoyo incondicional, por su ejemplo de lucha contante gracias por su amor, dedicación y paciencia. A mi padre gracias por confiar y creer en mí siempre, y por anhelar lo mejor para mi vida. A mi madre por acompañarme en cada noche agotadora de estudio desde que me propuse a estudiar la carrera de ingeniería.

A mis hermanos, por cada uno de sus consejos, por sus palabras de aliento que me guiaron durante toda mi vida.

A mi director de tesis, Ing. Rodrigo Fonseca, PhD, por su paciencia y esfuerzo, por guiarme siempre al cumplimiento de los objetivos propuestos, todos sus conocimientos, sus orientaciones y su motivación han sido fundamentales para mi formación como investigadora.

A mi querida universidad que me ayudo con el acceso a las diferentes bibliotecas digitales, para la obtención de toda la información pertinente para el desarrollo del presente trabajo investigativo.

Gracias a todas esas personas que, de una manera u otra, han sido claves en mi vida, familiares, maestros y amigos, quienes me acompañaron en el proceso de cumplimiento de este reto.

Gracias a la vida por este nuevo triunfo, que es un escalón más en mi vida profesional, gracias a todas las personas que apoyaron y creyeron en la realización de esta tesis.

ÍNDICE

CERTIFICADO	ii
AUTORÍA DE RESPONSABILIDAD.....	iii
AUTORIZACIÓN.....	iv
DEDICATORIA.....	v
AGRADECIMIENTO	vi
RESUMEN.....	x
CAPÍTULO I.....	1
INTRODUCCIÓN	1
1.1 Antecedentes.....	1
1.2 Problemática.....	3
1.3 Justificación.....	5
1.4 Objetivos.....	6
1.4.1 Objetivo General	6
1.4.2 Objetivos Específicos.....	7
1.5 Alcance.....	7
CAPÍTULO II	9
METODOLOGÍA Y MARCO TEÓRICO	9
2.1 Metodología.....	9
2.2 Marco Teórico	10
2.2.1 Antecedentes Investigativos	10
2.2.2 Fundamentación Filosófica.....	11
2.2.3 Red de Categorías	12
2.2.4 Fundamentación Científica de la Variable Independiente.....	13
2.2.5 Fundamentación Científica de la Variable Dependiente	17
2.2.6 Hipótesis	23
CAPÍTULO III.....	24

PLANIFICACIÓN DE LA REVISIÓN SISTEMÁTICA LITERARIA	24
3.1 Necesidad de la Investigación	24
3.2 Definición de las Preguntas de Investigación.....	25
3.3 Definición de la Estrategia de Búsqueda.....	26
3.4 Definición de los criterios de inclusión y exclusión.....	29
CAPÍTULO IV	31
EJECUCIÓN DE LA REVISIÓN SISTEMÁTICA LITERARIA.....	31
4.1 Depuración de Estudios Candidatos.....	31
4.2 Selección de Estudios	32
4.3 Selección de Estudios Primarios	33
CAPÍTULO V.....	34
EXTRACCIÓN Y SÍNTESIS DE RESULTADOS.....	34
5.1 Extracción de Resultados	34
5.2 Resultados de la SLR.....	70
5.2.1 Resultados Generales de la SLR Sobre Repositorios	70
5.2.2 Resultados Específicos de la SLR Sobre Repositorios.....	73
5.3 Síntesis de Resultados de la SLR	107
CAPÍTULO VI.....	112
CONCLUSIONES, RECOMENDACIONES Y FUTURAS LÍNEAS DE	
INVESTIGACIÓN.....	112
6.1 Conclusiones.....	112
6.2 Recomendaciones	116
6.3 Futuras Líneas de Investigación	117
Referencias.....	117

ÍNDICE DE TABLAS

Tabla 1. Estudios por Cadena de Búsqueda	27
--	----

Tabla 2 Selección de Estudios Primarios	33
Tabla 3 Fuentes de Publicación.....	71
Tabla 4 Modelos de Usabilidad, Atributos y métricas.....	84
Tabla 5 Atributos y métricas usados para medir la usabilidad en los estudios primarios	91
Tabla 6 Factores de usabilidad considerados en cada estudio primario.....	101
Tabla 7 Factores de usabilidad tomados en cuenta en el desarrollo de aplicaciones móviles de los estudios primarios	104

ÍNDICE DE FIGURAS

Figura 1. Red de Categorías de las Variables de Investigación	13
Figura 2. Etapas del desarrollo de software	18
Figura 3. Framework de usabilidad (ISO 9241-11)	21
Figura 4. Proceso de Estrategia de Búsqueda de Estudios.....	26
Figura 5. Proceso de Selección de Estudios Primarios	31
Figura 6. Distribución de Estudios Candidatos, Seleccionados y Primarios	33
Figura 7. Proceso de Extracción de Datos	34
Figura 8. Atributos de Usabilidad del Modelo PACMAD.....	38
Figura 9. Número de estudios primarios por año.....	73
Figura 10. Modelos de Usabilidad usados en los Estudios Primarios	83
Figura 11. Tipos de Aplicaciones evaluadas.....	93
Figura 12. Síntesis de Resultados de los Estudios Primarios.....	108

RESUMEN

Antecedentes: El uso de dispositivos móviles ha crecido sustancialmente en los últimos tiempos, principalmente para permitir a los usuarios acceder a diferentes servicios y realizar sus transacciones desde cualquier sitio. Desde esta perspectiva, la usabilidad de las aplicaciones móviles juega un papel fundamental. Sin embargo, las aplicaciones móviles presentan aún dificultades en este ámbito. Por tanto, se hace necesario considerar a la usabilidad como un factor primordial en el proceso de desarrollo de aplicaciones móviles. **Objetivo:** Investigar los métodos, técnicas y modelos que son usados para evaluar la usabilidad de las aplicaciones móviles multiplataforma. **Metodología:** Se realizó una revisión sistemática de literatura basada en las directrices de (Kitchenham & Charters, 2007). **Resultados:** La revisión de 42 estudios primarios permitió identificar aspectos relevantes sobre el proceso de evaluación de usabilidad de las aplicaciones móviles multiplataforma tales como: los métodos y modelos de evaluación tradicionales deben ser adaptados o combinados para abordar el enfoque móvil y no siempre son aplicados desde etapas tempranas del ciclo de vida del producto software. **Conclusiones:** Existen muchos métodos, técnicas y modelos para evaluar la usabilidad, sin embargo, ninguno cubre por completo los factores de usabilidad y las características que abarca el contexto móvil. Además, su aplicación muchas veces se realiza en etapas tardías del ciclo de vida de la aplicación. Esto se debe a que existe una falta de procesos iterativos de desarrollo-evaluación en etapas iniciales del ciclo de vida del producto software. La plataforma móvil es una característica influyente que debería ser considerada en el proceso de evaluación de las aplicaciones móviles multiplataforma. Finalmente la adopción de mecanismos automatizados para la evaluación de usabilidad requiere ser más investigada, para combinar más de un método y que integre un enfoque multiplataforma.

PALABRAS CLAVES:

- **USABILIDAD**
- **EVALUACIÓN DE USABILIDAD**
- **DISEÑO DE INTERACCIÓN**
- **APLICACIONES MÓVILES MULTIPLATAFORMA**
- **REVISIÓN SISTEMÁTICA DE LITERATURA**

ABSTRACT

Background: The use of the mobile devices has grown substantially in the recent years, mainly to allow the users to access different services and realize their transactions from anywhere. From this perspective, the usability of the mobile applications plays a fundamental. However, the mobile applications still present many difficulties in this aspect. Therefore, it is necessary to consider the usability as an important factor for the development of mobile applications. **Goal:** Research the methods, techniques or models that are used for evaluate the usability of the multiplatform mobile applications. **Method:** A systematic literature review was conducted based on the (Kitchenham & Charters, 2007) guidelines. **Results:** The review of 42 primary studies allowed us to identify relevant aspects about the current situation of the usability evaluation process of the mobile applications such as: the traditional evaluation methods or models must be adapted or combined to involve the mobile approach and not always there are applied since early stages of the software product life cycle; the existence of automated evaluation methods is limited, especially in a multiplatform approach; the specific usability factors are not always considered by developers and evaluators. **Conclusions:** There are many methods, techniques and models to evaluate usability, however, none of them completely covers the usability factors and the characteristics of the mobile context. In addition, its application is often done in later stages of the application life cycle. This is because exist a lack of iterative development-evaluation processes from the early stages of the software product lifecycle. The mobile platform is a influential characteristic and should be considered in the evaluation and design process. Finally, the adoption of automated mechanisms for the evaluation of usability requires to be more investigated, for the combination of more than one method and to integrate a multiplatform approach.

KEYWORDS:

- **USABILITY**
- **USABILITY EVALUATION**
- **INTERACTION DESIGN**
- **MULTIPLATFORM MOBILE APPLICATION**
- **SYSTEMATIC LITERATURE REVIEW**

CAPÍTULO I

INTRODUCCIÓN

1.1 Antecedentes

En los últimos años el crecimiento y uso de dispositivos móviles ha cambiado los hábitos de comunicación y acceso a la información de las personas (Di Giovanni, y otros, 2012). Estos dispositivos por medio de las aplicaciones móviles apoyan las actividades de la vida diaria de los usuarios (Di Giovanni, y otros, 2012). La gente utiliza los dispositivos móviles y aplicaciones para diferentes actividades, tales como, lectura, escuchar música, interactuar en las redes sociales, jugar, etc. (Angulo Cevallos, 2014).

Análogamente a las aplicaciones de escritorio, en las aplicaciones móviles el usuario espera que el producto software reúna ciertos criterios de calidad que satisfagan en gran medida sus necesidades (Enriquez & Casas, 2013). Dentro de la Ingeniería de Software, una de las primordiales características que debe tener un producto software para que sea exitoso entre los usuarios, es que el mismo sea de calidad, por lo que para los desarrolladores de software resulta relevante poder medir esa calidad (Enriquez & Casas, 2013).

La calidad de software es definida por la IEEE, Std. 610-1990, como “el grado con el que un sistema, componente o proceso cumple los requerimientos especificados y las necesidades o expectativas del cliente o usuario” (López, Cabrera, & Luz, 2008). Entonces, especificar la calidad de software para un sistema, es similar a determinar una lista de atributos de calidad de software requeridos para el sistema (Enriquez & Casas, 2013). Según el modelo de calidad ISO/IEC 9126:2001 la calidad del software está especificada por dos tipos de calidad: interna y externa. La calidad externa evalúa que el producto software satisfaga las necesidades de los usuarios, mientras que, la calidad interna evalúa las características intrínsecas o el total de los atributos que un producto software debe satisfacer. Este modelo de calidad externa e interna está constituido por las siguientes características: (1) Funcionalidad, (2) Confiabilidad, (3) Facilidad de uso, (4) Eficiencia, (5) Facilidad de Mantenimiento, (6) Portabilidad (Scalone, 2006).

En este trabajo en particular, se estudia la facilidad de uso o usabilidad, dado que esta característica está cobrando cada vez más importancia en el desarrollo de software (Ferré, 2000), específicamente en el de aplicaciones móviles, ya que se presentan

diversos desafíos (Di Giovanni, y otros, 2012), uno de ellos es mejorar la usabilidad y experiencia de los usuarios al usar las aplicaciones móviles (Enriquez & Casas, 2013).

La usabilidad de un sistema es un atributo de calidad que debería ser considerado como básico al momento de construir software y está asociado tanto a los usuarios, como a las necesidades y condiciones específicas (Ferré, 2000). Por tanto, la usabilidad no puede especificarse indistintamente del entorno de uso y los usuarios que vayan a utilizar el sistema (Ferré, 2000).

A menudo se relaciona a la usabilidad con las particularidades de los elementos de una interfaz gráfica de usuario basada en colores, ventanas, el diseño gráfico de sus animaciones e íconos (Ferré, 2000). No obstante, la usabilidad no solo está relacionado con la interfaz gráfica de usuario (Ferré, 2000). La usabilidad de un producto software también está ligada primordialmente a la interacción del mismo, es decir a la forma en que se lleva a cabo las operaciones con el sistema (Ferré, Juristo, Windl, & Constantine, 2001). Esta interacción no está solamente definida en la interfaz gráfica, sino que está incorporada en el código que efectúa las funcionalidades del sistema (Ferré, Juristo, Windl, & Constantine, 2001).

“La Ingeniería de la Usabilidad es una disciplina que proporciona métodos estructurados para lograr la usabilidad en el diseño de interfaces del usuario, durante el desarrollo del producto software” (Scholtz, 2004). La evaluación de la usabilidad es parte principal de este proceso (Scholtz, 2004). Existen técnicas para evaluar la usabilidad que permiten desarrollar sistemas más usables y alcanzar el nivel de facilidad de uso requerido por el usuario (Ferré, 2000).

La razón principal para la aplicación de las técnicas de usabilidad en el desarrollo de un producto software, es incrementar la eficiencia y satisfacción del usuario. De manera que, estas técnicas puedan permitir que los sistemas alcancen los objetivos del usuario y también ayudar a que los usuarios realicen sus tareas de manera eficiente (Ferré, Juristo, Windl, & Constantine, 2001).

La medición de la usabilidad de aplicaciones móviles se ha convertido en un tema de interés e investigación (Rodríguez, Sierra, & Jaramillo, 2015), y el presente trabajo busca conocer a fondo la situación actual sobre este tema, por lo que es preciso indagar en la literatura para obtener una cantidad de información adecuada sobre este tema y en especial que proceda de fuentes fiables. De manera que esta investigación

contribuya a brindar una visión completa, integral y válida de algunas pruebas existentes sobre la evaluación de usabilidad de aplicaciones en dispositivos móviles multiplataforma (Wohlin, Runeson, Höst, & Ohlsson, 2012).

1.2 Problemática

En el estándar ISO 9241 se define a la usabilidad como el nivel en el que un producto software puede ser usado por usuarios determinados para lograr objetivos concretos con efectividad, eficiencia y satisfacción en un contexto de uso definido (Ferré, 2000). En el producto software, a esta característica se la ha considerado como un atributo de calidad; por tanto, al momento de construir software debería ser considerada con la importancia que esta implica (Ferré, 2000).

Los factores principales que determinan la usabilidad de un producto software, corresponden a condiciones específicas, necesidades y los usuarios puntuales que utilizarán el sistema (Ferré, Juristo, Windl, & Constantine, 2001). Por tal motivo, estos factores deberían ser catalogados como artífices en la definición de las características más adecuadas para que un producto sea considerado como amigable y fácil de utilizar (Ferré, Juristo, Windl, & Constantine, 2001).

Con la proliferación del uso de los dispositivos móviles, la medición de usabilidad se ha tornado en una necesidad y por ende en un tema de investigación, dado que no están definidos aún estándares ni modelos de manera consistente, tal que, perfilen una guía práctica sobre los atributos, métricas y reglas necesarias para medir de forma adecuada la usabilidad de las aplicaciones móviles (Rodríguez, Sierra, & Jaramillo, 2015).

En la actualidad las directrices y métodos existentes fueron creados para el análisis de usabilidad de aplicaciones de escritorio tradicionales, lo que dificulta su adopción en el desarrollo de aplicaciones móviles (Rodríguez, Sierra, & Jaramillo, 2015) por las características específicas que poseen este tipo de dispositivos por ejemplo: las dimensiones reducidas, la capacidad de almacenamiento y procesamiento, el ancho de banda limitado, su sistema operativo que tiene sus particularidades en cuanto al manejo de la interfaz de usuario y su contexto de uso (Enriquez & Casas, 2013). Todas estas características únicas de los dispositivos móviles, tienen una influencia en la usabilidad de un producto software, por lo que es importante que sean tomadas en cuenta en el

proceso de desarrollo y evaluación de usabilidad de la aplicación móvil (Enriquez & Casas, 2013).

Además las técnicas de usabilidad de la IPO (Interacción Persona Ordenador), que permiten conseguir un nivel apropiado de usabilidad en el producto software, se aplican en el marco de métodos IPO, y no en el proceso de desarrollo (Ferré, 2005).

Es decir, que los desarrolladores a menudo aplican las técnicas de evaluación de usabilidad en las últimas etapas de desarrollo del producto software, donde los problemas de usabilidad llegan a ser muy costosos y muchas veces imposibles de arreglar (Ferré, Juristo, Windl, & Constantine, 2001).

Por otro lado, las pruebas de usabilidad por sí solas no son suficientes para la producción de un sistema muy fácil de utilizar, ya que estas pruebas no solucionan los problemas de diseño de interfaz (Ferré, Juristo, Windl, & Constantine, 2001).

Una de las causas de los problemas de diseño de interfaz es que, los desarrolladores al momento de crear la interfaz de usuario, se enfrentan a un dilema entre seguir la filosofía de interacción propia de cada plataforma y la creación de un diseño común de interacción multiplataforma (Barea, Ferré, & Villarroel, 2013). En la actualidad existen solo directrices genéricas para el diseño de interacción, las cuales no cubren estos aspectos (Barea, Ferré, & Villarroel, 2013), los mismos que están relacionados con la usabilidad. Por este motivo, es importante considerar cuidadosamente la interacción no solo en el diseño de la parte visible de la interfaz de usuario, sino también en el diseño del resto del sistema (Barea, Ferré, & Villarroel, 2013).

En conclusión, un buen diseño de la interfaz de usuario podría hacer que un producto software aumente su nivel de usabilidad, pero un sistema con un diseño de interacción pobre no logra mejorar su nivel de usabilidad solo cambiando la interfaz de usuario (Ferré, Juristo, Windl, & Constantine, 2001). Se evidencia que no es posible diseñar la lógica de negocio independientemente de la interacción y después unirlos (Ferré, Juristo, Windl, & Constantine, 2001). La relación que existe entre el diseño de la interacción y el diseño de software habitual, obliga a modificar este último para incorporar al primero (Ferré, 2000). Por tanto, es necesario prestar mucha atención a la usabilidad del sistema desde el inicio del desarrollo, y aplicar técnicas de usabilidad por ejemplo, evaluaciones heurísticas, inspecciones formales, cuestionarios, etc.,

desde las etapas iniciales del proceso de desarrollo (Ferré, Juristo, Windl, & Constantine, 2001).

1.3 Justificación

En la actualidad el mercado ofrece muchas opciones diferentes de aplicaciones de telefonía móvil (Humayoun, Ehrhart, & Ebert, 2013). Uno de los principales objetivos del mercado es la creación de aplicaciones que se puedan utilizar y funcionar perfectamente en las diferentes plataformas que existen actualmente en el mundo de dispositivos móviles (Humayoun, Ehrhart, & Ebert, 2013). De hecho, la selección de una tecnología adecuada y un adecuado marco para desarrollar la aplicación móvil son los retos a los que todos los equipos de desarrollo el mundo se enfrentan en este momento (Humayoun, Ehrhart, & Ebert, 2013).

Desde el punto de vista de la Ingeniería de Software, la usabilidad de las aplicaciones de software en los dispositivos móviles es uno de los factores importantes de calidad debido a esto la importancia de evaluar técnicas y metodologías para medir y analizar la usabilidad de las aplicaciones (Ferré, 2000). Otra razón principal para tomar en cuenta la aplicación de técnicas de usabilidad en el desarrollo de un producto software, es incrementar la eficiencia y satisfacción de los usuarios, por ende la productividad del sistema (Ferré, Juristo, Windl, & Constantine, 2001).

Por este motivo esta investigación se torna útil para los desarrolladores de software, para que al construir un producto software empleen adecuadamente estas técnicas o modelos, en particular para las evaluaciones de la usabilidad, y que de esta manera las aplicaciones sean de mejor calidad y los usuarios no tengan dificultad al usarlas (Ferré, 2000).

La usabilidad también es un factor significativo para la aceptación del sistema por el usuario, ya que si los usuarios no piensan que el sistema les ayudará a realizar sus tareas, ellos son menos propensos a aceptarlo (Ferré, Juristo, Windl, & Constantine, 2001). Es posible que no utilizarán el sistema en absoluto o lo utilizarán de manera ineficiente después del despliegue (Ferré, Juristo, Windl, & Constantine, 2001). De manera que, si un producto software no apoya a las tareas del usuario, no cumple con las necesidades del mismo, que es el principal objetivo que se busca en la construcción de software (Ferré, Juristo, Windl, & Constantine, 2001).

Por lo tanto las técnicas de usabilidad pueden ayudar a que el producto software alcance los objetivos del usuario, y lo apoye al realizar sus tareas. Además una buena usabilidad está ganando importancia en un mundo en que los usuarios tienen menos conocimientos de informática y no pueden permitirse el lujo de pasar mucho tiempo aprendiendo como funciona un sistema (Ferré, Juristo, Windl, & Constantine, 2001).

Por otro lado, la parte visual y de interacción con el usuario, son considerados cruciales tanto al usar aplicaciones de software, como al desarrollarlas. Por ende, con un diseño adecuado de las mismas, tomando en cuenta la usabilidad, se puede lograr que el nivel de calidad del producto software se incremente (Enriquez & Casas, 2013). También, es importante que el desarrollador realice una planificación y diseño previo a la implementación de este tipo de aplicaciones, debido a que entre los dispositivos móviles existe una heterogeneidad (Enriquez & Casas, 2013).

Por todas las implicaciones anteriormente mencionadas que ocurren al evaluar la usabilidad en aplicaciones móviles multiplataforma, es importante reunir información relevante y actual sobre este tema, y que proceda de fuentes confiables, por esto se ha pensado en una Revisión Sistemática de Literatura (SLR de los siglos en inglés Systematic Literature Review), la cual permite identificar, analizar e interpretar los datos disponibles en relación a investigaciones reportadas en las comunidades científicas (Kitchenham & Charters, 2007), para establecer el estado del arte sobre el tema y de esta manera contribuir para que en estudios posteriores se puedan armonizar los diseños de interfaz de usuario de las aplicaciones móviles multiplataforma, también para que los resultados contribuyan a los desarrolladores para el conocimiento y comprensión sobre la correcta aplicación de técnicas y modelos para medir e incrementar la usabilidad de las aplicaciones en dispositivos móviles multiplataforma.

1.4 Objetivos

1.4.1 Objetivo General

Realizar una revisión sistemática de literatura sobre evaluación de la usabilidad de aplicaciones en dispositivos móviles multiplataforma, para conocer a fondo la información relevante y la problemática actual referente a este tema y que proceda de fuentes fiables, aplicando guías de revisión sistemática literaria basadas en Kitchenham.

1.4.2 Objetivos Específicos

- Plantear una estrategia de búsqueda basada en el objetivo que plantea la SLR, tal que permita buscar y obtener los estudios candidatos de la investigación.
- Seleccionar un conjunto de estudios primarios, que responden a las características que plantea la SLR, para ubicar las fuentes principales de información, que contribuirán a dar respuesta a las preguntas de investigación.
- Extraer información relevante de los estudios primarios, para dar respuesta a las preguntas de investigación planteadas en la SLR, realizando una síntesis de resultados de las características principales y específicas de los estudios primarios.

1.5 Alcance

Se realizará una revisión sistemática literaria sobre la evaluación de la usabilidad en dispositivos móviles multiplataforma para identificar estudios que reporten resultados sobre la temática en cuestión y caracterizar cada estudio en función de los elementos de usabilidad. La SLR incluye las siguientes fases:

1. Planificación de la SLR

1.1 Establecimiento de la necesidad de la investigación

1.2 Definición de las Preguntas de investigación

2. Realización de la estrategia de búsqueda

2.1 Conformación del grupo de control: se constituye de un grupo de estudios que responden a las características que plantea la SLR, y facilitan la ubicación de palabras claves para definir las cadenas de búsqueda y garantiza la cobertura de la literatura.

2.2 Construcción de la cadena de búsqueda: inicia con el análisis de cada estudio del grupo de control a nivel de: título, resumen y palabras claves, para la conformación de las cadenas de búsqueda, utilizando los términos OR, para añadir sinónimos y AND para añadir nuevos términos.

2.3 Búsqueda de estudios candidatos: es la configuración de la cadena de búsqueda en cada base digital seleccionada, donde se especifica: fuentes de publicación, idioma, motores de búsqueda y método de búsqueda.

3. **Ejecución de la SLR:** una SLR ubica la información específica para contestar a las preguntas de investigación.
 - 3.1 Criterios de Inclusión y Exclusión.
 - 3.2 Proceso de selección de estudios primarios
 - 3.2.1 Depuración de estudios candidatos: es el refinamiento de los estudios obtenidos de cada base digital para ubicar posibles estudios duplicados.
 - 3.2.2 Selección de estudios: selección de estudios por cada grupo de estudios candidatos, esta selección primero será individual y luego se hará una validación cruzada.
 - 3.2.3 Selección de estudios primarios: es la consolidación de un listado de estudios seleccionados, donde se realizará un estudio de estos artículos, para que aplicando los criterios de inclusión y exclusión, se extraigan los estudios primarios.
4. **Extracción de datos:** los resultados de una SLR se sustentan en la información de los estudios primarios.
 - 4.1 Extracción de las características de los estudios primarios: extracción a nivel de características generales y específicas.
 - 4.2 Extracción de información relevante de cada estudio primario para dar respuesta a las preguntas de investigación planteadas: análisis más detallado de las características de cada estudio, para abstraer las respuestas a las preguntas de investigación.

Para el cumplimiento de los objetivos y el alcance de esta Revisión Sistemática de Literatura, se requiere una metodología experimental que será el medio que permitirá orientarnos hacia la solución del problema con la ayuda de herramientas y también una fundamentación teórica que dará un conjunto de conceptos que permiten sustentar y abordar el problema de la presente tesis. En el siguiente capítulo se detalla más concretamente la metodología y marco teórico usado para la Revisión Sistemática de Literatura.

CAPÍTULO II

METODOLOGÍA Y MARCO TEÓRICO

En este capítulo se especifica la metodología a ser aplicada en la investigación, considerando la aplicabilidad de los métodos y técnicas; luego, se realiza la descripción del marco teórico, basado en una red de categorías que respalda la presente investigación.

2.1 Metodología

Esta investigación requiere el uso de un método que la conduzca al conocimiento. La metodología a utilizar es la deductiva cualitativa con una investigación exploratoria, la cual se efectuará por medio de una revisión sistemática de literatura.

Método Deductivo: Es un método científico que va de lo general a lo particular ya que pretende desarrollar una teoría empezando por formular puntos de partida y deduciendo luego sus consecuencias (Chung, 2008). El método deductivo parte de datos generales aceptados como valederos, para deducir a través del razonamiento lógico y estadístico varias suposiciones, es decir, se partirá de un argumento o de la formulación de una hipótesis, en donde para comprobar su validez se compara con la información existente (Chung, 2008).

Para la presente tesis se requiere una recopilación importante de información con datos reales, aceptados como valederos, sobre evaluación de usabilidad de aplicaciones multiplataforma en dispositivos móviles, para que a través de un razonamiento lógico se logre generar un reporte con respecto a la información existente.

Investigación Cualitativa: Tiene como principal objetivo la recopilación y descripción de las cualidades de un fenómeno a partir de indagaciones (Chung, 2008), en el caso de la presente tesis este método es necesario, porque se va a realizar una descripción y un estudio minucioso de los datos encontrados sobre la evaluación de usabilidad de aplicaciones multiplataforma en dispositivos móviles.

Investigación exploratoria: La investigación exploratoria tiene por objetivo el estudio o examen de un problema para facilitar una investigación y aumentar el conocimiento sobre determinado fenómeno poco estudiado (Cazau, 2006). Para la presente investigación se requiere diagnosticar la situación del uso de técnicas de evaluación de usabilidad de aplicaciones multiplataforma en dispositivos móviles.

Revisión Sistemática de Literatura: Una revisión sistemática de literatura es una metodología para identificar, evaluar e interpretar las investigaciones relevantes y disponibles sobre una pregunta de investigación particular, área temática o fenómeno de interés (Kitchenham & Charters, 2007).

La SLR cubre las metodologías anteriormente mencionadas: la deductiva, cualitativa y exploratoria. La metodología deductiva, porque la SLR permite identificar vacíos del conocimiento, en donde se realiza una revisión minuciosa de un volumen de estudios relevantes, para la deducción de las respuestas a las preguntas de investigación planteadas, a través de un análisis y síntesis de estudios considerados como primarios. También cubre la metodología cualitativa, ya que permite extraer las características generales y específicas de los estudios primarios, y por último también la SLR envuelve la investigación exploratoria porque se requiere un proceso apropiado de búsqueda y selección de estudios candidatos y primarios.

Para la presente tesis se realizará una revisión sistemática literaria sobre la evaluación de la usabilidad de aplicaciones multiplataforma en dispositivos móviles, para identificar estudios primarios, relevantes y fiables, que reporten resultados sobre este tema, la SLR se fundamentará en las siguientes directrices, que se estructuran en un proceso que incluye las siguientes fases:

1. Planteamiento de las preguntas de investigación.
2. Establecimiento de la estrategia de búsqueda
3. Ejecución de la SLR:
 - 3.1 Criterios de Inclusión y Exclusión: Una SLR ubica la información específica para contestar a las preguntas de investigación.
 - 3.2 Proceso de selección de estudios primarios
4. Extracción de datos: Los resultados de una SLR se sustentan en la información de los estudios primarios.

2.2 Marco Teórico

2.2.1 Antecedentes Investigativos

La evolución de la tecnología de la información ha permitido que hoy en día cada vez se desarrollen más productos de software (Kitchenham, y otros, 2012). El rápido

crecimiento de la zona también ha ocasionado que numerosos proyectos de software hayan tenido problemas en cuanto a su funcionalidad, a los excesos de costes, a plazos incumplidos y su mala calidad (Kitchenham, y otros, 2012).

El desarrollo de aplicaciones para dispositivos móviles, en los últimos años ha tenido un crecimiento exponencial (Ferré, 2000). Este tipo de aplicaciones cuando son comercializadas se espera que posean un alto grado de aceptación por los usuarios, por lo que para los desarrolladores de software resulta importante poder evaluar esa calidad o efectuar pruebas de calidad (Enriquez & Casas, 2013). Uno de los factores más importantes dentro de la calidad de un sistema es su usabilidad (Enriquez & Casas, 2013).

La investigación en lo que respecta a la usabilidad de aplicaciones móviles se ha incrementado notablemente en los últimos años, debido a la demanda en el uso de dispositivos y aplicaciones móviles (Wasserman, 2010).

Entre las métricas y métodos que son usados en la actualidad para medir la usabilidad de las aplicaciones móviles se puede observar que fueron creados para aplicaciones de escritorio, por lo que su aplicación no se ajusta adecuadamente a las necesidades de los usuarios (Enriquez & Casas, 2013).

Para conocer el estado del arte y de la práctica sobre la evaluación de usabilidad de aplicaciones móviles, en la presente tesis se realizará una revisión sistemática de literatura, para encontrar las evidencias disponibles sobre los métodos o modelos que están siendo usados para este fin. También para saber qué tipo de problemas ocurren en el proceso de evaluación y como estos pueden ser solucionados. La SLR es muy importante dentro de una investigación ya que permite “analizar, identificar e interpretar las evidencias disponibles relacionadas a una pregunta específica de investigación”(Kitchenham, y otros, 2012).

2.2.2 Fundamentación Filosófica

La presente tesis estudia una realidad de la situación actual sobre la evaluación de usabilidad de aplicaciones multiplataforma en dispositivos móviles, para establecer el estado del arte sobre este tema, y así manifestar la importancia de tomar en cuenta este aspecto desde etapas iniciales del ciclo de vida de un producto software, y también conocer los métodos y técnicas más actuales que son usadas por los desarrolladores y expertos en usabilidad.

La usabilidad llega a ser un punto clave en la adopción de las aplicaciones móviles por los usuarios, por lo que los resultados de esta investigación podrían contribuir a los desarrolladores, para que tomen en cuenta la importancia del descubrimiento de problemas de usabilidad dentro de una etapa temprana de desarrollo, o también los ayudará a conocer que proceso de evaluación llevar a cabo en caso que se requiera evaluar la usabilidad después de la etapa de despliegue del software, y comprender cuál método o modelo es el más conveniente y aplicable para el tipo de aplicación, tipo de usuarios y contexto de uso.

La usabilidad de un producto es normalmente demostrada a través de sus interfaces, un conjunto de lineamientos y estándares han sido introducidos, pero para aplicaciones móviles estos lineamientos todavía faltan, están relativamente inexploradas o aún no están probados (Baharuddin, Singh, & Razali, 2013). Esto se debe a las características peculiares de las aplicaciones móviles y su contexto de uso por lo que es importante diseñar interfaces más efectivas, para que así la usabilidad de un producto de software incremente (Baharuddin, Singh, & Razali, 2013).

Por estos motivos la usabilidad es un tema que debe ser tratado, ya que si el usuario no percibe al sistema como usable, se le dificulta la aceptación del mismo (Ferré, Juristo, Windl, & Constantine, 2001).

2.2.3 Red de Categorías

Con el objetivo de buscar la congruencia en la fundamentación teórica de la presente tesis, es importante organizar una red de las categorías principales que intervienen en la explicación y entendimiento científico del tema en estudio; la red se muestra a continuación en la Figura 1.

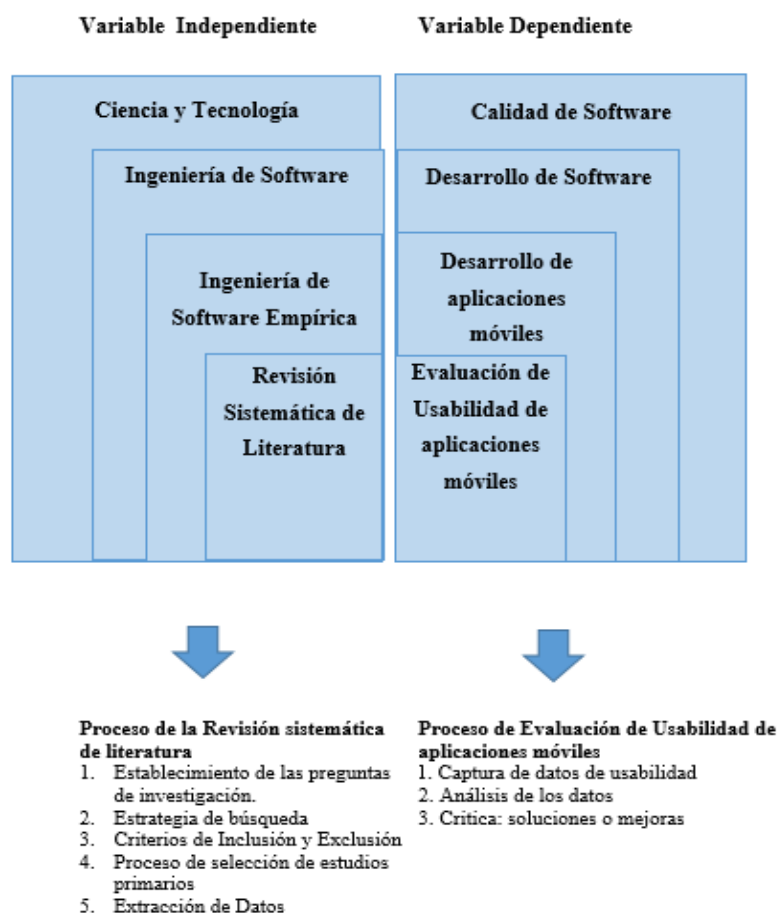


Figura 1. Red de Categorías de las Variables de Investigación

2.2.4 Fundamentación Científica de la Variable Independiente

Ciencia y Tecnología

La ciencia y tecnología se han convertido en las últimas décadas en el progreso de la sociedad, ya que constituyen un pilar de desarrollo social, cultural y económico de la vida en las personas de la sociedad moderna (Cañedo, 2000). El conocimiento científico y tecnológico hoy en día es una de las riquezas importantes de las sociedades (Cañedo, 2000).

Ciencia es aquella esfera de las actividades de la sociedad, cuyo propósito es la obtención de conocimientos sobre del mundo circundante. La ciencia está desarrollada por cuatro componentes esenciales: el factor humano, el factor social, el factor cognitivo y el factor material (Cañedo, 2000).

El factor humano está constituido por los científicos y todas las personas que trabajan con fines de la actividad científica. El factor social se compone por las relaciones que tienen los científicos en su marco de trabajo, estas son las sociedades,

los equipos y grupos de trabajo, etc. El factor cognitivo incluye los procesos requeridos para formar los conocimientos teóricos, prácticos, metodológicos, entre otros. Por último el factor material, que son los instrumentos y equipos utilizados como herramientas por los científicos en el proceso de formación del conocimiento, como por ejemplo un laboratorio (Cañedo, 2000).

La tecnología en cambio compone el sector de la actividad de la sociedad, que está interesada en la modificación del mundo circundante. El proceso de transformación de la realidad objetiva se realiza en un ciclo de cinco etapas que contiene tanto al servicio o producto como a los procesos necesarios para su generación (Cañedo, 2000).

Las cinco fases por las que un servicio o producto debe atravesar son: determinación de la necesidad, diseño y desarrollo del servicio o producto, producción o prestación del servicio, valoración del suministrador y el cliente y la última fase es la de análisis del perfeccionamiento del servicio o producto (Cañedo, 2000).

La unión entre la ciencia y la tecnología, y de ésta la fabricación material o la transformación de la ciencia en fuerza productiva, son atributos característicos del cambio que en la actualidad se maneja en las fuerzas productivas. Debido a ello, el progreso de la ciencia y tecnología es un aspecto del desarrollo del ser humano, como la fuerza principal productiva de la población (Cañedo, 2000).

Ingeniería de Software

La revolución de la tecnología de la información ha significado entre otras cosas que el software ha llegado a ser parte de más productos. El software es encontrado en productos cotidianos desde tostadoras hasta trasbordadores espaciales. Esto significa que el desarrollo de software cada vez va en incremento (Kitchenham, y otros, 2012).

El desarrollo de software no es una tarea fácil, es un proceso altamente creativo (Kitchenham, y otros, 2012). El rápido crecimiento en esta área ha provocado también identificar problemas encontrados en el software, por ejemplo: pérdida de funcionalidades, costos, poca calidad. Estos problemas y desafíos son identificados con el término “ingeniería de software”, que fue acuñado con la intención de crear una disciplina de la ingeniería enfocada al desarrollo intensivo de software (Kitchenham, y otros, 2012).

La ingeniería de software es formalmente definida por la (IEE, 1990) como: “ingeniería de software significa la aplicación de una mejora sistemática, disciplinada

y cuantificable para el desarrollo, operación y mantenimiento del software” (Kitchenham, y otros, 2012).

Ingeniería de Software empírica

Los estudios empíricos tradicionalmente han sido usados en ciencias sociales y psicología, es decir tienen que ver con el comportamiento humano. La ingeniería de software está también muy regulada por el comportamiento humano, a través de las personas que desarrollan software. Por lo tanto, se requiere hallar las normas o leyes formales en la ingeniería de software, excepto quizás cuando se centra en los aspectos técnicos específicos de un sistema (Wohlin, Runeson, Höst, & Ohlsson, 2012).

Los estudios empíricos son muy importantes para apoyar el logro de este objetivo y encajar en el contexto de la investigación en ingeniería de software industrial y académico, así como en una organización de aprendizaje, buscando siempre la mejora continua (Wohlin, Runeson, Höst, & Ohlsson, 2012).

Los estudios empíricos son un insumo importante para la toma de decisiones en la búsqueda de una mejora para la organización. Antes de la introducción de nuevas técnicas, métodos, u otras formas de trabajo, se prefiere una evaluación empírica de las virtudes de este tipo de cambios (Wohlin, Runeson, Höst, & Ohlsson, 2012).

Revisión Sistemática de Literatura

“Una revisión sistemática de literatura (SLR de los siglos en inglés Systematic Literature Review) es un medio de identificación, evaluación e interpretación de información disponible y pertinente con respecto a una pregunta en particular de una investigación, temática o fenómeno de interés (Kitchenham & Charters, 2007).

Muchas investigaciones inician con una SLR, ya que gracias al uso de esta metodología se puede dar una vista integral y válida de ciertas pruebas existentes, por esta razón este proceso debe realizarse de una manera rigurosa.

Kitchenham y Charters han adecuado ciertas directrices de la SLR a la ingeniería de software en un proceso de tres pasos: planificación, realización y presentación de informes de la SLR.

1. Planificación de la SLR: La planificación de la SLR incluye los siguientes pasos:

Identificación de necesidad de la Revisión: La necesidad de una SLR se origina de un investigador con la finalidad de comprender el estado del arte, utilizando

evidencias empíricas que contribuirán en las actividades de toma de decisiones o de mejora.

Especificación de la pregunta de investigación: Especificación de las preguntas de investigación, que son el foco para identificar estudios primarios del tema, la extracción de datos de los estudios y el análisis.

Desarrollo de un protocolo de revisión: En el desarrollo de un protocolo de revisión, es donde se definen los procedimientos necesarios para la SLR. Para el establecimiento de este protocolo se requiere los siguientes elementos: antecedentes y justificación, las preguntas de investigación, la estrategia de búsqueda de estudios primarios, criterios de selección de los estudios, procedimientos de selección de estudios, estrategia de extracción de datos, síntesis de los datos extraídos.

2. Realización de la SLR

La realización de la SLR implica establecer el protocolo de la revisión a la práctica, incluye los siguientes pasos:

Identificación de investigación: Se debe realizar la especificación de cadenas de búsqueda y su aplicación a las bases de datos, es importante una estrategia efectiva de búsqueda para obtener todos los estudios relevantes y evitar falsos positivos o artículos que aparentemente son de interés y posteriormente resulta que no lo son por el mismo motivo los falsos positivos deben ser excluidos.

Selección de estudios primarios: Los aspectos a tomar en cuenta al realizar la selección de estudios primarios son: criterios de inclusión y exclusión. Los criterios de búsqueda deben ser elaborados en la planificación de la SLR, pero es muy común que los criterios tengan que ser ajustados en el transcurso de la investigación.

Evaluación de la calidad del estudio: Esta etapa es muy importante ya que los estudios primarios pueden reportar resultados contradictorios, la evaluación se puede usar para analizar la causa de la obtención de esos resultados.

Extracción de datos y seguimiento: Una vez que la lista de estudios primarios este completa, se extraen los datos en un formulario adecuado de extracción de datos.

Síntesis de Datos: Aplicación de métodos para el análisis de los resultados de los estudios

3. Presentación de informes de la SLR

Como cualquier estudio empírico, la SLR se puede informar a un público específico, el informe tiene que adaptarse adecuadamente, incluye:

- Los cambios en el protocolo de estudio
- Las listas completas de los estudios primarios incluidos
- Datos sobre su clasificación
- Datos derivadores de cada estudio
- Publicación de los resultados en un artículo técnico

2.2.5 Fundamentación Científica de la Variable Dependiente

Calidad de Software

Desde el punto de vista de la Ingeniería de software, una de las principales particularidades que debe poseer una aplicación para ser considerada como exitosa entre los usuarios es que sea de calidad (Enriquez & Casas, 2013).

La calidad del software es el grado que el sistema cumple con un conjunto de atributos de calidad específicos requeridos para ese sistema. Uno de los atributos más importantes es la usabilidad, que indica la facilidad con la que los usuarios pueden utilizar un producto software (Enriquez & Casas, 2013).

Una herramienta importante para garantizar la calidad del software es el Plan de Calidad, que posee un conjunto de normas o estándares genéricos con procedimientos específicos. Los procedimientos podrían cambiar dependiendo de cada organización, pero es imprescindible que los mismos estén especificados y adaptados a los procesos concretos de la organización, para ser cumplidos (Scalone, 2006).

Los requisitos del software juegan un papel importante, ya que son una base de la medida de calidad y la escasa concordancia con estos requisitos demuestra una falta de calidad. Es decir, si no se sigue las metodologías o estándares de la ingeniería de software en el proceso de desarrollo siempre habrá falta de calidad en el producto software (Scalone, 2006).

Es importante distinguir la calidad del producto software y la calidad del proceso de desarrollo. Los objetivos o metas que se constituyan para la calidad del producto software son las que determinarán las metas a definir para la calidad del proceso de

desarrollo, es decir “la calidad del producto va a estar en función de la calidad del proceso de desarrollo” (Scalone, 2006). Sin un proceso de desarrollo de software adecuado es casi inalcanzable conseguir un buen producto (Scalone, 2006).

Desarrollo de Software

El proceso de desarrollo de software es la descripción de una secuencia de actividades que deben ser realizadas por los desarrolladores con el fin de producir y mantener un producto software eficaz y eficiente que reúna los requisitos del cliente (Donaldson & Siegel, 2000).

El desarrollo de software es una actividad muy importante dentro de la computación. Para desarrollar software se requiere resolver algunas etapas de un proceso que es conocido como el ciclo de vida del desarrollo de software (Scalone, 2006). Estas etapas del desarrollo de software se muestran a continuación en la Figura 2.



Figura 2. Etapas del desarrollo de software

Fuente: (Scalone, 2006)

Para tener éxito en el desarrollo de software existen algunos requisitos básicos que deben ser tomados en cuenta (Kitchenham, y otros, 2012):

- Comprensión del proceso de software y producto
- Definición de las cualidades del proceso y producto
- Evaluación de los éxitos y fracasos
- La retroalimentación de información para el control del proyecto.
- Aprender de la experiencia
- Embalaje y reutilización de experiencia relevante

Desarrollo de Aplicaciones móviles

Los dispositivos móviles son aparatos electrónicos que forman parte de nuestra vida diaria, poseen muchas funcionalidades y características especiales, una memoria limitada, diferentes capacidades de procesamiento, y cada vez son más potentes los dispositivos que salen al mercado (Angulo Cevallos, 2014).

Para desarrollar software para este tipo de dispositivos se debe tener en cuenta las restricciones que poseen tanto en hardware y software, por ejemplo: sus dimensiones reducidas, poca capacidad de almacenamiento y ancho de banda, tienen bajo poder de cómputo, etc. (Enriquez & Casas, 2013).

Entre los dispositivos móviles existe una heterogeneidad en sus aplicaciones, por lo que es importante que el desarrollador realice una planificación y diseño previo a la implementación del sistema (Enriquez & Casas, 2013).

Los desarrolladores de software al crear una aplicación móvil deberían tener en cuenta el público hacia el cual va dirigido, y también las características de interfaz propias de cada plataforma, ya que al desarrollar aplicaciones móviles, no existe un diseño formal o directrices generalizadas para conseguir un diseño de interacción móvil multiplataforma (Barea, Ferré, & Villarroel, 2013).

Existen varios tipos de aplicaciones móviles que son creadas por los desarrolladores las mismas se exponen a continuación:

Aplicaciones Móviles Multiplataforma: El desarrollo de aplicaciones multiplataforma tiene como objetivo compartir la codificación de los distintos sistemas para distintas plataformas para mejorar la relación entre el beneficio y el costo (Barea, Ferré, & Villarroel, 2013).

Aplicaciones Nativas: Las aplicaciones móviles nativas poseen archivos ejecutables que están descargados directamente en el dispositivo o también son pueden ser descargadas desde la tienda de aplicaciones propia de cada plataforma y almacenados localmente (IBM, 2012).

Aplicaciones Web móviles: Las aplicaciones web móviles son corridas dentro de un navegador, la tecnología utilizada para el desarrollo puede ser JavaScript, CSS, HTML. Este tipo de aplicaciones tiene muchas ventajas ya que no dependen del sistema operativo para adecuarse, es decir son independientes. El tiempo de respuesta

de cada tarea depende o es afectada por la velocidad del internet ya que es un sistema cliente-servidor (Enriquez & Casas, 2013).

Aplicaciones Híbridas: Las aplicaciones híbridas son una combinación del desarrollo web y nativo, este tipo de aplicaciones son desarrolladas con lenguajes de desarrollo web como JavaScript o HTML y un framework para crear aplicaciones híbridas como titanium, steroids, etc., estas características hacen que las aplicaciones puedan ser usadas en diferentes plataformas (IBM, 2012). Una importante ventaja de este tipo de aplicaciones es que se puede reutilizar no solo código y su costo de desarrollo es menor que el de las aplicaciones nativas (IBM, 2012).

Evaluación de Usabilidad de Aplicaciones móviles

El estándar ISO9241 define a la usabilidad como “el grado en el que un producto puede ser utilizado por usuarios específicos para conseguir objetivos específicos con efectividad, eficiencia y satisfacción en un determinado contexto de uso”. Como se puede apreciar, en esta definición la usabilidad del sistema es un atributo inherente al software, es decir, no puede especificarse independientemente del entorno de uso y de los usuarios concretos que vayan a utilizar el sistema (Coello, Ferreira González, & Urrútia , 2011).

Existen atributos, características o propiedades que se consideran en el estándar ISO9241, para que una aplicación o sistema sea usable:

- Satisfacción del usuario, el grado en que se encuentra conforme y el sistema alcanza los objetivos determinados.
- Efectividad de la aplicación, la precisión con la que cumple los objetivos específicos.
- Eficiencia de la aplicación, el correcto uso de los recursos empleados para llegar a cumplir los objetivos.
- Facilidad de Aprendizaje con la que los usuarios alcanzan los objetivos específicos la primera vez que usan el sistema o aplicación.
- Errores que el usuario comete cuando usa la aplicación.
- Facilidad del usuario para memorizar el uso de la aplicación, la curva de aprendizaje.

- Accesibilidad a la aplicación, tomando en cuenta posibles limitaciones físicas, visuales, etc.

Los métodos y métricas actuales que miden usabilidad no se pueden aplicar de manera directa al contexto móvil (Enriquez & Casas, 2013).

En el estándar ISO9241 para medir o especificar la usabilidad es importante descomponer los atributos y el contexto de uso, en componentes que deben ser medibles y verificables. En la Figura 3 se puede observar el framework de usabilidad que propone esta norma, donde se encuentran las diferentes relaciones entre el usuario, el producto, el contexto de uso y los atributos del sistema (Enriquez & Casas, 2013).

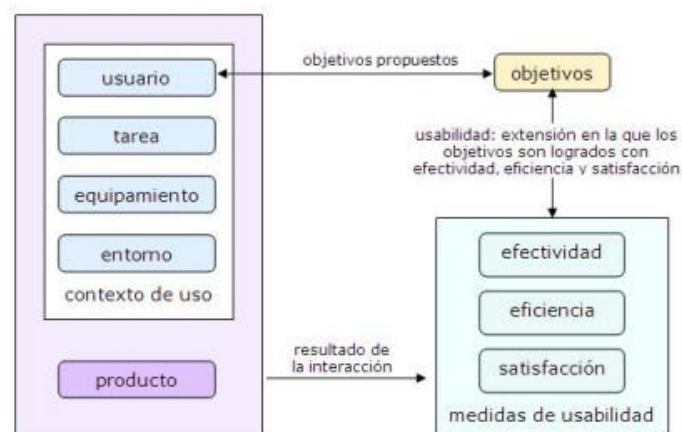


Figura 3. Framework de usabilidad (ISO 9241-11)

Fuente: (Enriquez & Casas, 2013).

Entre las metodologías más usadas actualmente para la evaluación de usabilidad de dispositivos se encuentran:

Evaluaciones con expertos: Este tipo de evaluaciones, requieren ser analizadas por un grupo de expertos que evalúan las interfaces del usuario y demás problemas de usabilidad que pueden existir en un sistema (Enriquez & Casas, 2013).

Evaluaciones Heurísticas: Un grupo de expertos inspeccionan el diseño de la interfaz de la aplicación basándose en principios determinados de usabilidad previamente establecidos, verificando de esta manera el cumplimiento de cada uno de ellos (Enriquez & Casas, 2013). Las diez heurísticas o principios de usabilidad más populares son las denominadas Reglas de Nielsen, que miden los siguientes atributos:

- Visibilidad del estado de un sistema
- Adecuación del lenguaje
- Control y libertad para el usuario
- Consistencia y estándares
- Prevención de errores
- Reconocimiento más que memoria
- Flexibilidad y eficiencia de uso
- Diálogo estético y diseño minimalista
- Soporte a los usuarios para reconocimiento, diagnóstico y recuperación de errores
- Ayuda y documentación

Recorridos cognitivos: Se trata de que un experto recorre un escenario específico de tareas y de cómo las realizaría determinado tipo de usuario (Enriquez & Casas, 2013). Existen dos tipos de recorridos:

- Un recorrido cognitivo individual
- Un recorrido cognitivo en grupo, con usuarios, desarrolladores, etc.

Inspección formal de usabilidad: Se utiliza la metodología de inspecciones de código, son más frecuentes, contemplan con mayor especificación los conceptos más técnicos, entre ellos se encuentran:

- Inspecciones Formales de Usabilidad, que enfatizan la búsqueda de errores recorriendo las tareas con los objetivos específicos del usuario.
- Inspecciones de Características, revisan solo un conjunto de características del sistema.
- Inspecciones de Consistencia, por medio de una comparación de sistemas similares, se desea asegurar la consistencia.
- Inspecciones de Estándares, garantizan que los sistemas estén acorde a los estándares de la industria.

Revisión de guías de estilo: Consiste en la revisión del sistema donde se verifica que cumpla con las reglas establecidas por la institución o entorno. Existen dos tipos de revisiones de guías de estilo:

- Guías de comprobación: aseguran la revisión adecuada de los principios de usabilidad.
- Guías de comprobación basadas en escenarios: se lleva a cabo por medio de tres escenarios: usuario experto, usuario novato, y el manejo de errores, para cada uno de los perfiles se debe comprobar ciertos aspectos.

Test de Usabilidad: Los test de usabilidad se usan básicamente para la estimación de la facilidad con la que una persona aprende a usar una aplicación o sistema, tomando en cuenta los frecuentes problemas para que estos sean solucionados (Enriquez & Casas, 2013). Primero se selecciona un conjunto de usuarios o no usuarios de la aplicación, esto depende de los objetivos del test de usabilidad y de cuáles serán los perfiles de usuario que van a manejar, posteriormente se pide a los usuarios que realicen tareas determinadas (Ferré, 2000). A continuación el grupo de diseño y desarrollo del sistema van tomando nota de las dificultades que tuvieron los usuarios, el número de errores (Ferré, 2000).

Los test de usabilidad pretenden mejorar la usabilidad de un producto, para lo cual toman distintas métricas como son la eficiencia, la efectividad, la satisfacción del usuario (Ferré, 2000) y son generalmente realizados en laboratorios con el equipamiento necesario para formar el entorno de trabajo donde el usuario procederá utilizar el producto y el grupo de expertos a monitorizar y evaluar la interacción (Ferré, 2000).

2.2.6 Hipótesis

La aplicación de una revisión sistemática de literatura, permitirá la comprensión y establecimiento del estado del arte de la Evaluación de usabilidad de aplicaciones multiplataforma en dispositivos móviles.

Señalamiento de Variables:

- **Variable Independiente:** Revisión sistemática de literatura
- **Variable Dependiente:** Evaluación de usabilidad de aplicaciones multiplataforma en dispositivos móviles.

Para el cumplimiento de la hipótesis planteada para este trabajo de investigación, se desarrollará una Revisión Sistemática de Literatura que es detallada con cada una de sus fases en los siguientes capítulos de la presente tesis.

CAPÍTULO III

PLANIFICACIÓN DE LA REVISIÓN SISTEMÁTICA LITERARIA

La planificación de la Revisión Sistemática de Literatura es la primera etapa del proceso según (Kitchenham & Charters, 2007) y (Wohlin, Runeson, Höst, & Ohlsson, 2012). Esta etapa es muy importante y consta de cuatro actividades: identificación de la necesidad de la SLR, especificación de las preguntas de investigación, definición de la estrategia de búsqueda, finalmente la definición de criterios de inclusión y exclusión.

3.1 Necesidad de la Investigación

Como se ha indicado en la problemática de la presente investigación, existen carencias en el proceso de evaluación de usabilidad de aplicaciones móviles multiplataforma. El principal problema se centra en la falta de una adecuada aplicación de métodos y técnicas en el proceso de evaluación de usabilidad, el mismo que muchas veces es aplicado por los desarrolladores al final del ciclo de vida del producto software, lo cual produce aplicaciones poco usables y usuarios insatisfechos (Ferré, Juristo, Windl, & Constantine, 2001).

Por otro lado, uno de los desafíos a los que se enfrentan los desarrolladores de aplicaciones móviles es la heterogeneidad en el diseño de interacción de una aplicación en los dispositivos móviles según la plataforma en la que se ejecuta (Barea, Ferré, & Villarroel, 2013). Por lo tanto es importante realizar una revisión sistemática cuyo objetivo sea identificar los estudios relevantes que reporten evaluaciones de la usabilidad de aplicaciones en dispositivos móviles multiplataforma, para conocer los métodos y técnicas que actualmente se usan para este fin, también para saber en qué medida el diseño de interacción de las aplicaciones móviles difiere entre las plataformas existentes y cómo esto influye en los usuarios para el uso de una aplicación móvil. En conclusión toda la información importante que se encuentre en los estudios primarios sobre este tema que pueda contribuir a los desarrolladores a saber la aplicación correcta de métodos y técnicas para que la usabilidad de las aplicaciones móviles se incremente y los usuarios cada vez estén más satisfechos con sus aplicaciones.

3.2 Definición de las Preguntas de Investigación

La especificación de las preguntas de investigación es una parte importante en la etapa de planeación del proceso de la SLR, ya que estas definen el objetivo de la búsqueda como alcanzable y todo el proceso de la SLR debe enfocarse en responder estas preguntas de investigación que están directamente relacionadas con la problemática que aborda la presente tesis (Kitchenham & Charters, 2007).

Para cumplir el objetivo deseado en este trabajo de investigación, se plantearon las siguientes preguntas de investigación:

- **RQ1:** ¿Cuáles son las técnicas existentes para evaluar la usabilidad de aplicaciones en dispositivos móviles?
- **RQ2:** ¿Qué modelos de usabilidad son comúnmente utilizados en la evaluación de usabilidad de aplicaciones móviles?
- **RQ3:** ¿Qué atributos de usabilidad son considerados en la evaluación de usabilidad de aplicaciones móviles y que métricas se utilizan para medir dichos atributos?
- **RQ4:** ¿Qué tipo de aplicaciones móviles son evaluadas y a través de qué actividades?
- **RQ5:** ¿Qué ventajas e inconvenientes se generan al momento de evaluar la usabilidad de las aplicaciones móviles en los estudios?
- **RQ6:** ¿Con qué frecuencia se consideran los factores de usabilidad en los estudios existentes?
- **RQ7:** ¿Con qué frecuencia se considera el diseño de interacción respecto a los factores de usabilidad en el desarrollo de aplicaciones móviles?
- **RQ8:** ¿En qué medida difiere el diseño de interacción de las aplicaciones móviles entre las plataformas existentes?

Generalmente las preguntas de investigación se expresan con el formato PICOC, lo cual en esta investigación no es aplicable porque PICOC es definido dentro de una revisión sistemática de literatura específicamente para la comparación de dos tratamientos versus una salida. En el caso de la presente tesis se está buscando información concreta en los artículos, así que, plantear una pregunta con PICOC, no parece tener mucha utilidad (Fonseca, 2014).

3.3 Definición de la Estrategia de Búsqueda

La estrategia de búsqueda planteada consta de las siguientes actividades: conformación del grupo de control, construcción de la cadena de búsqueda y búsqueda de los estudios candidatos. Ver Figura 4.

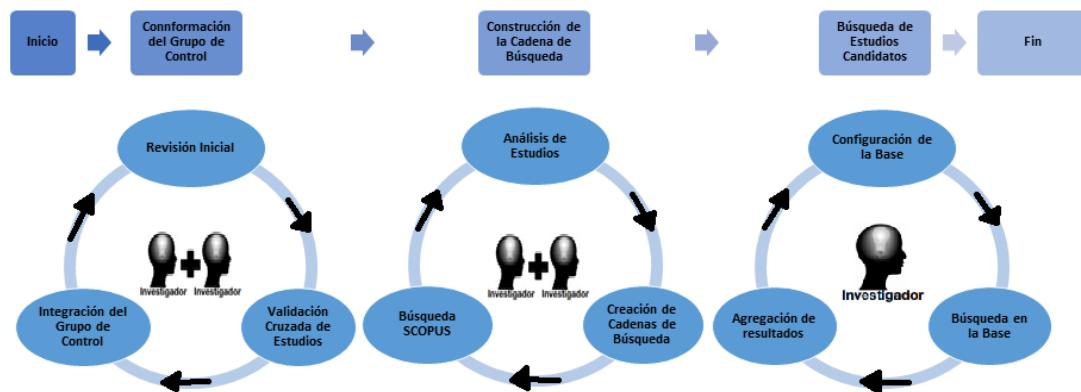


Figura 4. Proceso de Estrategia de Búsqueda de Estudios

La etapa de **Conformación del grupo de control**, se refiere a un conjunto de estudios que responden a las particularidades y objetivos que plantea la presente SLR. El objetivo del grupo de control es facilitar la localización y establecimiento de palabras claves que contribuyen a la formación de la cadena de búsqueda de la SLR, de esta manera se asegura una cobertura de literatura conveniente y correcta, por la variedad de palabras claves que se manejan en los artículos (Fonseca, 2014). Los estudios del grupo de control no necesariamente conformarán el grupo de los estudios primarios, ya que para ser considerado como tal tendrán que atravesar un proceso de selección más exhaustivo que se detallará más adelante.

Para la conformación del grupo de control inicialmente se hizo una revisión inicial de la literatura individual, la misma que fue llevada a cabo por tres investigadores. Posteriormente los estudios ubicados por cada investigador fueron sometidos a un proceso de validación cruzada entre los mismos investigadores. De esta manera se garantiza que los estudios del grupo de control cumplan con los criterios de inclusión y exclusión de la presente SLR.

Dentro de la etapa de conformación del grupo de control se obtuvo catorce estudios de la disciplina de ingeniería de software. Cada uno de estos estudios fue analizado para abstraer los términos generales que contribuyeron a conformar las cadenas de

búsqueda. En la sección Anexos la Tabla A.1 muestra los estudios del grupo de control y los términos abstraídos de cada estudio.

La **construcción de la cadena de búsqueda**, empieza con el análisis del grupo de control, cada estudio de este grupo se revisó a detalle a nivel de título, resumen y las palabras claves, con el objetivo de hallar: términos generales de cada estudio, palabras comunes entre los estudios y si es posible términos comunes referidos al propósito de la SLR. Los términos obtenidos del grupo de control permiten conformar una o algunas cadenas de búsqueda, con el uso de los operadores lógicos: OR para agregar sinónimos y AND para aumentar nuevos términos.

Cada una de las cadenas de búsqueda es probada en las bases de datos digitales especificadas, y según los resultados que se obtienen con las cadenas, se realizan las modificaciones respectivas a la cadena anterior, con el propósito de lograr como resultados de la búsqueda la mayoría de los estudios pertenecientes al grupo de control. Finalmente se elige una cadena de búsqueda, la que alcanza los mejores resultados. Este proceso es validado por dos investigadores.

Usando la base de datos digital SCOPUS se conformaron ocho cadenas de búsqueda (ver Anexos Tabla A.2). De las ocho, fue seleccionada la cadena 8 (CD8), la misma que incluye once de los catorce estudios del grupo de control. En la Tabla 1, se puede observar los siguientes datos: el identificador de la cadena, el número de estudios que resultan de cada cadena en la base de datos SCOPUS, los estudios del grupo de control recuperados de la búsqueda de la cadena.

Tabla 1.

Estudios por Cadena de Búsqueda

CADENA	#ESTUDIOS	ESTUDIOS GRUPO DE CONTROL
CD1	602	EC1, EC2, EC3, EC4, EC5, EC8, EC11
CD2	396	EC1, EC2, EC3, EC4, EC5, EC8
CD3	858	EC1, EC2, EC3, EC4, EC5, EC8, EC11
CD4	804	EC1, EC2, EC3, EC4, EC5, EC8, EC9, EC11
CD5	359	EC1, EC2, EC3, EC4, EC5, EC7, EC8, EC11
CD6	753	EC1, EC2, EC3, EC4, EC5, EC8, EC9, EC11, EC14
CD7	787	EC1, EC2, EC3, EC4, EC5, EC7, EC8, EC10, EC11, EC14
CD8	368	EC1, EC2, EC3, EC4, EC5, EC7, EC8, EC9, EC10, EC11, EC14

Las cadenas CD1 y CD3 fueron descartadas por el número alto de estudios que devolvían y solo cubrían la mitad de los estudios del grupo de control, de manera similar la cadena CD2, devolvía menos número de estudios pertenecientes al grupo de control y por ese motivo también fue descartada.

Las cadenas CD4, CD6, CD7 del mismo modo fueron descartadas, ya que si bien es cierto cubrían un número aceptable de estudios del grupo de control, pero la cantidad de artículos que devolvía era muy elevada, la razón es que las cadenas no se asociaron bien al objetivo de la SLR, por tanto los resultados también eran un tanto incoherentes. Finalmente a partir de estas cadenas se añadió variantes y se reorganizó ciertos términos para obtener la cadena CD8, la misma que devolvía 368 artículos y cubría once de los catorce estudios del grupo de control.

La **búsqueda de los estudios candidatos** se la realiza por medio de la configuración de la cadena seleccionada en cada base de datos digital, tomando en cuenta las siguientes consideraciones:

- **Fuentes de Publicación:** Trabajos presentados como artículos en conferencias internacionales, journals, capítulos de libros, revistas científicas y que se encuentren disponibles en la web. Específicamente en las bases de datos bibliográficas a las que se encuentra suscrita la Universidad de las Fuerzas Armadas ESPE.
- **Idioma:** se consideraron exclusivamente publicaciones que están escritas en inglés.
- **Motores de búsqueda:** los motores de búsqueda considerados fueron las siguientes bases digitales: Springer, Scopus y ACM
- **Método de búsqueda:** la cadena de búsqueda fue adecuada al formato de lenguaje propio de las bases de datos digitales indicadas.

Estas bases de datos digitales se establecieron como motores de búsqueda ya que como se mencionó anteriormente son a los que está afiliada la Universidad de las Fuerzas Armadas ESPE, razón por la cual se nos ha permitido el acceso a los diferentes estudios para la presente investigación. Sin embargo algunos estudios no se los pudo obtener desde estas fuentes, por lo que se procedió a contactar directamente a los autores de los estudios a través del correo electrónico o también se los pudo conseguir por medio del acceso del director de la presente tesis. Como resultado de las búsquedas

en las bases digitales se obtiene un conjunto de estudios candidatos que son organizados según la base de datos digital. En el caso de Springer por la gran cantidad de estudios que devolvía la cadena se refinó para que los resultados que devuelva la cadena estén dentro de las categorías: “Computer Science” y “HCI”.

3.4 Definición de los criterios de inclusión y exclusión

Los criterios de inclusión y exclusión constituyen los parámetros que deben cumplir los estudios para ser considerados o no en la investigación, estos criterios ayudan a ubicar la información pertinente para contestar las preguntas de investigación planteadas.

Criterios de inclusión: Los criterios de inclusión definidos para a presente revisión sistemática de literatura son los siguientes:

- El artículo describe información sobre evaluación de la usabilidad en aplicaciones multiplataforma para dispositivos móviles.
- El artículo hace referencia a la importancia de medir la usabilidad en aplicaciones multiplataforma para dispositivos móviles, en las diferentes etapas del proceso de ciclo de vida del producto software.
- El artículo describe el método o técnica utilizada para evaluar la usabilidad en aplicaciones móviles.
- El artículo describe los problemas suscitados al momento de evaluar aplicaciones multiplataforma para dispositivos móviles.
- El artículo que proponga metodologías, técnicas o modelos para evaluar la usabilidad en aplicaciones móviles y su validez sea probada experimentalmente con usuarios finales, aleatorios o expertos.
- Aquellos artículos que contengan información detallada sobre el proceso de evaluación de usabilidad de aplicaciones móviles.
- Los estudios realizados desde el año 2014 hasta el 2016.

Criterios de exclusión: Los criterios de exclusión definidos para a presente revisión sistemática de literatura son los siguientes:

- Artículos que no hagan referencia a estudios de la usabilidad de aplicaciones en dispositivos móviles

- Aquellos artículos que muestren guías de diseño de interacción que se centren en aspectos meramente de diseño de interfaz de usuario en aplicaciones multiplataforma y que omitan información sobre medición de usabilidad.
- El artículo que este escrito en otro idioma que no sea inglés.
- Aquellos artículos que posean solo propuestas de técnicas o métodos para evaluar usabilidad y cuya validez no sea evaluada.
- Aquellos estudios que no respondan ninguna pregunta de investigación de la SLR.
- Artículos que no contengan información detallada de cómo realizaron la evaluación de usabilidad de aplicaciones móviles.
- Artículos que hayan sido publicados antes del 2014.

Con los estudios candidatos obtenidos se procede al proceso de selección de estudios primarios que se detallará en el siguiente capítulo de la presente tesis que es el de la Ejecución de la SLR.

CAPÍTULO IV

EJECUCIÓN DE LA REVISIÓN SISTEMÁTICA LITERARIA

Posteriormente de la aplicación de la estrategia de búsqueda, inicia la etapa de ejecución de la SLR sobre los repositorios, en donde se desarrolla el proceso de selección de los estudios primarios, procedimiento que comprende de las siguientes actividades: Depuración de estudios candidatos, Selección de estudios candidatos y Selección de estudios primarios. Ver la Figura 5.



Figura 5. Proceso de Selección de Estudios Primarios

4.1 Depuración de Estudios Candidatos

El proceso de depuración de los estudios candidatos se realiza con el fin de hallar posibles estudios duplicados, empieza con un refinamiento de los estudios obtenidos en cada base de datos digital, para posteriormente realizar una depuración cruzada de las coincidencias encontradas entre las bases digitales. Finalmente se realizan unos listados depurados de los estudios candidatos por cada base de datos digital. Estas tareas son aprobadas por dos investigadores.

En la etapa de depuración de los estudios candidatos, se ubicó en Scopus un estudio que parecía ser duplicado ya que tenía el mismo título y autores, pero el resumen de cada uno era diferente, para estar seguro de que no son duplicados se procedió a leer ambos artículos y se descubrió que son diferentes y que uno es la mejora del primer estudio, por lo tanto, ambos fueron tomados en cuenta. En la depuración cruzada también se pudieron observar varias coincidencias entre los resultados de las bases de datos digitales, 1 entre las tres bases digitales: Springer, Scopus y ACM, 26 artículos coincidentes entre Scopus y ACM, 3 artículos coincidentes entre Springer y Scopus, y tan solo 1 entre Springer y ACM. En la Figura

6, se puede observar las coincidencias de estudios candidatos encontradas en los resultados de las búsquedas de las diferentes bases digitales.

4.2 Selección de Estudios

Posteriormente por cada grupo de estudios candidatos en las distintas bases de datos bibliográficas digitales, se realiza la selección de estudios, que es llevada a cabo por dos investigadores. En esta actividad de selección se inspeccionó el título y resumen de cada estudio candidato, los investigadores seleccionan solo aquellos estudios que consideran que podrían contribuir en la presente SLR. A paso seguido se procedió a realizar la validación cruzada entre los investigadores por cada uno de los estudios de las bases de datos digitales. Esta tarea de validación ayuda a confirmar las coincidencias y desacuerdos entre los investigadores, sobre la inclusión o exclusión de los estudios candidatos. Finalmente se componen los listados de los estudios seleccionados por cada base de datos digital.

En el caso de la presente SLR, se eligió 206 estudios, los cuales son señalados como estudios seleccionados, y se descartaron 394. El proceso de elección y eliminación de los estudios candidatos se caracteriza por un porcentaje alto de acuerdos entre los investigadores, 90, 17%. En lo respecta a los 206 estudios seleccionados, los investigadores tuvieron acuerdo al seleccionar 152 artículos y desacuerdo en 54. Mientras que, en los 394 estudios descartados, tuvieron acuerdo en 389 artículos y desacuerdo en 5. En resumen, el proceso de selección de estudios tuvo un porcentaje de acuerdo del 90,17% y desacuerdo un 9,83%. Así mismo se puede observar en la Figura 6, la distribución de los estudios seleccionados en las diferentes bases de datos digitales y sus coincidencias.

Se ha podido apreciar que los artículos obtenidos en la base digital Springer tienen un comportamiento inusual, dado que muestran temáticas muy diversas. Es decir, algunos artículos muestran temáticas aproximadas al objetivo planteado en la SLR, sin embargo otros presentan tópicos totalmente alejados al objetivo de estudio. Por tanto se considera que los resultados obtenidos en esta base digital no son representativos comparados con los obtenidos en las otras bases digitales utilizadas.

4.3 Selección de Estudios Primarios

En la selección de estudios primarios, interviene solo un investigador. En esta etapa se unifica todos los artículos seleccionados en un solo listado. Después, se procede a descargar y obtener la mayor cantidad posible de estudios seleccionados desde las bases de datos digitales. Posteriormente se realizó un estudio detallado de los artículos, con el cumplimiento de los criterios de inclusión y exclusión de la SLR. Finalmente es compuesto un listado con los estudios primarios. Cabe mencionar que se realizó un proceso de verificación por un segundo investigador, para la validación de los estudios escogidos como primarios.

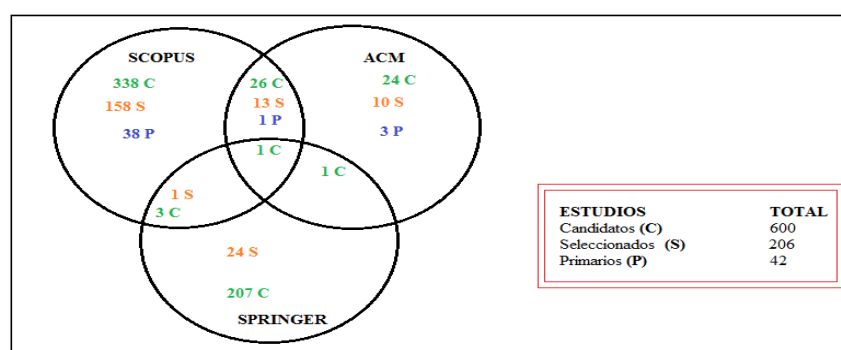


Figura 6. Distribución de Estudios Candidatos, Seleccionados y Primarios

En la presente SLR, dentro de los estudios seleccionados, fueron elegidos 42 estudios primarios (Ver Anexos, Tabla A.3), descartados 160 estudios y no fueron analizados 4 estudios, ya que fue imposible acceder a sus reportes. La Tabla 2 muestra la relación entre los estudios primarios, respecto a los estudios seleccionados, descartados y estudios sin analizar.

Tabla 2

Selección de Estudios Primarios

ESTUDIOS	CANTIDAD	PORCENTAJE
SELECCIONADOS	206	100%
DESCARTADOS	160	78%
PRIMARIOS	42	20%
SIN ANALIZAR	4	2%

Con la obtención de los estudios primarios se procede al proceso de extracción de datos y resultados de la SLR, que es explicado y detallado más específicamente en el siguiente capítulo que es el de Extracción y Síntesis de Resultados.

CAPÍTULO V

EXTRACCIÓN Y SÍNTESIS DE RESULTADOS

5.1 Extracción de Resultados

La extracción de datos de la presente SLR se sustenta principalmente de los estudios primarios. Toda la información obtenida tendrá concordancia con el objetivo de la investigación, y deberá responder al menos una de las preguntas de investigación.

Con el fin de obtener la información requerida para la investigación se planea un proceso de extracción de datos realizado por un investigador y validado por un segundo investigador. Este proceso se muestra en la Figura 7, consta de dos principales actividades: Extracción de las características generales de los estudios primarios, Extracción de las respuestas para cada pregunta de investigación por estudio primario.

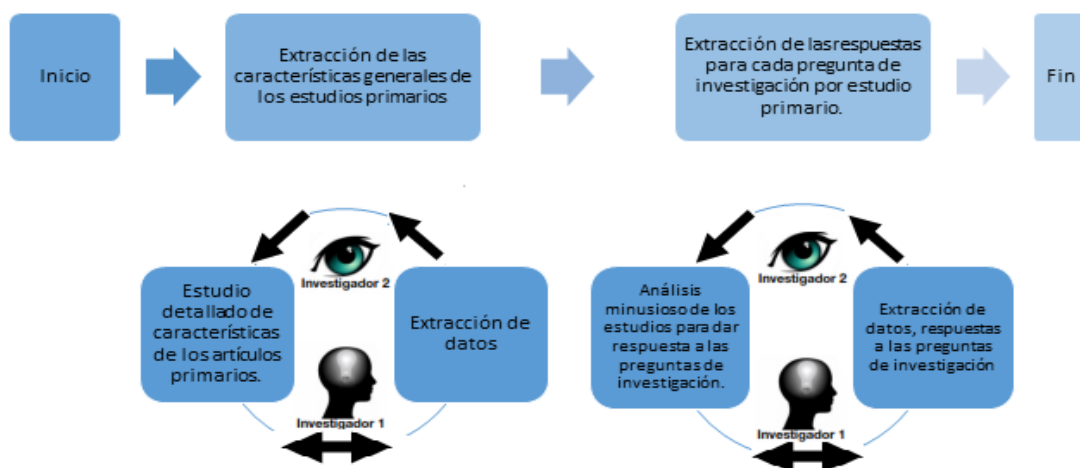


Figura 7. Proceso de Extracción de Datos

El proceso de extracción de datos inicia con un estudio detallado de las características importantes, tanto generales como específicas de los estudios primarios. Las características generales se refieren a los detalles más universales de los estudios primarios, como por ejemplo, año de publicación, disciplina a la que pertenece, etc., y las características más específicas en cambio se refieren a las peculiaridades de cada artículo, como por ejemplo el tipo de metodología de evaluación que utilizan los investigadores, los modelos de usabilidad aplicados en los estudios, es decir, información que puede dar respuesta a las diferentes preguntas de investigación. Para la extracción y clasificación de estos datos, se utilizan: (1) listados de resumen, por

cada estudio primario, donde se muestran, las particularidades más relevantes de los estudios y (2) tablas de datos, que contienen las respuestas a las preguntas de investigación por cada estudio primario, ver Anexos Tabla A.4.

Finalmente se procede al análisis detallado de cada estudio primario, para lo cual se va a utilizar estos listados y tablas que contienen información principal de cada estudio primario respondiendo a las preguntas de investigación para posteriormente realizar el análisis y resumen de cada estudio. El principal objetivo de este análisis es determinar las características generales y específicas de cada estudio que contribuyan a responder las preguntas de investigación y así definir el estado de arte sobre la evaluación de usabilidad de aplicaciones móviles.

A continuación se encuentra el listado de resúmenes de los estudios primarios mencionado anteriormente:

EP1 (Gresse von Wangenheim, y otros, 2016): An usability Score for Mobile Phone Applications based on Heuristics

En este estudio los autores presentan un conjunto de heurísticas de usabilidad en un checklist. Estas heurísticas son una modificación de las de Nielsen en 1994, las mismas que fueron originalmente creadas para aplicaciones de computadoras de escritorio, estas son: visibilidad del estado del sistema, encuentro entre el sistema y el mundo real, control de usuario y facilidad, consistencia y estándares, prevención de errores, reconocimiento en lugar de recuerdo, flexibilidad y eficiencia de uso, diseño estético y minimalista, reconocer ayuda a los usuarios, diagnóstico y recuperación de errores, ayuda y documentación. En este estudio los autores presentan nuevas heurísticas (compatibilidad entre diferentes plataformas, poca interacción humana con dispositivo, interacción física y ergonómica, legibilidad y diseño) y eliminan tres heurísticas tradicionales que consideraron que no contribuyen a la medición de usabilidad, estas son: (Prevención de errores, diagnóstico y recuperación de errores, y finalmente ayuda y documentación).

Este checklist con este conjunto de heurísticas es validado a través de un estudio empírico en el cual los resultados de 247 evaluaciones heurísticas han sido analizados estadísticamente, con puntuaciones para facilitar el proceso de evaluación. Las aplicaciones evaluadas fueron las aplicaciones gratuitas de Google Play y Apple Store, que presentaban altas y bajas puntuaciones de los usuarios.

Las nuevas heurísticas añadidas a las de Nielsen con respecto a las características específicas de los dispositivos móviles han sido confirmadas en general como relevantes para medir usabilidad. A excepción de la heurística “poca interacción humana/dispositivo”, que se refiere a que debido al tamaño de los dispositivos móviles y su contexto de uso, es importante que el esfuerzo de interacción sea reducido, (por ejemplo para reducir la entrada de datos pueden verse características de autocompletado, cajas de textos con sugerencias, etc.). Esta heurística todavía requiere un fuerte estudio ya que no es confirmada como importante al medir usabilidad.

Los resultados de esta evaluación pueden ser usados para medir la usabilidad de aplicaciones móviles desde etapas tempranas del proceso de diseño y de esta forma facilitar las evaluaciones de una manera más rentable.

EP2 (Ibrahim, Hussain, & Binti, 2014): Evaluating Mobile Banking Application: Usability Dimensions and Measurements

Los autores en este estudio proponen un conjunto de dimensiones y mediciones de usabilidad para evaluar las aplicaciones de banca móvil. Crearon cinco dimensiones de usabilidad y doce criterios relevantes que pueden ser usados para evaluar este tipo específico de aplicaciones. Para plantear estos criterios se basaron en un estudio de literatura.

Las cinco dimensiones son: Eficiencia, Efectividad, Confiabilidad, Facilidad de Aprendizaje y Satisfacción de Usuarios, cada una con diferentes sub-criterios que los autores consideraron importantes al momento de medir la usabilidad de estas aplicaciones móviles. Sin embargo en este estudio no prueban las dimensiones, solo es una propuesta en la que los autores fueron muy cuidadosos al momento de indagar en estudios previos, revistas y conferencias. Este estudio fue considerado porque respondía algunas preguntas de investigación sobre los métodos, modelos, atributos y métricas de evaluación de usabilidad.

EP3 (Faliagka, Rigou, & Sirmakessis, 2014): A usability study of iPhone built-in applications

Este estudio se enfoca en la usabilidad de aplicaciones nativas de iPhone, donde se identifican los principales problemas reportados, por medio de una evaluación de cinco aplicaciones nativas: Contactos, Reloj, Cámara, Calendario y Safari. El método

usado es una evaluación de laboratorio, que es un experimento conducido con un equipo especial, donde a los participantes les dieron un dispositivo iPhone 4, para que completen cinco tareas, cada una enfocada en cada aplicación, mientras son preguntados por “think aloud” y grabados. Después de la evaluación, los usuarios resolvieron un cuestionario online. Esta evaluación es enfocada en los principios de Norman y Nielsen en el 2010, donde los atributos medidos son: visibilidad, retroalimentación, consistencia, operaciones no destructivas, descubrimiento, escalabilidad, confiabilidad. Entre los resultados y problemas más frecuentes se encuentra que estas aplicaciones nativas del iPhone no cumplen con el principio de visibilidad ni con el de retroalimentación.

Los resultados del estudio confirman la lista de problemas de usabilidad reportados en los estudios, y hacen énfasis en la importancia de conformar principios de usabilidad. Los resultados también indican que cada diseño de interacción y plataforma móvil difiere en gran medida, a pesar de que cada dispositivo es necesario que cumpla con principios de usabilidad independientes de la plataforma se sigue observando en las últimas décadas su efecto heterogéneo.

EP4 (Joyce, Lilley, Barker, & Jefferies, 2015): Smartphone application Usability Evaluation: The Applicability of Traditional Heuristics

En este estudio los autores investigan la aplicabilidad de las heurísticas tradicionales de Nielsen, en una evaluación de usabilidad de aplicaciones móviles, para cumplir este fin realizan un estudio y análisis de cada heurística propuesta por Nielsen.

Los resultados indican que este conjunto de heurísticas pueden permitir el descubrimiento de cuestiones de usabilidad en una etapa temprana del ciclo de vida de la aplicación móvil, pero también no toman en cuenta áreas vitales de las aplicaciones móviles como su ambiente, el contexto de uso, el ingreso de contenido y el uso del micrófono, cámara y sensores.

Sin embargo, una de las ventajas de este método de evaluación heurística es que es de bajo costo, efectivo y relativamente es un procedimiento rápido de evaluación de usabilidad. En el presente estudio los autores argumentan que sin modificación alguna del método podría no ser efectivo cuando se evalúa aplicaciones móviles, pero sí de define un conjunto de heurísticas eficaces en la evaluación de usabilidad para aplicaciones en dispositivos móviles podría permitir a los evaluadores descubrir un

número considerable de problemas de usabilidad antes de la aplicación de métodos más costosos usados en las últimas etapas del ciclo de vida. Lo ideal sería que el conjunto de heurísticas sean económicas, rápidas de usar y fáciles de aprender.

EP5 (Saleh, Binti Isamil, & Binti Fabil, 2015): Extension of PACMAD Model for Usability Evaluation Metrics Using Goal Question Metrics (GQM) Approach

Los autores del estudio proponen el modelo PACMAD, para evaluar la usabilidad, el mismo que es probado en dos aplicaciones web. PACMAD fue presentado como modelo integral para evaluar la facilidad de uso de aplicaciones móviles basado en los siguientes factores de usabilidad: usuarios, contexto de uso, y tareas.

PACMAD, extiende los modelos de usabilidad tales como: Nielsen y la ISO, para el contexto de aplicaciones móviles. El modelo PACMAD posee siete componentes: Efectividad, Eficiencia, Satisfacción, Facilidad de Aprendizaje, Facilidad de Recuerdo, Errores y Carga cognitiva. Cabe especificar que la Carga cognitiva que se refiere al procesamiento cognitivo que el usuario necesita utilizar la aplicación. Se puede medir a través de una herramienta de evaluación: el Índice de Carga de Tareas de la NASA. En la Figura 8 se puede observar los atributos del modelo PACMAD.



Figura 8. Atributos de Usabilidad del Modelo PACMAD

Fuente: (Saleh, Binti Isamil, & Binti Fabil, 2015)

En el estudio utilizaron los métodos de evaluación de usabilidad: observación y un cuestionario. El experimento se dividió en dos partes; primero recogieron los datos objetivos a través de la observación y, a continuación recogieron los datos subjetivos a través de un cuestionario.

Para desarrollar las métricas de usabilidad los autores del estudio usaron el modelo GQM (Goal Question metric), métrica pregunta-objetivo, es una formación jerárquica para evaluar aplicaciones móviles, que está basado en un método de manejo de

objetivos para desarrollar las métricas. Los resultados de este estudio muestran que el modelo PACMAD, ayuda al desarrollador de aplicaciones móviles, en diferentes etapas del proceso de desarrollo para evaluar la usabilidad de la aplicación.

EP6 (Cruz Zapata, Fernández, Idri, & Toval, 2015): Empirical Studies on Usability of mHealth Apps: A Systematic Literature Review

Este estudio investiga los procesos de evaluación de usabilidad empíricos, a través de una revisión sistemática de literatura. Un total de 22 estudios seleccionados y relacionados con las aplicaciones móviles de salud son analizados.

Los autores investigan cuatro métodos empíricos de evaluación de usabilidad: cuestionarios, entrevistas, registros y “think aloud”. Los cuestionarios y entrevistas son métodos similares ya que ambos consisten en un conjunto de preguntas que se basan en recuerdos subjetivos de los usuarios y escalas de calificación.

El protocolo “Think aloud” es un método directo por el cual se puede evaluar la usabilidad de una aplicación. Al aplicar este protocolo los datos verbales obtenidos y las interacciones de los usuarios son transcritas y registradas para su análisis, generando un informe de respuestas no estructuradas motivo por el cual este método ha sido objeto de críticas en cuanto a su fiabilidad y validez.

En la literatura los autores obtuvieron varios hallazgos: el atributo atraktividad en el modelo ISO 9126-1 es la característica menos evaluada, la misma que no debería ser ignorada ya que causa un efecto positivo en la usabilidad y afecta en la decisión de empezar a usar la aplicación. Una recomendación de los autores de este estudio es que el desarrollo de aplicaciones móviles debería ser más enfocado al tipo de usuarios finales por ejemplo los adultos mayores, que tienen más dificultades de aprender a usar dispositivos móviles y sus aplicaciones.

Los resultados del estudio también mostraron que la falta de procesos iterativos de diseño-evaluación provocan los problemas de usabilidad y que la mayoría de aplicaciones evaluadas son de la plataforma Android. Por otro lado en este estudio se descubrió que los métodos de evaluación de usabilidad empíricos empleados, podrían mejorarse mediante la adopción de mecanismos automatizados y que los procesos de evaluación también deben ser investigados para la combinación de más de un método.

EP7 (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015): User Experience Evaluation of Mobile Spiritual Applications for Older People: An Interview and Observation Study

En este estudio se realiza una evaluación de usabilidad de una aplicación móvil espiritual para adultos mayores. Los métodos usados por los autores para la evaluación de usabilidad son: una entrevista, la cual es conducida con ocho personas adultas mayores, combinado con un estudio de campo y el protocolo “think aloud”.

La entrevista, fue realizada por los participantes que fueron usuarios reales de la aplicación desde su casa o lugar de trabajo, es decir, el estudio fue conducido en un ambiente natural donde la gente mayor se sentía comfortable. A continuación, el estudio se dividió en dos sesiones; la primera sesión es la entrevista a las personas mayores para reunir el perfil demográfico y las experiencias con aplicaciones espirituales móviles. En la segunda sección los usuarios comenzaron a usar la aplicación móvil espiritual llamada “iQuran”, guiados por un conjunto de tareas, también se les pidió a los usuarios que "piensen en voz alta" durante las sesiones de prueba. Finalmente los investigadores realizaron entrevistas a los usuarios para obtener sus experiencias y conocer sus opiniones o sentimientos después de usar aplicación móvil espiritual.

Los autores concluyeron que las actividades de entrevistas con gente adulta mayor para evaluar la usabilidad es una tarea compleja, por lo que es necesario aplicarse en una forma diferente y apropiada. Por otro lado las actividades de evaluación realizadas en un ambiente real y usuarios reales facilita la investigación sobre la usabilidad de aplicaciones espirituales.

EP8 (Tomé Klock & Gasparini, 2015): A Usability Evaluation of Fitness-Tracking Apps for Initial Users

Este estudio presenta los problemas de usabilidad encontrados en cinco aplicaciones móviles para el monitoreo de la actividad física de las personas, estos inconvenientes son descubiertos a través de una evaluación heurística basada en criterios de ergonomía.

Los criterios ergonómicos son una guía para diseñar y evaluar la interfaz de usuario, y cada criterio consiste en una definición, una razón de ser, ejemplos y algunos comentarios para evitar la ambigüedad. Bastien y Scapin crearon un conjunto de

dieciocho criterios ergonómicos y los ocho principales usados en este estudio son: (1) Orientación, (2) la carga de trabajo, (3) el control explícito, (4) Adaptabilidad, (5) Gestión de errores, (6) La consistencia, (7) Significado de los códigos y (8) Compatibilidad.

Para conducir la evaluación heurística se determinó cinco aplicaciones gratuitas de seguimiento de actividad física más descargadas en Play Store hasta la segunda mitad del año 2014. Las aplicaciones son: RunKeeper, Nike + Running, Runtastic, Runtastic Pedometer y Endomondo Sports Tracker.

Con estas aplicaciones seleccionadas para la evaluación, una lista de quince actividades principales fue creada. Los resultados de estas evaluaciones heurísticas indican que se debería realizar más evaluaciones de usabilidad que involucren a usuarios a través de experimentos controlados, para de esta forma entender los problemas de usabilidad relacionados con los usuarios y las consecuencias de estos problemas.

EP9 (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015): Usability Evaluation Method For Mobile Learning Application Using Agile: A Systematic Review.

En este estudio los autores revisan previas investigaciones que han sido realizadas para evaluar la usabilidad de aplicaciones móviles de aprendizaje, y cómo estas mejoras pueden ser integradas en un proceso de desarrollo ágil, para obtener aplicaciones más efectivas y usables.

Los autores observaron que las pruebas de usabilidad móvil se realizan mediante diferentes enfoques heurísticos, un enfoque incluye estudios de campo y el otro incluye el uso de diseñadores experimentados. Estos dos enfoques pueden lograr tener más iteraciones en el ciclo de desarrollo ágil y con la ayuda de especialistas experimentados, diseñadores y usuarios, se podría comprobar la facilidad de uso en una iteración antes de desarrollar la siguiente iteración. Con este enfoque, la evaluación de usabilidad y procesos ágiles de desarrollo son compatibles y pueden trabajar juntos.

En el estudio los autores encontraron información sobre un marco de diseño llamado CRUISER “Cross-discipline User Interface design and Software Engineering”, que integra la evaluación de la usabilidad dentro del proceso de ciclo de vida del producto software, es decir, integra la ingeniería de software con la interacción

humano-máquina (HCI) en un proceso ágil de desarrollo. Esto se realiza mediante el uso de los primeros prototipos o primeros diseños, que luego son refinados en las etapas posteriores del ciclo de vida del sistema. Se encuentra también que, la usabilidad de una aplicación es aún mejor con la iteración de la evaluación de usabilidad durante todas las etapas del desarrollo, esto implica desde el requisito inicial de la producción del sistema.

La incorporación de un proceso de desarrollo ágil y una evaluación de usabilidad en las aplicaciones móviles, ha producido un impacto significativo para que las aplicaciones móviles de aprendizaje sean más amigables y fáciles de usar.

EP10 (Hsiu Ching, 2015): Developing Mobile Application Design of Virtual Pets for Caring for the Elderly.

En este estudio se desarrolla una aplicación con funciones de cuidados de la salud y acompañamiento para ancianos. El diseño incluye una mascota virtual para asistir al usuario de la aplicación móvil. Para el desarrollo de la aplicación móvil, los autores realizaron un estudio de las preferencias de este tipo de usuarios en cuanto a la interfaz gráfica, y modelos de interacción para ancianos, con el fin de convencerlos sobre la facilidad de uso de esta aplicación.

Después del desarrollo de esta aplicación, los autores evaluaron la usabilidad de la misma para lo cual, primero establecieron tareas a los usuarios finales. Posteriormente de completar las tareas, un cuestionario fue respondido por los usuarios, con preguntas sobre la completitud de tareas, y opiniones de la aplicación. Para puntuar cada pregunta del cuestionario ellos utilizaron la escala Likert.

La usabilidad fue puntuada basándose en los principios de Nielsen: facilidad de aprendizaje, eficiencia, satisfacción, tasa de errores y facilidad de recordar.

En este estudio, se considera las funcionalidades de cuerpo y mente de los ancianos, por lo que toman en cuenta: funciones motrices, función de percepción, función cognitiva. Al igual que se considera el principio de diseño de interfaz de aplicaciones móviles, propuesto por Shneiderman, que plantea un conjunto de principios que son: (1) Luchar por la coherencia. (2) Permitir a los usuarios frecuentes usar atajos. (3) Ofrecer retroalimentación informativa. (4) Diseño de diálogo para dar cierre. (5) Ofrecer un manejo simple de error. (6) Permiso fácil para la reversión de

acciones. (7) Apoyar centro de control interno. (8) Reducir la carga de la memoria a corto plazo.

Después de la evaluación se considera que el efecto de la aplicación en los usuarios es generalmente satisfactorio, también que las imágenes marcadas con textos son más fácilmente reconocidos por el usuario y mejoran la eficiencia de la interfaz de usuario. En general en este estudio, los resultados encontrados podrían ser de ayuda a los diseñadores y desarrolladores de aplicaciones móviles para desarrollar una interfaz interactiva para los ancianos.

EP11 (Choi, Suk Lee, & Ho Park, 2015): Usability of Academic Electronic Medical Record Application for Nursing Students' Clinical Practicum

En este estudio se evalúa la usabilidad de una aplicación llamada AEMR (Academic Electronic Medical Record), una aplicación permite el acceso de los estudiantes de enfermería a la documentación de los registros para la práctica clínica.

Para la evaluación de la usabilidad de la aplicación, a cada uno de los usuarios, en este caso los estudiantes, le proporcionaron quince minutos introductorios de una vista general de cómo usar la aplicación. Posteriormente a los usuarios les asignaron quince tareas y mientras eran realizadas se aplicó el método "Think Aloud" para medir la habilidad del usuario para interactuar con la tecnología preguntándoles que están pensando sobre la interacción con la tecnología. La duración de cada tarea fue revisada y la interacción de la aplicación con el usuario fue grabada.

Como resultados de esta evaluación obtuvieron que: las oportunidades de documentación durante la entrevista con el paciente eran valiosos, la aplicación requiere un esfuerzo mínimo de usuario para acostumbrarse a la misma, los iconos de la aplicación fueron intuitivos, característica que ayudó a los usuarios a encontrar la información mucho más rápido.

EP12 (Foo, y otros, 2015): Mobile task management tool that improves workflow of an acute general surgical service

El objetivo de este estudio fue evaluar el impacto de una herramienta de gestión de tareas móvil en el flujo de trabajo clínico, dentro de un servicio de cirugía general mediante el análisis de la captura de datos sobre la usabilidad de la aplicación.

La aplicación evaluada es Cortex en la plataforma iOS, que fue desarrollada para digitalizar el flujo del paciente y proporcionar visibilidad en tiempo real sobre la toma de decisiones clínicas y la ejecución de tareas.

La usabilidad de esta aplicación fue medida a través de un cuestionario electrónico, que debía ser respondido en un periodo de tiempo específico, donde se requería que los participantes indiquen que tipo de usuarios son: consultor senior/junior, registrador o funcionario, y su nivel de satisfacción con la aplicación. La encuesta consistió en un conjunto de preguntas puntuadas a través de la escala Likert de cinco puntos. Las respuestas van desde muy desacuerdo hasta muy de acuerdo.

Los resultados de la encuesta han demostrado que el sistema fue bien recibido por el personal subalterno clínico tanto en su concepto y en la práctica. En conclusión, este estudio demostró que es posible digitalizar el flujo de trabajo de un equipo quirúrgico. Esta aplicación móvil se ha descubierto que no es sólo una herramienta útil para el usuario final, sino que también captura la información crucial para el hospital, a un nivel operativo que permitirá la mejora en la prestación de asistencia a los pacientes.

EP13 (Nawi, Isa Hamzah, Chy Ren, & Tamuri, 2015): Adoption of mobile technology for teaching preparation in improving teaching quality of teachers

El objetivo de este estudio fue identificar la disposición de los maestros para utilizar los teléfonos móviles en la preparación de la enseñanza. El estudio también examinó el nivel de satisfacción de los maestros al utilizar las aplicaciones de tecnologías móviles desarrolladas para el propósito de la enseñanza y el aprendizaje en el aula.

En este estudio se utilizó un método mixto para recoger datos. Se realizó un cuestionario de satisfacción y entrevistas. Para efectos de interpretación de la evaluación de aplicaciones, los investigadores utilizaron una escala Likert de cinco puntos.

Un total de 31 maestros participaron en responder el cuestionario y siete profesores fueron entrevistados para obtener datos de apoyo. El cuestionario de satisfacción incluía una serie de criterios incluyendo: sentirse feliz usando la aplicación, ahorrar tiempo, si la aplicación es de acuerdo con las necesidades del usuario, si obtuvieron una mayor comprensión del tema, si les gustaría utilizar la aplicación y si la recomendarían a otros posibles usuarios.

En los resultados de este estudio, los usuarios coincidieron en el tema de la edad de los profesores, el grupo afirmó que la aplicación es inapropiado para los profesores más adultos, porque consideran que es más complejo para ellos utilizar los teléfonos móviles para la preparación de la enseñanza, debido al pequeño tamaño de la fuente de texto de las aplicaciones.

Los resultados también muestran que el uso de las aplicaciones para enseñanza en los teléfonos móviles puede ayudar a los maestros a suavizar la preparación de las clases. En general, todos los criterios de la satisfacción de los usuarios tenían puntajes muy altos, que indican que los encuestados coincidieron en que están satisfechos con esta aplicación.

EP14 (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014): Usability testing of a monitoring and feedback tool to stimulate physical activity

Los autores de este estudio evalúan la usabilidad de la aplicación web móvil “It’s life”, que es una herramienta de seguimiento y retroalimentación para estimular la actividad física. En este estudio la usabilidad fue evaluada en un método mixto en las siguientes cuatro fases:

1. Evaluación heurística: esta fase requirió la ayuda de seis expertos tecnológicos como evaluadores. Los expertos iniciaron con la lectura del manual de usuario donde obtuvieron una idea general de la aplicación y después de fueron a fondo estudiando cada pantalla y anotando sus observaciones y puntuando diez heurísticas de usabilidad. Los resultados de esta fase se utilizan para ajustar el manual de usuario y desarrollar un segundo prototipo de la aplicación.

2. Evaluación de laboratorio: esta fase requirió la ayuda de cinco pacientes que debían realizar el procedimiento de “Think Aloud” para evaluar la capacidad de uso del manual, la aplicación y el servidor. Finalmente todos los participantes completaron un cuestionario de usabilidad del sistema.

3. Evaluación de campo: esta evaluación se realizó con 20 pacientes o en un contexto real. Después de cada interacción, los participantes fueron entrevistados acerca de sus experiencias. Las entrevistas fueron grabadas en audio y transcritas. Al final del período de intervención, los participantes completaron un cuestionario.

4. Segunda evaluación de laboratorio: esta fase de evaluación se la realizó con usuarios finales, en un laboratorio. Finalmente un post estudio de la usabilidad del sistema fue completado por medio de un cuestionario.

Los resultados de esta evaluación combinada han demostrado la importancia de un enfoque de métodos mixtos al medir la usabilidad de aplicaciones móviles, ya que se dieron a conocer cuestiones diferentes en el laboratorio en comparación con la prueba de la vida real. Las interfaces pueden ser muy eficaces, eficientes y deseables en una situación de laboratorio, pero si falla la comunicación entre los diferentes componentes de la herramienta en una situación de la vida real, la satisfacción será baja.

EP15 (Tsiaousis & Giaglis, 2014): Mobile websites: Usability evaluation and design

Este estudio muestra un modelo que formula una hipótesis sobre los efectos del contexto ambiental en la usabilidad de aplicaciones web móviles. Para validar esta hipótesis empíricamente, los autores realizan una evaluación de usabilidad de laboratorio en donde son simuladas las distracciones del ambiente.

En la evaluación de usabilidad se pretendió simular un contexto de uso real, cada participante recibió un pequeño tutorial del uso de la aplicación después de cada sesión de evaluación, a los participantes les preguntaron sobre el grado en que ellos consideraron compleja cada tarea por medio del cuestionario NASA-TLX, mientras el número de errores cometidos, la completitud de las tareas, el tiempo que tomaba realizar las tareas fue grabado en tiempo real.

También a los participantes se les preguntó su grado de satisfacción en el uso del sitio web móvil usando el cuestionario SUMI. Finalmente se pidió a los usuarios evaluar las condiciones del contexto simulados, con el fin de validar que son percibidos como estaba previsto.

Los resultados de este estudio muestran que a través de un experimento de laboratorio que simula las características de uso reales, se puede diseñar y controlar las distracciones durante las sesiones experimentales. En la literatura de HCI, se reconoce el impacto del contexto ambiental como atributos en usabilidad móvil, sin embargo falla al dar un nivel medible de influencia de cada tipo de distracción y la mayoría de estudios no capturan o controlan el contexto de uso.

EP16 (Hussain, Abdul, & Zainol, 2014): A usability testing on jFakih Learning Games for hearing impairment children

Este estudio evalúa la usabilidad y aceptación de una aplicación móvil llamada “JFakih”, que es un juego de aprendizaje diseñado para niños con discapacidad auditiva. La prueba de usabilidad se llevó a cabo con diez niños de edad entre 9-15 años, con discapacidad auditiva. Los autores de este estudio emplean varias técnicas de evaluación: cuestionarios, observación y entrevistas, basado en nueve elementos de aceptación. Los cinco atributos de usabilidad propuestos por Nielsen son usados y se derivan con otros elementos que han sido determinados en previos estudios. Los elementos son: facilidad de uso, habilidad de aprendizaje, ayuda, eficiencia, disfrute del usuario, rendimiento del sistema, nivel de juego, navegación y amigable al usuario. Las técnicas para medir usabilidad fueron empleadas de la siguiente forma:

Observación: con el fin de crear un ambiente confortable y agradable para los niños, la evaluación fue llevada realizada en sus aulas de clases. Los niños interactuaron con la aplicación y aspectos de usabilidad fueron observados por un evaluador.

Entrevista: los niños fueron guiados por un instructor para responder varias preguntas relacionadas a su satisfacción y opinión sobre el juego, sus respuestas y opiniones fueron grabadas.

Cuestionario: basado en la entrevista y observación, los investigadores usaron cuestionario SUMI “Inventario de medición de usabilidad de software” para puntuar los elementos mencionados anteriormente.

Los resultados muestran que esta aplicación de aprendizaje es atractiva y usable para niños con discapacidad auditiva, sin embargo, se requiere una mejora para incrementar el disfrute de los usuarios mientras están jugando el juego.

EP17 (Skorin-Kapov, Dobrijevic, & Piplica, 2014): Towards evaluating the quality of experience of remote patient monitoring services: A study considering usability aspects

En este estudio los autores estudian los factores y métricas para medir la QoE “Calidad de experiencia” basado específicamente es aspectos de usabilidad, para una aplicación móvil para el servicio de monitoreo de pacientes.

El estudio es realizado con 26 usuarios con un prototipo de la aplicación de salud “EMH” Ericsson Mobile Health, que se basa en el control de signos vitales a través de sensores médicos. Las técnicas de evaluación de usabilidad aplicadas en este estudio son: “Think aloud” y cuestionarios.

Los autores aseguran que una de las ventajas de usar “think aloud”, para evaluar la usabilidad es que provee las opiniones más honestas de los usuarios durante el uso de la aplicación. En cuanto a las dimensiones medidas o evaluadas en este estudio son las propuestas por Nielsen.

En este estudio los autores consideran que existen factores que influyen en los usuarios al momento del uso de una aplicación de tipo móvil como: los relacionados con la calidad de servicio de red, almacenamiento y visualización de las mediciones de los pacientes, los relacionados al diseño de interfaz del paciente/médico, los relacionados con el contexto de uso y los factores relacionados con el usuario individual (por ejemplo, motivación, experiencias previas). En consecuencia, se sostienen que existe una necesidad de estudiar y modelar la relación entre los diversos factores de influencia percibida por el usuario.

Los resultados de esta evaluación muestran la fuerte correlación entre calidad de experiencia con: la efectividad percibida de la interfaz móvil, la facilidad percibida de llevar a cabo una tarea, y la motivación para el uso de este servicio.

Un nuevo examen de los factores de influencia relacionados con la historia y el estado de salud de los usuarios se necesita en futuros estudios, para poder justificar las declaraciones relativas a la complejidad de la tarea de usuarios individuales.

EP18 (de Lima Salgado & Pimenta Freire, 2014): Heuristic evaluation of mobile usability: A mapping study

En este artículo los autores realizan un estudio sistemático de mapeo sobre un método en particular de evaluación de usabilidad, este es el método de evaluación heurística.

En este estudio se investiga y analiza sobre cuáles son las técnicas más usadas de evaluación heurística en los últimos años, el número de evaluadores que conducen esta técnica de evaluación en los estudios, entre otras características importantes de este tipo de evaluación.

Las heurísticas tradicionales más usadas en las evaluaciones de usabilidad de aplicaciones móviles son las propuestas por Nielsen y Molich, pero las propuestas de nuevas heurísticas para interfaces móviles crece sustancialmente.

La evaluación heurística propuesta por Nielsen y Molich, ha sido definida como un método de inspección de bajo costo y útil para evaluar la usabilidad del software. Este método requiere que evaluadores inspeccionen algunos elementos específicos de la interfaz de usuario, siendo guiados por tareas pre definidas y comparan la interfaz con una lista de heurísticas, con el fin de encontrar problemas de usabilidad.

Por otro lado en cuanto a la pregunta planteada por los investigadores sobre cuántos evaluadores es adecuado que lleven a cabo el proceso de evaluación heurística, los autores encontraron que en la mayoría de estudios el número de evaluadores esta entre tres o cuatro. También se analizaron las escalas de gravedad de los inconvenientes reportados en los estudios y un total de nueve estudios de investigación clasifica la gravedad de los problemas utilizando la escala propuesta por Nielsen, que van desde 1 = cosmético, 2 = menor, 3 = mayor y 4 = importante catastrófico. Dos estudios en cambio usan la escala Likert y tan solo uno de los estudios categoriza las severidades de los problemas encontrados en: baja, media o crítica. Los otros estudios no proporcionaron detalles sobre el índice de gravedad o escala utilizada.

Los autores encontraron que el uso de métodos de evaluación heurística ordinaria quizá no siempre es apropiado para evaluar todo tipo de interfaces móviles y quizás no abarquen características específicas de la interfaz móvil.

EP19 (Mirkovic, Kaufman, & Ruland, 2014): Supporting cancer patients in illness management: Usability evaluation of a mobile app

Los autores del estudio realizan una evaluación de usabilidad de una aplicación móvil para el control de la enfermedad de pacientes con cáncer llamada “Connect Mobile”.

El estudio de usabilidad es realizado con un prototipo de la aplicación con siete usuarios finales, donde los objetivos de la evaluación fueron identificar cuestiones de diseño, funcionalidad y satisfacción con la aplicación. También se conduce unas entrevistas para obtener: una retroalimentación sobre la usabilidad de la aplicación, para identificar la necesidad de nuevas características para el nuevo sistema en cuanto

a requerimientos de diseño de la interfaz, la aceptación de la aplicación móvil y sus características en el control diario de la salud.

Los usuarios realizan un conjunto de tareas para posteriormente resolver un cuestionario estandarizado que mide la percepción de los usuarios. Las heurísticas usadas en la evaluación de usabilidad son las propuestas por Bertini, que son una modificación de las heurísticas de Nielsen con el objetivo de capturar factores contextuales en la computación móvil.

En el estudio revelaron un total de 27 cuestiones de diseño de interfaz, la mayoría de violaciones estaban relacionados con la facilidad de entrada y la legibilidad de la pantalla. Los comentarios de los participantes también indicaron los requisitos de los pacientes con cáncer para los sistemas de apoyo y cómo estas necesidades se ven influidas por diferentes factores relacionados con el contexto.

Por otro lado se encontró que al utilizar la aplicación móvil los participantes que son dueños del mismo tipo de dispositivo móvil y tienen previa experiencia con el mismo, tienen un incremento en el nivel de autoconfianza y menos peticiones de ayuda mientras realiza las tareas. En cambio quienes no tenían o era limitada la experiencia con este tipo de dispositivos, necesitaron la orientación de rutina con la aplicación.

EP20 (Aizawa, y otros, 2014): Comparative study of the routine daily usability of foodlog: A smartphone-based food recording tool assisted by image retrieval

En este estudio se realiza un estudio comparativo de la aplicación “FoodLog”, que es un registro de alimentos con dos tipos diferentes de interfaces de usuario: una basada en imágenes y otra en modo de texto. Los autores realizan una evaluación para conocer cual tipo de interfaz es más usable, mediante el método de cuestionario.

Se reclutaron a 18 estudiantes universitarios (edad = 18-24 años), los cuales llevan a cabo la grabación de alimentos por primera vez. Las participantes utilizaron “FoodLog” sobre una base diaria durante 1 mes. Los participantes fueron al laboratorio e instalaron la aplicación con asistencia basada en imágenes o aplicación solo basada en ingreso de texto.

Ellos usaron las aplicaciones en su vida diaria, después de dos semanas fueron al laboratorio e intercambiaron la aplicación para probar ambas interfaces. En dos semanas más los participantes completaron un cuestionario, usando una escala de 5

puntos para la calificación: 5 = positivo, 4 = débilmente positivo, 3 = neutral, 2 = débilmente negativo, 1 = negativo.

Dentro de los resultados los autores observaron que en la evaluación subjetiva, la aplicación con interfaz basada en imágenes tuvo mayores relaciones con la facilidad de uso, la diversión, la frecuencia de la navegación, y la intención de continuar. En particular, los últimos tres factores recibieron puntuaciones significativamente más altas en comparación con la aplicación basada en textos.

La evaluación subjetiva demostró que esta nueva herramienta, que explota completamente el uso de imágenes, es un método prometedor para la grabación o registro de los alimentos.

EP21 (Kluth, Krempels, & Samsel, 2014): Automated usability testing for mobile applications

En este estudio los autores comparan una herramienta automatizada llamada AUT para evaluar la usabilidad de las aplicaciones móviles de la plataforma iOS con el método de evaluación de observación de usuarios. El método AUT no requiere de expertos, de manera que reduce costos, tiempo y gastos de personal.

Esta herramienta funciona con la captura automática de las interacciones de los usuarios mediante un framework, el cual crea un registro de la interacción de los usuarios en cada sesión y lo envía a un servidor central. Los problemas de usabilidad son reconocidos en una secuencia de interacciones con el sistema, donde se inspecciona que las características del sistema cumplan con los patrones de diseño de interacción especificados en el framework.

AUT está separada en cinco clases de métodos: evaluación, inspección, cuestionario, modelos analíticos y simulaciones. AUT posee cuatro diferentes pasos de automatización: no automática, captura automática, análisis automático y crítica automática. Cada paso de automatización es consecutivo, además el esfuerzo de AUT es estimado formalmente (tareas explícitas para participantes) e informalmente (participantes usan objetivos del sistema sin ninguna tarea adicional).

Las fases del método AUT son: (1) Fase de preparación, (2) Fase de captura, (3) Fase de Análisis y (4) Fase Crítica.

La herramienta AUT utiliza cuatro patrones de diseño de HCI: Ley de Fitts, Entrada errónea, Carga Navegacional y Toque Accidental. Los autores en el presente

estudio comparan los métodos de observación de usuarios y la herramienta AUT, evaluando el prototipo de la aplicación “Bike Sharing”.

Entre los resultados encontrados en este estudio comparativo están que en el clásico método observación de usuario, se hace difícil capturar todos los ingresos de datos en la interfaz de la pantalla táctil y la obtención de datos es tedioso porque hay que clasificar manualmente los problemas dentro de los patrones de diseño de interacción, en cambio AUT genera automáticamente los registros de las interacciones y da una visión en conjunto de todos los problemas por cada patrón con un gráfico de interacción apropiado.

EP22 (Kalz, y otros, 2014): Smartphone apps for cardiopulmonary resuscitation training and real incident support: A mixed-methods evaluation study

En este estudio se realiza una evaluación de usabilidad de una aplicación de resucitación cardiopulmonar por medio del método de cuestionario SUS. Los autores conducen una evaluación secuencial combinando la perspectiva de médicos expertos y usuarios finales.

La aplicación evaluada es CPR “cardiopulmonary resuscitation”, dos expertos médicos evalúan el contenido médico de esta aplicación de las plataformas móviles “Android” y “iOS”, en una segunda fase se realiza la evaluación de usabilidad con el cuestionario SUS, con usuarios finales de las aplicaciones.

La escala de sistema de usabilidad (SUS) es una herramienta simple para evaluar la usabilidad de un diverso conjunto de tecnologías. Consiste en diez preguntas con una escala Likert de cinco puntos, donde las direcciones de los ítems son cambiados con cada pregunta. Un valor específico de usabilidad fue calculado por cada aplicación. También los autores evalúan la calidad hedónica (hedónico que causa placeres satisfacción), con la herramienta “ReactionDeck”, que es un conjunto de herramientas que se basa en la evaluación de aspectos como "deseos" y "diversión" de los usuarios con los productos.

Los resultados de este estudio muestran que la evaluación que se realizó y sus resultados son un primer paso para optimizar el desarrollo y la evaluación de aplicaciones móviles para la reanimación y apoyo en incidentes reales, ya que en este estudio prestan particular atención en el contenido, así como la facilidad de uso y la calidad hedónica. Sin embargo, una desventaja de este estudio es que los autores no

podieron sacar conclusiones acerca de la eficacia de la aplicación en una situación de emergencia real.

EP23 (Olsina, Santos, & Lew, 2014): Evaluating mobile app usability: A holistic quality approach

En este estudio realizan una evaluación de usabilidad de la aplicación móvil “Facebook”, a través de un enfoque integral de calidad para la evaluación de la usabilidad y la experiencia del usuario de aplicaciones móviles, ya que con la literatura investigada consideran que estas no están bien relacionadas hacia las metas del sistema y del sistema en uso.

Los autores caracterizaron y representaron las características más relevantes de usabilidad, de experiencia de usuarios y de entidades de contexto de uso, para establecer los atributos que serán medidos en la evaluación, los mismos que se basan en 2Q2U v2.0, que es un modelo de calidad que especifica sub-características de usabilidad.

Los inspectores fueron midiendo la usabilidad de la aplicación basado en las sub - características y atributos establecidos en una escala de 0 hasta 100% y utilizaron colores para definir la gravedad de los problemas.

Entre los resultados de esta evaluación encontraron que este enfoque holístico de evaluación de la usabilidad por medio de estrategias específicas contribuye a los diseñadores a entender los problemas de usabilidad con eficacia y a producir mejores soluciones de diseño.

Por otro lado, en las directrices de diseño y de interfaz de usuario en Apple y Google, la relación entre las entidades de las aplicaciones móviles con conceptos de usabilidad y experiencia de usuario no está definitivamente explícita en modelos, ni es representado en el trabajo de usabilidad, ni en otras investigaciones relacionadas con la calidad, por ejemplo, Nielsen enumera muchas características y listas de control de aplicaciones móviles en las que se requiere ciertos contextos de uso, pero no utilizan puntos de vista de calidad y métodos de modelización. Por lo tanto, la capacidad para una aplicación coherente utilizando un marco conceptual y estrategias para evaluar o mejorar una aplicación móvil son bastante limitados.

EP24 (Cruz Zapata, Hernandez Niñirola, Idri, Fernandez-Aleman, & Toval, 2014): Mobile PHRs compliance with android and ios usability guidelines mobile systems

En este artículo los autores evalúan la conformidad de las guías de diseño de interfaz de usuario de Android y las Guías de Interfaz Humana de iOS, con 24 aplicaciones PHRs (Aplicaciones Móviles de Registro de Salud Personal).

Este proceso se lleva a cabo a través de una revisión sistemática de literatura en donde se estudia y analiza las aplicaciones seleccionadas en el grado que cumplen con estos principios de diseño de interfaz.

Para esta evaluación utilizan el método cuestionario para medir la concordancia de las aplicaciones con los principios de diseño de interfaz, el mismo está compuesto de 13 preguntas, la evaluación es conducida por expertos.

Los resultados de usabilidad pueden estar sesgados en la dirección del interés general en el diseño de la aplicación, que no necesariamente corresponden a la opinión de los usuarios reales. Por otra parte, el cuestionario se basa en las directrices de diseño de iOS y Android, cuyo objetivo es aumentar la retención de usuarios y la satisfacción del cliente, buscando que sea más fácil que los usuarios aprendan a utilizar la aplicación de la forma más rápida e intuitiva como sea posible. Los resultados también muestran que un 46% de las 24 aplicaciones no usan ninguno de los principios recomendados.

EP25 (Yáñez Gómez, Cascado Caballero, & Sevillano, 2014): Heuristic Evaluation on Mobile Interfaces: A New Checklist

En este trabajo los autores proponen un checklist de heurísticas para medir usabilidad, basado en la compilación de evaluaciones heurísticas existentes en la literatura, para adaptarlas a la interfaz móvil. Con estas heurísticas se puede ofrecer una herramienta para trabajar mejor, no solo para la evaluación sino serviría también como una lista de mejores prácticas.

La propuesta de este checklist es probado experimentalmente, con la colaboración de dos ingenieros de software sin ningún conocimiento de usabilidad y un grupo de diez usuarios quienes comparan la usabilidad de un primer prototipo de una aplicación diseñada sin estas heurísticas propuestas, y una segunda después de aplicar el checklist propuesto.

La Evaluación heurística tiene una gran ventaja sobre las otras técnicas, su implementación es fácil, rápida y barata, también es adecuada para toda fase del ciclo de vida del software y no requiere planeación previa. Además no es mandatorio para los evaluadores que sean expertos en usabilidad, a diferencia de las pruebas de laboratorio que parecen incapaces de asegurar por completo la facilidad de uso en este paradigma móvil que es el contexto de uso, y muchas veces requieren expertos.

Los resultados muestran la utilidad de este checklist para evitar vacíos de usabilidad incluso con los desarrolladores no entrenados. Sin embargo la lista de comprobación heurística propuesta debe ser validada a fondo en investigaciones futuras en relación con diferentes aspectos.

EP26 (Lyles, Sarkar, & Osborn, 2014): Getting a Technology-Based Diabetes Intervention Ready for Prime Time: a Review of Usability Testing Studies

En este artículo los autores realizan una revisión de artículos publicados entre el 2009 – 2013, sobre evaluación de usabilidad de sistemas web o móviles para el apoyo de pacientes con diabetes.

Entre la literatura encontraron 23 artículos de los cuales: 11 evalúan la usabilidad con métodos mixtos, 9 evalúan cuantitativamente la usabilidad y 3 usan métodos cualitativos. De estos artículos se extraen las características específicas y resultados importantes sobre las evaluaciones y experiencias de los usuarios. En este estudio se muestra información importante para la presente tesis como por ejemplo los métodos utilizados por otros autores para medir la usabilidad entre ellos están:

Cuestionario de usabilidad: la encuesta o cuestionario de usabilidad está basado en una puntuación que los usuarios dan a la aplicación indicando, sus opiniones, la facilidad de uso percibida y su asistencia o experiencias usando una tecnología específica.

Evaluación de calidad: Una evaluación cualitativa puede detectar problemas de usabilidad no anticipados, que no se pueden descubrir en una medida de usabilidad cuantitativa. Casi la mitad (12 de 23, 52%) de los estudios incluidos utilizaron métodos cualitativos, que se centran específicamente en entrevistas en grupos o individuales.

Think aloud: es un método para evaluar la usabilidad del sistema o aplicación, en donde se les pide a los participantes utilizar una aplicación o sistema mientras piensa continuamente en voz alta.

Recorrido Cognitivo: Un recorrido cognitivo implica la evaluación de usar una aplicación por primera vez (sin entrenamiento formal) del sistema, en el que un experto (1) establece un objetivo a lograr, (2) realiza las búsquedas de acciones disponibles, (3) selecciona una acción o tarea que le parece más relevante, y (4) realiza la acción y evalúa el sistema entero con el fin de avanzar hacia la meta original.

En muchos de los estudios los autores optan por usar métodos combinados de evaluación de usabilidad, para cubrir todos los aspectos que consideran importantes, por ejemplo datos cuantitativos, cualitativos, experiencias y satisfacción de los usuarios.

Los autores de este estudio concluyeron que cada método de evaluación posee distintas ventajas y que es fundamental llevar a cabo una evaluación de usabilidad robusta usando una combinación de metodologías con el fin de lograr una evaluación integral, este estudio también afirmó que lo ideal sería seguir un proceso iterativo centrado en el usuario en el proceso de diseño del sistema.

EP27 (Maren Witt, 2014): A set of quantitative user experience metrics for multi-modal dialog systems

Los autores de este artículo presentan un framework para evaluar aplicaciones móviles de voz multimodales. Este framework permite el análisis de la experiencia del usuario a profundidad, para realizar mejoras en el diseño del sistema.

La usabilidad de este framework propuesto ha sido evaluada con datos de una aplicación móvil multimodal de voz. El proceso de evaluación se llevó a cabo en dos estudios: el primer estudio de usabilidad se realizó en una semana con 10 participantes, cada participante recibió instrucciones iniciales sobre como descargar y usar la aplicación. Se pidió a los participantes realizar un seguimiento de todos los alimentos que comían y el envío de mensajes de correo electrónico al director del estudio, con retroalimentación o los problemas suscitados al usar la aplicación. Al final de la semana cada participante realizó una encuesta, tomando en cuenta las métricas del framework propuesto.

La segunda evaluación de usabilidad consistió en 40 participantes quienes recibieron instrucciones de descarga y dos párrafos de descripción de la aplicación vía mail, en contraste con el primer estudio aquí no tuvieron oportunidad de recibir directamente la retroalimentación de usuarios.

Entre las métricas evaluadas se encuentran: métricas del rendimiento del sistema, métricas del comportamiento del usuario, desglosadas en atributos específicos. La aplicación móvil multimodal de voz evaluada es “Calio”, que se trata de una aplicación que permite a los usuarios realizar un seguimiento de su ingesta de alimentos. La comida puede ser buscada por cuatro diferentes vías: habla, forma de texto libre, por medio de escaneo de código de barras y usando favoritos o comidas previamente guardadas.

Este framework añade métricas de usuarios específicos que son muy útiles en la predicción de necesidades usuarios e identificando sus patrones de comportamiento.

EP28 (Vélez, Boakye, Kanter, & Bakken, 2014): A usability study of a mobile health application for rural Ghanaian midwives

En este estudio se realiza un estudio descriptivo de usabilidad del prototipo de la aplicación “mClinic”, que es una aplicación móvil para parteras en Ghana, que captura los datos para la gestión de la atención al paciente, como también la toma de decisiones. Este proceso de evaluación está conformado de tres fases:

(1) Evaluación híbrida en laboratorio en vivo, que identifica problemas de usabilidad. Los participantes completaron tres tareas especificadas en función de las metas y los casos de uso destinados para mClinic, mientras que un inspector fue tomando notas de la interacción de los participantes con la aplicación.

(2) Cuestionario de usabilidad después de terminar las tres tareas previstas para conocer la usabilidad percibida y el control del usuario.

(3) Finalmente una Entrevista, se llevó a cabo con cada una de las usuarias, también se les pidió expresar sus pensamientos acerca de la facilidad de uso y la utilidad del prototipo mClinic.

Los problemas de usabilidad identificados durante la realización de las tres tareas y durante las entrevistas se clasificaron utilizando una lista de heurísticas de usabilidad de Nielsen, para dispositivos móviles. Cada problema identificado se clasificó en una escala de 0 a 4.

Entre los resultados del estudio, encontraron que la evaluación heurística es un método comúnmente utilizado para la identificación de problemas de usabilidad. Las usuarias del sistema consideraron que mClinic era útil, fácil de usar y fácil de controlar.

El uso de la salud móvil puede proporcionar mecanismos para mejorar la eficiencia y eficacia de la atención prestada por una partera, mediante el apoyo en la toma de decisiones clínicas y la disponibilidad de recursos de información, mientras se reducen sus cargas administrativas.

EP29 (Hussain, Hashim, & Nordin, 2014): mGQM: Evaluation Metric for Mobile and Human Interaction

Este artículo examina cuestiones de usabilidad de aplicaciones móviles y propone un modelo basado en métricas que pueden ser usadas para evaluar aplicaciones de dispositivos móviles.

El proceso de creación del modelo se realiza primero con una Revisión Sistemática de Literatura (SLR) para extraer las características de usabilidad que deben tener las aplicaciones de dispositivos móviles.

GQM (Goal Question Metric), métrica pregunta-objetivo, ha sido seleccionado como metodología para derivar las métricas de usabilidad para aplicaciones móviles. Las dimensiones creadas en la SLR iniciarán el método de generación de preguntas y finalmente se derivarán en métricas que propone este modelo.

Para validar este modelo basado en métricas, se realizó un estudio de usabilidad de las aplicaciones Facebook y CoPilot en iOS y Android, que constaba de tres pasos: (1) Experimento, que fue realizado para reunir datos objetivos, (2) Cuestionario para obtener datos subjetivos y (3) Entrevista para obtener también datos subjetivos.

Los resultados mostraban que los participantes estaban felices con las aplicaciones en el iPhone en comparación con las del teléfono Android. El modelo que se presenta en este documento contribuye como una opción para el campo de la evaluación relacionada con los dispositivos móviles.

EP30 (Arnhold, Quade, & Kirch, 2014): Mobile applications for diabetics: A systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older

Los autores en este artículo muestran una evaluación de usabilidad de aplicaciones para el control de la diabetes disponibles para los sistemas operativos iOS y Android.

El método de evaluación de usabilidad es la heurística con expertos, basada en Nielsen y Barnum. Las bases para esta evaluación de usabilidad fueron definidas por

un conjunto de criterios especiales considerando el proceso de interacción, el diseño de interfaz y la comprensión del contenido. Los expertos calificaron cada sub-criterio y la expresión de sus características en una escala de cinco puntos Likert.

Dentro de esta evaluación de la usabilidad, se investigó exclusivamente los criterios de usabilidad. No se evaluó ni la calidad de los contenidos, ni funciones, ni su eficacia. Por otra parte, hay que mencionar que una evaluación de la usabilidad no puede pretender abarcar todas las situaciones posibles de uso que posiblemente pueden ocurrir.

La prevalencia de diabetes aumenta con la edad. Por lo tanto, los ancianos son un amplio grupo de usuarios que podrían beneficiarse de aplicaciones para la diabetes. Sin embargo, varios estudios han demostrado una falta de aceptación y un uso insatisfactorio de las tecnologías móviles innovadoras entre personas de este grupo de edad.

Además, las habilidades cognitivas y físicas van disminuyendo con la edad y el resultado de las necesidades son considerablemente diferentes a las de los usuarios jóvenes. Los resultados de las evaluaciones de usabilidad a las aplicaciones de diabetes elegidas muestran que para los usuarios o pacientes de 50 o más años de edad es moderadamente buena.

EP31 (Lira, de Souza, Ferreira, & Carvalho, 2014): Experimenting on the cognitive walkthrough with users

Este artículo presenta un caso de estudio cuyo objetivo es investigar qué variante del método de evaluación de usabilidad “Think aloud” (pensamiento en voz alta retrospectivo y pensamiento en voz alta concurrente) se integra mejor con el método de Recorrido Cognitivo con usuarios.

Se realizó una evaluación con doce usuarios quienes debían resolver tres tareas y con un evaluador u observador. La aplicación evaluada se llama “Plantas Digital”, sirve para la identificación o reconocimiento automático de una planta con el uso de imágenes de las hojas. El proceso de evaluación inicia con el método:

Recorrido cognitivo que provee un proceso de evaluación que consiste en la inspección de cada acción posible del usuario para realizar una tarea determinada, incluye cuatro heurísticas a ser medidas por un experto en usabilidad mientras se combina con el protocolo “Think aloud” y sus variantes, los evaluadores deben

observar y coleccionar datos desde el inicio hasta el final de la interacción de usuarios con el software evaluado. El usuario debe hablar todo lo que está pensando con relación al software de evaluación. Las variantes del protocolo “Think-aloud” se caracterizan por:

Pensar en voz alta Concurrente (Concurrent Think-aloud): los instructores ven al usuario hablar mientras está interactuando con el software.

Pensar en voz alta Retrospectivo (Retrospective Think-aloud): se instruye al usuario a hablar mientras mira un video de su interacción con el software.

Los resultados también indican que: el recorrido cognitivo con usuarios se integra bien con ambas variantes del protocolo “Think aloud”; el protocolo pensar en voz alta retrospectivo encuentra más problemas de usabilidad; y el protocolo pensar en voz alta concurrente es ligeramente más rápido de realizar y es más económico. Sin embargo esto es solo un caso de estudio y se requiere mucha investigación para verificar si estos resultados son estadísticamente significantes

EP32 (Beul-Leusmann, y otros, 2014): Usability evaluation of mobile passenger information systems

Los autores en este artículo investigan el prototipo de un sistema de información de pasajeros, a través de una evaluación de usabilidad. Este sistema es comparado con otra aplicación líder de movilidad en Alemania. Las dos aplicaciones son de la plataforma iOS, y el método usado para su evaluación es una evaluación de campo y de laboratorio, seguido de un cuestionario puntuado con la escala de usabilidad del sistema SUS.

Las condiciones en que las aplicaciones fueron evaluadas son: la primera dentro de un laboratorio y la segunda evaluación fue con las características de campo en donde los participantes usan la aplicación en el exterior mientras caminan.

Durante las evaluaciones los participantes fueron instruidos para “pensar en voz alta”, es decir, ellos tenían que mencionar comentarios espontáneos de la interacción.

En el laboratorio la interacción fue adicionalmente grabada con un sistema controlado de grabación, pero debido a las restricciones técnicas, esto no fue aplicable en el campo, por lo que un cuestionario fue realizado.

El cuestionario se realizó con la herramienta SUS o Escala de usabilidad del sistema, que es fácil de aplicar y posee diez ítems en los cuales los usuarios indican el grado de acuerdo o desacuerdo con los factores importantes de usabilidad establecidos.

Los resultados de esta evaluación muestran que la retroalimentación al usuario es beneficiosa en el aprendizaje sobre las expectativas de los usuarios, por otro lado encontraron que los participantes invierten más tiempo en realizar las tareas en un ambiente de campo que en la evaluación de laboratorio. La evaluación de la usabilidad reveló que los problemas encontrados son de fácil mejora, ya que son problemas de tipo estético y también se mencionaron problemas técnicos del servicio basado en la ubicación del sistema.

EP33 (O'Malley, Dowdall, Burls, Perry, & Curran, 2014): Exploring the usability of a mobile app for adolescent obesity management

El principal objetivo de este estudio es evaluar la usabilidad de una aplicación móvil llamada “Reactivate” para el manejo de la obesidad en adolescentes, basándose en las siguientes características: eficacia técnica, eficiencia y satisfacción de usuario.

El proceso de evaluación inicia con diez participantes que tienen obesidad, a quienes se les instruyó para que cumplan ocho tareas, con el fin de evaluar la eficacia técnica. Cada tarea requiere que el participante obtenga o introduzca datos específicos que se utilizaría en una tarea típica. La tarea se completó cuando el evaluador indica que se ha cumplido el objetivo de la tarea (ya sea con o sin éxito), o cuando el participante solicitó y recibió orientación suficiente para justificar la puntuación del escenario como un error crítico.

Con el fin de probar la eficiencia relativa del usuario al usar la aplicación se mide el tiempo medio que un usuario tarda en completar una tarea en comparación con un usuario experto de la aplicación.

En la evaluación de satisfacción de los usuarios, cada participante completó el inventario estandarizado de la medición de la usabilidad del software SUMI, que mide cinco aspectos de la satisfacción del usuario: eficiencia, efectos, utilidad, capacidad de control y facilidad de aprendizaje.

Todos los adolescentes completaron cada una de las tareas con éxito y manifestaron que utilizarían “Reactivate” para controlar su crecimiento en el tiempo,

para su motivación y para la fijación de objetivos, ellos describen a esta aplicación móvil como importante.

Los autores concluyeron que las pruebas de usabilidad son un paso importante para refinar el desarrollo de la aplicación “Reactivate”, que puede ser utilizada en el tratamiento de la obesidad.

EP34 (Reynoldson, y otros, 2014): Assessing the quality and usability of smartphone apps for pain self-management

En este artículo se evalúa dos aplicaciones seleccionadas para la autogestión del dolor, utilizando criterios de evaluación de calidad y pruebas de usabilidad con usuarios potenciales. Se ha contado con 41 participantes con experiencia de episodios crónicos o recurrentes de dolor.

Los autores midieron la facilidad de uso por medio de un cuestionario basado en heurísticas diseñadas para identificar defectos de usabilidad, empleando Likert y SUS, el diseño de interfaz y la usabilidad también fueron medidos después de ser descargadas las aplicaciones, usando los criterios de esta lista de heurísticas.

El cuestionario SUS (escala para medir usabilidad del sistema) es aplicado a los participantes, a partir de las pruebas de usabilidad, los participantes indican una preferencia por un diseño de la interfaz que emplea un esquema de color más claro y una fuente de texto en particular. La variación en la calidad de las aplicaciones y la falta el compromiso en el desarrollo clínico fueron inconvenientes encontrados a través de la evaluación de estas aplicaciones.

En conclusión se obtuvo información útil que podría ser beneficiosa, es necesario como desarrolladores de productos software involucrar a los usuarios en un proceso temprano de desarrollo, así como el establecimiento de formas de combinar las necesidades o requerimientos del usuario final con el contenido basado en la evidencia, para proporcionar aplicaciones de alta calidad y que pueden utilizarse para auto-tratamiento del dolor.

EP 35 (Biederer, Arguel, Liu, & Lau, 2014): From Web-Based to Mobile: Experiences of Developing a Personally Controlled Health Management System

El objetivo de este estudio fue desarrollar una versión móvil de "Healthy.me", un sistema de gestión de la salud controlada personalmente, basada en la versión Web existente de la aplicación y el análisis de su usabilidad.

La extensión móvil de PCHMS fue diseñada de forma iterativa, desde su versión basada en la web y se evaluó durante un periodo de seis meses con 17 participantes quienes completaron un estudio de usabilidad que incorpora el protocolo pensar en voz alta, actividades de tareas, ejercicios de cartas de clasificación, entrevistas y un cuestionario de aceptación de la tecnología, midiendo así la utilidad percibida y la facilidad de uso percibida de la versión móvil de la aplicación.

El estudio es realizado con las características de evaluación de laboratorio, combinado con el protocolo "Think aloud", es decir mientras los usuarios realizaban las actividades y tareas fueron observados e interrogados.

En la siguiente fase del estudio se estableció la "Tarjeta de clasificación" que es una actividad que se basa en clasificar las características favoritas de la versión móvil de PCHMS.

En el componente de la entrevista, los participantes respondieron preguntas sobre su experiencia general con el PCHMS móvil. Finalmente para medir la percepción de los participantes sobre la aceptación de la aplicación, se desarrollaron dos cuestionarios, basados en una versión extendida del Modelo de Aceptación de Tecnología TAM 2.

Como resultados de esta evaluación se encontró que la mayoría de los participantes definen a la aplicación como fácil de usar, por ello tuvieron poca dificultad al navegar por las diferentes funciones a través de los botones y fichas proporcionadas. Sin embargo los participantes mayores de 40 años de edad experimentaron más dificultades en completar las tareas de usabilidad.

EP 36 (Wein, 2014): Visual Recognition in Museum Guide Apps: Do Visitors Want It?

En este estudio, el reconocimiento visual (VisRec) se evalúa como un método para acceder a la información general sobre las obras de arte en una aplicación móvil de guía de museo, por medio de un experimento de campo.

Los objetivos de la investigación son evaluar la usabilidad de VisRec en comparación con los códigos QR y códigos numéricos, para identificar y explicar las

preferencias de los visitantes. En primer lugar, un prototipo fue desarrollado para demostrar la eficacia técnica de VisRec.

Una evaluación de usabilidad de campo es desarrollada en dos museos. Después de la visita y de que los usuarios hayan usado la aplicación, ellos calificaron con respecto a cinco dimensiones de usabilidad basadas en el cuestionario SUS, calificadas en una escala Likert de cinco puntos.

Los resultados ponen de manifiesto la dependencia del contexto de la utilidad percibida y la variabilidad en la importancia de los factores de usabilidad. Los resultados muestran una clara preferencia por VisRec entre los participantes (53%); sólo el 14% prefiere los códigos QR.

Los resultados reales revelaron una clara preferencia entre los participantes para VisRec en comparación con los códigos numéricos y códigos QR. Esto pone de relieve la madurez técnica de VisRec para su uso en la vida real y confirma la hipótesis de que VisRec podría ser un método más natural, intuitivo y discreto para interactuar con obras de arte durante una visita a un museo para obtener la información deseada.

EP37 (Mandia Grossi, Torres Pisa, & Marin, 2014): Oncoaudit: development and evaluation of an application for nurse auditors

El objetivo de este estudio es desarrollar una aplicación web en dispositivos móviles y evaluar su usabilidad. Esta aplicación sirve para buscar los medicamentos de quimioterapia y apoyar a las auditorías del hospital.

La aplicación web se evaluó por auditores de enfermería utilizando el cuestionario SUS y también se evaluó por informáticos profesionales de la salud a través de una evaluación heurística basada en las guías de Nielsen.

En la evaluación de satisfacción percibida del usuario, con los auditores de enfermería, se aplicó el cuestionario de la Escala de usabilidad del sistema SUS, que es un método ampliamente aceptado por su fiabilidad y validez. Este cuestionario contiene una simple escala de diez puntos, lo que permite la evaluación de percepciones subjetivas.

En la evaluación de la usabilidad heurística el evaluador inspeccionó la aplicación, utilizando las heurísticas como una guía para identificar posibles problemas y el siguiente paso fue clasificar la gravedad del problema.

La evaluación de la enfermera indicó que las versiones web y móvil abordan las necesidades del usuario. En la evaluación de la usabilidad, 14 problemas fueron identificados en la versión móvil y ocho en el sistema web. Se evaluó la herramienta con respecto a la satisfacción del usuario y facilidad de uso, lo que ayudó a promover mejoras en el software antes de que esté disponible al público. Incluso con el número limitado de los evaluadores, estaba claro que Oncoaudit puede ser usado en la práctica de la auditoría y hacer que este proceso sea más rápido y completo.

EP 38 (Luhur & Widjaja, 2014): Location-based social networking media for restaurant promotion and food review using mobile application

Este documento se centra en el desarrollo y evaluación de una aplicación móvil bajo la plataforma Android para buscar restaurantes y promociones, es basado en la localización y redes sociales.

Los métodos de evaluación que son utilizados en este proyecto son: la evaluación heurística y las pruebas de usabilidad que actúan como complementarias entre sí para averiguar los problemas de usabilidad que se pueden eliminar para aumentar la posibilidad de éxito en lugar de mostrar los resultados repetitivos.

La evaluación heurística es una inspección sistemática de un diseño de interfaz de usuario para la facilidad de uso, en donde interactúan los llamados evaluadores o expertos con la aplicación, se llaman evaluadores o expertos porque son personas que tienen habilidades y experiencia con respecto a la usabilidad y la interfaz de usuario.

La evaluación de usabilidad, es el segundo método de ensayo, el administrador de la prueba explica los escenarios de sesión y prueba, y le pide a los participantes que llenen brevemente un cuestionario de antecedentes. Luego los participantes leen los escenarios de prueba y tratan de completar las tareas. Finalmente después de cada tarea, el administrador de la prueba le pidió al participante que califique la interfaz en una escala Likert de 5 puntos.

El resultado de ambas pruebas muestra que la aplicación es muy fácil de utilizar. A pesar de que hay algunos problemas de usabilidad que no son graves y que se pueden eliminar de inmediato mediante la implementación de las recomendaciones de los evaluadores expertos y los usuarios.

Es aconsejable aumentar el número de expertos evaluadores para la evaluación heurística y aumentar el número de participantes para las pruebas de usabilidad para obtener una retroalimentación más a fondo del sistema.

EP 39 (Xu, Ding, Huang, & Chen, 2014): A Pilot Study of an Inspection Framework for Automated Usability Guideline Reviews of Mobile Health Applications

La evaluación Heurística o guía de revisión ha demostrado ser un método eficaz entre muchos métodos de evaluación de usabilidad. El proceso de evaluación heurística por profesionales de la usabilidad, puede ser costoso y consume tiempo, especialmente en las actualizaciones frecuentes de prototipos generados por iteraciones rápidas.

En este trabajo se pretende automatizar específicamente la evaluación heurística de las aplicaciones móviles de salud, el primer paso es identificar directrices o heurísticas en el dominio de la salud para los dispositivos móviles. Para ayudar a encontrar posibles problemas de usabilidad en una etapa temprana y reducir la carga de trabajo de expertos en usabilidad, se propone un marco de inspección para realizar evaluaciones automatizadas de las aplicaciones móviles de salud. El marco de inspección se basa en las directrices de usabilidad de la Sociedad de Sistemas de Información y Gestión en la salud “HIMSS” para aplicaciones de salud móvil.

En primer lugar, se tradujo las descripciones de alto nivel de las directrices de usabilidad dentro de métricas operativizadas que pueden ser medidas por el software.

En segundo lugar, se demostró que la traducción es significativa proporcionando un análisis detallado de las métricas sugeridas y estudios de casos reales. Se espera que este marco puede ser utilizado para hacer contribuir en la usabilidad de las aplicaciones móviles de salud y que sea adaptable cuando se desarrollen nuevos productos en el campo. A pesar de los resultados alentadores de este estudio, los autores mencionan que este marco de inspección automatizado no captura por completo todos los problemas de usabilidad, ni podrá sustituir a los expertos humanos en usabilidad.

EP 40 (López-Gil, Urretavizcaya, Fernández-Castro, & Losada, 2014): Integrating Field Studies in Agile Development to Evaluate Usability on Context Dependant Mobile Applications

El desarrollo ágil es una tendencia actual en Ingeniería de Software que se siente alentado por las necesidades especiales de los sistemas interactivos que requieren la incorporación de la ingeniería de la usabilidad en sus procesos de desarrollo. En este sentido, la forma de evaluación de usabilidad de aplicaciones interactivas móviles debe llevarse a cabo en el marco de las metodologías ágiles.

En este estudio se realiza el desarrollo de la aplicación FinMyPlace, que es una aplicación para ayudar a los usuarios a encontrar su ubicación física dentro de una universidad, la metodología ágil de desarrollo que se utilizará para la creación de este producto software se llama InterMod.

InterMod controla el desarrollo de software a través de los objetivos del usuario, es decir, los deseos de los usuarios que pueden ser suplidas por uno o más requisitos funcionales y no funcionales. Cada objetivo de usuario es una unidad lógica completa, así como una parte del resultado final. Además, cada objetivo de usuario se cumple por medio de modelos desarrollados por los equipos de trabajo a lo largo de las actividades realizadas en las diferentes iteraciones del sistema.

Posteriormente de la creación de la aplicación, los autores integran el análisis o evaluación de la aplicación “FindMyPlace” con el método de desarrollo ágil InterMod expuesto anteriormente. Para el proceso de análisis de la aplicación varias técnicas de evaluación de usabilidad son aplicadas con el fin de reunir las medidas de eficacia, eficiencia, y la satisfacción de los usuarios. Entre las técnicas de evaluación empleadas en el estudio están: una evaluación de campo, el protocolo pensar en voz alta, un cuestionario de satisfacción y entrevistas.

Este estudio permitió detectar distintos patrones de interacción y factores contextuales que son difíciles o imposibles de simular en laboratorio. La integración de las evaluaciones de campo en procesos ágiles de desarrollo, como el que presentan los autores de este estudio, es consistente con el objetivo de mejorar la calidad del producto a través de evaluaciones adicionales.

Los resultados de este estudio de campo realizado de una aplicación móvil utilizando metodologías ágiles de desarrollo señalan que las evaluaciones de

usabilidad se han integrado satisfactoriamente con las metodologías ágiles, de una manera coherente; así como también se encontró que el estudio de campo se ajusta a las necesidades crecientes de los desarrollos ágiles.

EP41 (Angulo Cevallos, 2014): Case study on mobile applications UX: effect of the usage of a crosss-platform development framework

El objetivo de este estudio es comparar un enfoque multiplataforma y nativo móvil para la determinación si el enfoque de desarrollo seleccionado tiene algún impacto en los usuarios en términos de experiencia de usuario.

Para cumplir este objetivo en este trabajo se desarrollaron dos versiones de una misma aplicación: una usando un framework que genera código iOS y Android automáticamente y otra que desarrolla versiones nativas de la misma aplicación. Estas aplicaciones son evaluadas con 37 usuarios con la combinación de los métodos: evaluación de usabilidad de laboratorio y estudio longitudinal.

El test de usabilidad fue llevado a cabo por los usuarios donde se les otorgó cinco tareas a completar en la aplicación. Durante la evaluación el tiempo que tarda en realizar cada tarea es medido, toman notas y graban con el fin de entender las acciones de los usuarios. La evaluación de laboratorio se realizó para evaluar la primera impresión de los usuarios cuando no han interactuado con la aplicación antes, después de cada evaluación de laboratorio hicieron un cuestionario de satisfacción.

Después de las tareas los usuarios fueron entrevistados por un cuestionario de satisfacción, para conocer sus impresiones de la aplicación. También se ejecutó un estudio longitudinal, en donde los participantes tienen cinco días para usar la aplicación libremente bajo cualquier circunstancia, después su nivel de satisfacción también fue evaluado. Los cuestionarios se ejecutaron para obtener las opiniones generales del sistema sobre características específicas.

Los resultados del estudio muestran que un buen nivel de experiencia de usuario puede ser obtenido si el framework de desarrollo multiplataforma es escogido cuidadosamente en términos de proveer un estilo de interacción adaptado para cada plataforma, y un equipo de desarrollo con expertos en experiencia de usuarios. Pero hay más posibilidades de conseguir una mejor experiencia de usuarios mediante el desarrollo de una aplicación nativa, estas son más fuertes para la plataforma iOS que

para Android donde los usuarios pueden estar acostumbrados a una mayor diversidad de estilos de interacción.

Con las evaluaciones de usabilidad realizadas se analiza la experiencia de usuarios desde diferentes contextos de uso, lo que puede brindar información valiosa para entender posibles problemas o dificultades que los usuarios quizás tengan mientras lleven a cabo tareas específicas. Los modelos existentes de evaluación de la usabilidad deben ser personalizados para abordar el enfoque móvil, sin embargo el uso de un modelo establecido puede ser útil como una línea base para comparar dos versiones de una misma aplicación.

EP 42 (Assal, Hurtado, Imran, & Chiasson, 2015): What's the deal with privacy apps? A comprehensive exploration of user perception and usability

En este artículo se exploró la privacidad móvil a través de una encuesta y una evaluación de usabilidad de tres aplicaciones móviles de preservación de privacidad. Este estudio explora el conocimiento de los riesgos de privacidad, así como las actitudes y motivaciones de los usuarios para proteger su privacidad en dispositivos móviles. Se encontró que los usuarios tienen modelos mentales incompletos de los riesgos de privacidad asociados con tales dispositivos.

Para explorar los diferentes aspectos de la privacidad móvil y obtener una comprensión más completa del tema, se realizó varias evaluaciones de las siguientes aplicaciones que son herramientas que preservan a privacidad: Chatsecure, ObscuraCam y Tor.

El estudio inicia con la investigación sobre el conocimiento de los usuarios, preferencias y hábitos con respecto a la privacidad móvil. Posteriormente para cada aplicación se empleó diferentes técnicas de evaluación de usabilidad:

ObscuraCam: se condujo un estudio de laboratorio con 15 participantes

ChatSecure: se condujo un estudio de laboratorio con 20 participantes

Tor: particularmente se evalúan las aplicaciones Orbot y ProxyMob con una combinación de técnicas de usabilidad: recorrido cognitivo con una evaluación heurística.

Los resultados de las evaluaciones muestran que en Chatsecure se notan varios problemas de usabilidad ya que la aplicación no es intuitiva, su interfaz es inconsistente, es decir, viola varios estándares de guías de diseño. En la aplicación

ObscuraCam se revelan problemas de interfaz porque las características de la aplicación no están claramente visibles a los usuarios. Por otro lado en las aplicaciones de Tor evaluadas: Orboot y ProxyMob se encuentra que Orboot no es intuitivo ni es entendible incluso para usuarios técnicos, la aplicación habla un lenguaje muy técnico; y ProxyMob no es visible al usuario una vez instalado y también tiene problemas de interfaz.

En general los resultados revelaron problemas de usabilidad que pueden afectar negativamente a los datos de la seguridad de los usuarios, hasta de los usuarios más técnicamente conscientes.

En resumen, se ha extraído la información más importante de cada estudio primario, características que contribuyen a responder las preguntas de investigación planteadas, por ejemplo el método o técnica empleada por los investigadores de las evaluaciones de usabilidad de aplicaciones móviles, también las ventajas e inconvenientes suscitados en los estudios, y demás información que se consideró relevante de cada estudio.

5.2 Resultados de la SLR

La presente revisión sistemática de literatura, permitió encontrar importantes hallazgos, en lo que respecta a la evaluación de usabilidad de aplicaciones móviles. El proceso de extracción de datos de los estudios primarios muestra las características más significativas tanto generales como específicas en cada artículo. Esta sección se divide en: Resultados generales, Resultados específicos, y Análisis de resultados.

5.2.1 Resultados Generales de la SLR Sobre Repositorios

Dentro de las características generales se ha extraído las características del estilo como año de publicación y fuente de publicación.

Los hallazgos que son evidenciados de los estudios primarios, corresponden por un lado a la cronología de las propuestas para la evaluación de usabilidad de aplicaciones móviles multiplataforma y por otro lado a las fuentes de publicación, donde los autores difunden los resultados de sus investigaciones. La Tabla 3, muestra las fuentes de publicación, su identificador y el año de la publicación.

Tabla 3**Fuentes de Publicación**

Estudio	Fuente de Publicación	Año
EP1	International Journal of Mobile Human Computer Interaction	2016
EP2	IEEE Information Technology and Multimedia (ICIMU)	2014
EP3	Behaviour & Information Technology	2015
EP4	Lecture Notes in Computer Science	2015
EP5	International Conference on Information Integration and Web-based Applications Services	2015
EP6	Journal of Medical Systems	2015
EP7	Journal of Theoretical and Applied Information Technology	2015
EP8	International Conference Communications in Computer and Information Science	2015
EP9	Jurnal Teknologi	2015
EP10	Lecture Notes in Computer Science	2015
EP11	Healthcare Informatics Research	2015
EP12	ANZ Journal of Surgery	2015
EP13	International Journal of Instruction	2015
EP14	Patient Prefer Adherence	2014
EP15	International Journal of Mobile Communications	2014
EP16	Information and Communication Technology for The Muslim World (ICT4M)	2014
EP17	International Journal of Mobile Human Computer Interaction	2014
EP18	Lecture Notes in Computer Science	2014
EP19	JMIR mhealth and uhealth (mobile and ubiquitous health)	2014
EP20	Journal of Diabetes Science and Technology	2014

EP21	International Conference on Web Information Systems and Technologies	2014
EP22	Journal of Medical Internet research	2014
EP23	Lecture Notes in Computer Science	2014
EP24	Journal of Medical Systems	2014
EP25	The Scientific World Journal	2014
EP26	Current Diabetes Reports	2015
EP27	Speech Communication	2014
EP28	JMidwifery Womens Health.	2014
EP29	Communications in Computer and Information Science	2014
EP30	Journal of Medical Internet research	2014
EP31	International conference on Human-computer interaction with mobile devices & services	2014
EP32	Lecture Notes in Computer Science	2014
EP33	JMIR mhealth and uhealth (mobile and ubiquitous health)	2014
EP34	Journal Pain Medicine	2014
EP35	e-Journal of Health Informatics	2014
EP36	ACM Conference on Human Factors in Computing Systems	2014
EP37	Acta Paulista de Enfermagem	2014
EP38	CASCE 2013 - International Conference on Advances Science and Contemporary Engineering, Jakarta, Indonesia	2014
EP39	ACM Conference Wireless Health	2014
EP40	International Conference on Human Computer Interaction	2014
EP41	E.T.S. de Ingenieros Informáticos (UPM)	2014
EP42	ACM Mobile and Ubiquitous Multimedia	2015

Como se puede apreciar las fuentes de publicación utilizadas son revistas, congresos, y simposios relacionados con las ciencias de la computación y la Interacción humano computador o humano móvil.

En lo que respecta a la cronología se puede observar en la Figura 9 que en los estudios primarios los datos sobre evaluación de usabilidad de aplicaciones móviles va en disminución, ya que existen muchos estudios entre el 2014 y el 2015, pero nada más 1 realizado en el 2016. Sin embargo este tema sigue siendo investigativo ya que dentro de estos estudios primarios pocos hablan sobre evaluación de usabilidad automatizada, para ser exactos de los 42 estudios primarios nada más 2 hablan sobre metodologías automatizadas, la mayor parte usan metodologías tradicionales por medio de expertos o pruebas con usuarios.

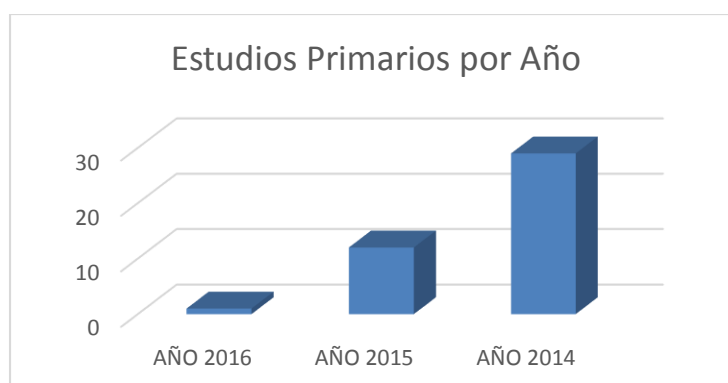


Figura 9. Número de estudios primarios por año

5.2.2 Resultados Específicos de la SLR Sobre Repositorios.

El estudio detallado de los estudios primarios evidencio ciertos aspectos comunes que los describen. Estas características se han considerado como criterios de clasificación de los estudios primarios y son: los tipos de metodologías para evaluar usabilidad de cada estudio, modelos de usabilidad más populares para evaluar usabilidad, factores de usabilidad considerados en los estudios, tipos de aplicaciones evaluadas, entre otras particularidades que contribuyen a dar respuesta a las preguntas de investigación que se exponen a continuación.

RQ1: ¿Cuáles son las técnicas existentes para evaluar la usabilidad de aplicaciones en dispositivos móviles?

En los estudios primarios se encontraron diferentes métodos y técnicas de evaluación de usabilidad de aplicaciones en dispositivos móviles usadas por los autores, estos se detallan a continuación:

Evaluación heurística: es un método de inspección rápido, barato y fácil para encontrar problemas de usabilidad (Gresse von Wangenheim, y otros, 2016), para su

aplicación se requiere que un grupo de expertos en usabilidad evalúen la conformidad del diseño de interfaz de la aplicación móvil con un conjunto de heurísticas de usabilidad con el fin de encontrar los posibles problemas de la aplicación (de Lima Salgado & Pimenta Freire, 2014) (Arnhold, Quade, & Kirch, 2014). Este tipo de evaluación está centrado en aplicaciones de escritorio pero es adecuado para las características específicas de las aplicaciones móviles (Yáñez Gómez, Cascado Caballero, & Sevillano, 2014).

La evaluación heurística también puede ser llevada a cabo sin la necesidad de involucrar usuarios finales (Gresse von Wangenheim, y otros, 2016), y es relativamente económica y fácil de aprender (Joyce, Lilley, Barker, & Jefferies, 2015). Es un método adecuado para toda fase del ciclo de vida del software, no requiere planeación previa y además no es mandatorio para los evaluadores que sean expertos en usabilidad (Yáñez Gómez, Cascado Caballero, & Sevillano, 2014).

Las heurísticas propuestas por Nielsen y Molich son los métodos más usados en las evaluaciones heurísticas de usabilidad en aplicaciones móviles, pero las propuestas de nuevas heurísticas específicas para interfaces móviles han crecido sustancialmente (de Lima Salgado & Pimenta Freire, 2014).

Evaluación de Laboratorio: Es un método donde los participantes son observados por un investigador a lo largo de su rendimiento en cada tarea asignada (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014). El investigador puede registrar la comunicación no verbal relevante de los usuarios (por ejemplo, seguridad o expresiones faciales confusas), todo el proceso de evaluación, es grabado en vídeo (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014).

Evaluación de Campo: Es un método de evaluación de usabilidad de aplicaciones móviles que es llevado a cabo con usuarios reales, en un contexto de uso real. En el experimento los expertos usan métodos observacionales y toman notas, mientras los usuarios realizan las tareas asignadas (López-Gil, Urretavizcaya, Fernández-Castro, & Losada, 2014).

Es importante mencionar que en uno de los estudios primarios integran la evaluación de campo dentro un proceso de desarrollo de aplicaciones, esto es utilizando la metodología ágil de desarrollo de software “InterMod”, que controla y

combina los objetivos del usuario, conocidos como requerimientos funcionales y no funcionales con la evaluación de campo, dentro de la etapa de desarrollo (López-Gil, Urretavizcaya, Fernández-Castro, & Losada, 2014).

Cuestionarios: Es un método para evaluar la usabilidad subjetiva, es decir, la que es percibida por el usuario y también su grado de satisfacción con la aplicación móvil (Saleh, Binti Isamil, & Binti Fabil, 2015). Los cuestionarios presentan un conjunto de preguntas desarrolladas para que los usuarios las respondan posteriormente de haber usado la aplicación, para su interpretación se evalúa el grado en que la aplicación asiste a los elementos de medición y se le asigna una calificación para cada pregunta (Cruz Zapata, Fernández, Idri, & Toval, 2015). Es un método económico y fácil de administrar a los participantes (Angulo Cevallos, 2014). Los tipos de cuestionarios de usabilidad más populares son:

- **SUS:** La Escala de usabilidad del sistema (SUS) propuesta por Brooke en 1996, brinda una herramienta para medir usabilidad. Este cuestionario está compuesto por una escala de 10 elementos, que dan una visión global de la medición subjetiva de usabilidad. Cada pregunta incluye 5 posibles respuestas van de “muy en desacuerdo” hasta “muy de acuerdo” (Wein, 2014).
- **SUMI:** El Inventario de la medición de usabilidad del software SUMI, contiene 50 elementos que tienen que ser respondidos en una escala de tres puntos (de acuerdo, indeciso y en desacuerdo). SUMI considera 5 sub-escalas globales para medir la usabilidad: eficacia, afecto, utilidad, control y facilidad de aprendizaje (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014).
- **PSSUQ:** Consiste en una escala de satisfacción general consta de 19 elementos divididos en tres sub-escalas: utilidad del sistema, calidad de la información y de calidad de interfaz. Se califican en una escala de 7 puntos van desde muy en desacuerdo = 1 hasta muy de acuerdo = 7 (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014).
- **“UEQ” Cuestionario de Experiencia de usuario:** Este cuestionario brinda una rápida medición de la experiencia de los usuarios de productos interactivos, permitiendo a los usuarios expresar sus sentimientos e

impresiones sobre el producto al interactuar con el mismo. UEQ considera aspectos hedónicos: aspectos de calidad no orientados a tareas, como por ejemplo impresión: estética, simulación, etc.; y aspectos pragmáticos: tareas orientadas a aspectos de calidad: eficiencia, facilidad de aprendizaje, etc. (Angulo Cevallos, 2014).

- **Cuestionario Ad-hoc:** Cuestionario adicional para evaluar otros aspectos no cubiertos por SUS, UEQ, etc. (Angulo Cevallos, 2014).

Algunos autores en los estudios primarios realizan cuestionarios basados en aspectos o criterios de usabilidad que consideran importantes y quizás los cuestionarios anteriormente mencionados no cubren. Estos cuestionarios utilizan una escala llamada Likert de cinco puntos para responder a cada pregunta dónde: 1 = Totalmente en desacuerdo, 2 = En desacuerdo, 3 = Ligeramente en desacuerdo, 4 = De acuerdo y 5 = Totalmente de acuerdo (Nawi, Isa Hamzah, Chy Ren, & Tamuri, 2015).

Entrevistas: es un método de evaluación de usabilidad similar a los cuestionarios ya que ambos consisten en un conjunto de preguntas que se basan en datos subjetivos de los usuarios (Hussain, Hashim, & Nordin, 2014), es decir, sirven para conocer sus opiniones y sentimientos después de usar las aplicaciones móviles (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015).

Protocolo “Think aloud” (Pensar en voz alta): El protocolo “pensar en voz alta” es un método directo para evaluar la usabilidad de una aplicación. Todos los datos verbales obtenidos y las interacciones de los usuarios tienen que ser transcritas para su análisis. “Think aloud” genera un informe de las respuestas no estructuradas de los usuarios, y por esta razón es que el protocolo ha sido objeto de fuertes críticas en cuanto a su fiabilidad y validez (Cruz Zapata, Fernández, Idri, & Toval, 2015). Jaspers en el 2009, afirma que un cuestionario complementario al usuario después de la sesión de evaluación “think aloud” y al menos contar con dos evaluadores podría ayudar a la interpretación de la información subjetiva proporcionada por los usuarios (Cruz Zapata, Fernández, Idri, & Toval, 2015).

El protocolo “Think aloud” tiene sus variantes (Lira, de Souza, Ferreira, & Carvalho, 2014):

- **Pensar en voz alta concurrente (Concurrent Think aloud):** se instruye al usuario a hablar mientras esta interactuando con el software (Lira, de Souza, Ferreira, & Carvalho, 2014).
- **Pensar en voz alta Retrospectiva (Retrospective Think-aloud):** se instruye al usuario a hablar mientras mira un video de su interacción con el software (Lira, de Souza, Ferreira, & Carvalho, 2014).

Recorrido Cognitivo: Un recorrido cognitivo implica la evaluación de uso por primera vez de una aplicación, es decir sin entrenamiento formal del sistema, en el que un experto: (1) establece un objetivo a lograr, (2) busca acciones disponibles, (3) selecciona una acción o tarea que parece más relevante, y (4) realiza la acción y evalúa el sistema entero o aplicación para hacer un avance hacia la meta original (Lyles, Sarkar, & Osborn, 2014).

Estudio longitudinal: En este método los participantes tienen un tiempo determinado para usar la aplicación libremente bajo cualquier circunstancia, después de culminar este periodo de tiempo de uso, los investigadores evalúan la satisfacción de los usuarios (Angulo Cevallos, 2014). Los estudios longitudinales son importantes en la evaluación de aplicaciones móviles, ya que considera los diferentes contextos donde la aplicación sería usada típicamente, característica difícil de reproducir en un ambiente controlado (Angulo Cevallos, 2014).

Combinación de métodos de evaluación de usabilidad: En numerosos estudios los autores decidieron combinar más de un método para evaluar la usabilidad de las aplicaciones móviles.

Por ejemplo los autores, realizan una evaluación de laboratorio, conducida con un equipo especial, donde a los participantes se les dio un dispositivo móvil, para empezar a ejecutar las tareas, mientras realizaban las mismas, tenían que “pensar en voz alta”, en este proceso fueron grabados. Después de la evaluación los usuarios fueron interrogados, por medio de un cuestionario online (Faliagka, Rigou, & Sirmakessis, 2014). Al igual, (Aizawa, y otros, 2014) decidieron realizar una evaluación de usabilidad de laboratorio de la aplicación FOODLOG, con el establecimiento de un conjunto de tareas y al finalizar completaron un cuestionario de satisfacción.

Generalmente los métodos de evaluación observacionales de campo y laboratorio son complementados con cuestionarios, entrevistas o el protocolo “think aloud”

(Biederer, Arguel, Liu, & Lau, 2014), (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015), (Beul-Leusmann, y otros, 2014), (Saleh, Binti Isamil, & Binti Fabil, 2015), (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015), (Tsiaousis & Giaglis, 2014), (Wein, 2014), (López-Gil, Urretavizcaya, Fernández-Castro, & Losada, 2014) & (Abdul, Hussain, & Zainol, 2014).

Por otro lado, ciertos autores han decidido combinar otros métodos de evaluación de usabilidad, por ejemplo en un estudio primario los autores investigan cuatro métodos empíricos de evaluación de usabilidad en dispositivos móviles: cuestionario, entrevista, registros y “think aloud” (Cruz Zapata, Hernandez Niñirola, Idri, Fernandez-Aleman, & Toval, 2014). Otros autores en el EP14, evaluaron la usabilidad en un método mixto que constaba en las siguientes 4 fases: (1) evaluación heurística (2) evaluación en laboratorio, (3) evaluación de campo, (4) una última evaluación de laboratorio con usuarios finales seguido de un cuestionario de satisfacción (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014).

La importancia de un enfoque de métodos mixtos permite que diferentes cuestiones de usabilidad se den a conocer (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014) y la evaluación es más robusta e integral (Lyles, Sarkar, & Osborn, 2014).

Otros métodos:

En dos de los estudios primarios se mencionan métodos para evaluar la usabilidad de aplicaciones más específicas, como aplicaciones de aprendizaje, y aplicaciones multimodales. El primer estudio es realizado por (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015), donde mencionan una evaluación de usabilidad de aplicaciones de aprendizaje usando el “Sistema de Evaluación de Usabilidad” (SUE), que es una metodología que permite una evaluación fiable mediante la combinación de la evaluación sistemática de inspección y pruebas de usuario, donde señalaron cuatro dimensiones que las aplicaciones de aprendizaje deben tener para ser consideradas usables. Esto incluye la presentación, hipermedialidad, proactividad de la aplicación y la actividad del usuario (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).

En el segundo estudio primario anteriormente mencionado, en cambio se propone un framework para evaluar la usabilidad de aplicaciones multimodales, este marco añade métricas específicas de usuario, que se ha demostrado que son útiles en la

conservación de la predicción de usuarios y la identificación de los patrones de comportamiento del usuario (Maren Witt, 2014).

Evaluaciones Automatizadas:

En dos estudios primarios mencionan métodos de evaluación de aplicaciones móviles automatizados, el primero se llama AUT (Kluth, Krempels, & Samsel, 2014), y el segundo es un método de inspección automatizado para aplicaciones móviles de salud (Xu, Ding, Huang, & Chen, 2014).

AUT (Automated usability testing): es una metodología propuesta por Ivory & Hearst en el 2001, para evaluar la usabilidad de las aplicaciones móviles de la plataforma iOS, que está separada en cinco tipos de métodos: evaluación, inspección, cuestionario, modelos analíticos y simulaciones. AUT posee cuatro diferentes pasos de automatización: no automática, captura automática, análisis automático y crítica automática. Cada paso de automatización es consecutivo, además el esfuerzo de AUT es estimado formalmente por (las tareas explícitas para participantes) e informalmente (los participantes usan objetivos del sistema sin ninguna tarea adicional (Kluth, Krempels, & Samsel, 2014). Las fases de AUT son las siguientes:

(1) **Fase de preparación** del proyecto donde el desarrollador integra el marco de captura AUT dentro de su aplicación.

(2) **Fase de captura** automática para generar registros con los datos de las interacciones de los usuarios.

(3) **Fase de Análisis** y reconocimiento de los patrones, se conforma con las especificaciones del problema de los patrones de HCI, los mismos son aplicados a cada registro encontrado y marcado como problema.

(4) **Fase Crítica** que se refiere al reconocimiento de patrones de diseño de HCI y relaciona una solución en forma de mejores prácticas para cada problema dado.

Inspección automatizada: Es un framework, que actualmente está soportado sobre aplicaciones de Salud de la plataforma Android, constituye un proceso de evaluación automatizado, sin intervención de usuarios. Los autores del framework han traducido en lenguaje de descripción de alto nivel, las directrices de usabilidad de la Sociedad de Sistemas de Información y Gestión en la salud (HIMSS) en métricas cuantitativas que podrían medirse automáticamente por el software (Xu, Ding, Huang, & Chen, 2014).

RQ2: ¿Qué modelos de usabilidad son comúnmente utilizados en la evaluación de usabilidad de aplicaciones móviles?

Existen diferentes estándares o modelos propuestos que sirven para cuantificar y evaluar la usabilidad dentro de la Interacción humano-máquina y la Ingeniería de Software, pero la usabilidad no se ha definido de forma consistente a través de estos modelos y estándares (Enriquez & Casas, 2013). Los modelos o estándares existentes proponen una guía de atributos y métricas a utilizar en la medición de usabilidad, sin embargo estos no cubren todos los aspectos de la usabilidad (Enriquez & Casas, 2013). Entre los modelos de usabilidad que son comúnmente usados para la evaluación de usabilidad de aplicaciones móviles están los siguientes:

Modelo de Usabilidad de Nielsen: este modelo propone un conjunto de heurísticas tradicionales como guías de la interfaz de usuarios para aplicaciones de computadoras de escritorio, también aplicadas a diseño de interfaces de aplicaciones móviles (Gresse von Wangenheim, y otros, 2016).

Modelo ISO 9241-11: este modelo define la usabilidad como “el grado en el cual un producto puede ser usado por usuarios específicos para cumplir objetivos específicos con efectividad, eficiencia y satisfacción en un determinado contexto de uso” mide atributos como eficiencia, efectividad y satisfacción (Gresse von Wangenheim, y otros, 2016).

Modelo de Shneiderman 2005: es un modelo que propone la medición de los siguientes atributos: rapidez de rendimiento, tiempo de aprendizaje, tiempo de retención, tasa de error por usuarios y satisfacción subjetiva (Ibrahim, Hussain, & Binti, 2014).

Modelo de medición integrada de calidad en uso QUIM (Quality in use integrated measurement): es un modelo consolidado para evaluar usabilidad (Enriquez & Casas, 2013). QUIM es construido en una organización jerárquica, se basa en la descomposición de factores de usabilidad (Ibrahim, Hussain, & Binti, 2014). El modelo incluye 10 factores de usabilidad, desafortunadamente QUIM todavía no es óptimo y necesita ser validado (Saleh, Binti Isamil, & Binti Fabil, 2015).

Modelo de Condos et al. 2002: es un modelo de usabilidad que propone las siguientes métricas: navegación, contenidos, información de la arquitectura,

prevención de errores, presentación, tasa de entrada y visualización de menús (Ibrahim, Hussain, & Binti, 2014).

Modelo de métrica pregunta objetivo mGQM de Azham, 2011: este modelo de usabilidad consta de métricas de usabilidad para evaluar las medidas cuantitativas y cualitativas de las aplicaciones móviles (Hussain, Hashim, & Nordin, 2014). El modelo GQM es una formación jerárquica para evaluar aplicaciones móviles, este modelo está basado en un método de manejo de objetivos para desarrollar las métricas. GQM inicia definiendo un objetivo, refina este objetivo en preguntas y posteriormente define métricas que brindan información para responder a estas preguntas (Saleh, Binti Isamil, & Binti Fabil, 2015). Este modelo considera importantes los siguientes factores: simplicidad, exactitud, tiempo tomado, características, seguridad y que tan atractiva es la aplicación (Ibrahim, Hussain, & Binti, 2014).

Modelo de Baharuddin et al. 2013: este modelo de usabilidad propone las siguientes métricas: efectividad, eficiencia, satisfacción, utilidad, estética, facilidad de aprendizaje, simplicidad, que tan intuitiva, comprensible y atractiva es la aplicación (Ibrahim, Hussain, & Binti, 2014).

Modelo de Usabilidad PACMAD: es un modelo para evaluar la usabilidad de aplicaciones móviles propuesto por Harrison, basado en los siguientes factores de usabilidad: usuarios, contexto de uso y tareas. El objetivo de este modelo es extender los modelos existentes de usabilidad, como Nielsen o ISO para el contexto específico de aplicaciones móviles (Saleh, Binti Isamil, & Binti Fabil, 2015). Los atributos que medidos en PACMAD son: efectividad, eficiencia, satisfacción, facilidad de aprendizaje, facilidad de recordar, errores y carga cognitiva (Saleh, Binti Isamil, & Binti Fabil, 2015).

Métricas para estándares de usabilidad en computación (Music): modelo que posee un conjunto de métricas para evaluar estándares de usabilidad en Computación, Music identifica métricas relacionadas con la medición del desempeño de la usabilidad del software, incluyen medidas de rendimiento, tales como efectividad en las tareas de los usuarios, eficiencia temporal y proporción del período productivo. Sin embargo, el modelo Music no puede indicar con precisión los atributos de usabilidad tales como la facilidad de aprendizaje y la satisfacción (Saleh, Binti Isamil, & Binti Fabil, 2015).

Modelo de Aceptación de Tecnología TAM2: consiste en un conjunto de cuestionarios que fueron desarrollados mediante la adaptación de dos dimensiones: "utilidad percibida" y "facilidad de uso percibida" (Biederer, Arguel, Liu, & Lau, 2014).

Modelo propuesto por Ardito: este modelo para medir la usabilidad de las aplicaciones móviles de aprendizaje se basa en cuatro dimensiones que las aplicaciones de este tipo deben satisfacer. Esto incluye presentación, hipermedialidad, proactividad de la aplicación y actividad del usuario (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).

Modelo propuesto por Bertini: este modelo es una modificación de las heurísticas de Nielsen con el objetivo de capturar factores contextuales en la computación móvil. Las heurísticas de usabilidad son (1) visibilidad del estado del sistema, (2) concordancia entre el sistema y el mundo real, (3) consistencia y mapeo, buen diseño ergonómico y minimalista, (4) facilidad de entrada, (5) legibilidad de la pantalla y flexibilidad, (6) eficiencia de uso y personalización, (7) estética, privacidad y convenciones sociales, y (8) administración realista de errores.

Los **modelos de usabilidad** de las aplicaciones móviles, son un conjunto de guías propuestas por diferentes autores, que contienen lineamientos y atributos a ser medidos dentro de un proceso de evaluación de usabilidad, cumpliendo estos atributos un producto software se considerará usable. Dentro de los estudios primarios que realizan evaluaciones de usabilidad basadas en modelos, se puede observar que el modelo más conocido y usado es el de Nielsen, ver en la Figura 10.

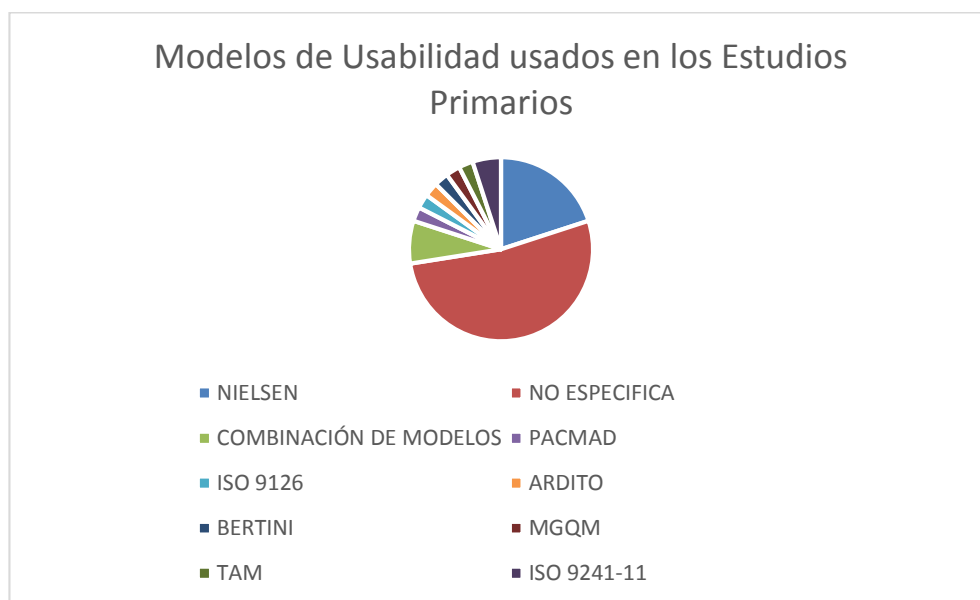


Figura 10. Modelos de Usabilidad usados en los Estudios Primarios

Como se puede observar en la Figura 10, la mayoría de estudios no especifica el modelo usado al momento de evaluar la usabilidad, esto se debe a que los autores decidieron evaluar por medio de tareas específicas, y miden los atributos que ellos consideran importantes sin basarse en un modelo en especial. El modelo que es muy usado y más popular es el de Nielsen, muchos autores decidieron realizar evaluaciones heurísticas basadas en estos criterios propuestos que principalmente fueron para evaluar aplicaciones tradicionales de escritorio, y después adaptados al contexto móvil. El modelo ISO 9241-11 es un modelo que mide la efectividad, eficiencia y satisfacción. Como se puede apreciar cada uno de los modelos ISO, posee sus variantes en cuanto a los atributos a ser medidos al momento de evaluar la usabilidad.

En los estudios primarios que realizan evaluaciones de usabilidad basadas en modelos, se puede observar que el modelo más conocido y usado es el de Nielsen (Arnhold, Quade, & Kirch, 2014), (Joyce, Lilley, Barker, & Jefferies, 2015), (de Lima Salgado & Pimenta Freire, 2014), (Luhur & Widjaja, 2014). Sin embargo varios autores optaron por la combinación de los modelos existentes: el modelo de Nielsen con el modelo ISO 9241-11 (Saleh, Binti Isamil, & Binti Fabil, 2015), o por la modificación de las heurísticas de Nielsen para capturar factores del contexto móvil como el estudio de (Gresse von Wangenheim, y otros, 2016) y (Mirkovic, Kaufman, & Ruland, 2014). La combinación y personalización de los modelos existentes de usabilidad indica la falta de modelos formalizados en el proceso de evaluación de

usabilidad de aplicaciones móviles, ya que si bien es cierto existen diferentes estándares o modelos propuestos dentro de la Interacción humano-máquina y la Ingeniería de Software, pero en la usabilidad móvil no se han definido de forma consistente.

RQ3: ¿Qué atributos de usabilidad son considerados en la evaluación de usabilidad de aplicaciones móviles y que métricas se utilizan para medir dichos atributos?

Los datos extraídos de los estudios primarios para esta pregunta de investigación pertenecen a los atributos de los modelos de usabilidad más popularmente usados indicados en la RQ2. Estos atributos y métricas se resumen en la Tabla 4.

Tabla 4

Modelos de Usabilidad, Atributos y métricas

MODELOS DE USABILIDAD	ATRIBUTOS Y MÉTRICAS
<p>Heurísticas Tradicionales basadas en Nielsen (2004) (Gresse von Wangenheim, y otros, 2016), (Joyce, Lilley, Barker, & Jefferies, 2015).</p>	<p>Visibilidad del estado del sistema: para cada interfaz de usuario es importante mostrar en cada pantalla la información clara y concisa sobre el estado del sistema (Gresse von Wangenheim, y otros, 2016). El sistema siempre debe mantener informados a los usuarios sobre lo que están haciendo, en un tiempo razonable (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Encuentro entre el sistema y el mundo real: el sistema debe hablar el idioma del usuario, con palabras, frases y conceptos familiares para el usuario (Gresse von Wangenheim, y otros, 2016) (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Control y libertad para el usuario: las interrupciones deben ser tratadas naturalmente por el sistema, ofreciendo al usuario la posibilidad de reiniciar inmediatamente sus acciones (Gresse von Wangenheim, y otros, 2016). Los usuarios a menudo escogen funciones del sistema por error y necesitaran marcar “una salida de emergencia” para abandonar un estado no deseado (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Consistencia y estándares: la interfaz móvil debe ser similar a la interfaz de escritorio en términos de botones, logos y esquemas de color, así que la experiencia es consistente a través de todas las plataformas (Gresse von Wangenheim, y otros, 2016). Los usuarios no deberían tener que preguntarse si diferentes palabras, situaciones o acciones, (Continua)</p>

MODELOS DE USABILIDAD	ATRIBUTOS Y MÉTRICAS
	<p>significan lo mismo, generalmente las normas de consistencia siguen las convenciones de la plataforma (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Prevención de errores: la necesidad para confirmar acciones arriesgadas en procedimientos complejos y poder evitar errores accidentales (Gresse von Wangenheim, y otros, 2016), (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Reconocimiento en lugar de recuerdo: minimizar la carga de los usuarios, haciendo que los objetos, acciones y opciones sean visibles. El usuario no tiene que recordar información, el sistema debe ser visible o fácilmente recuperable cuando sea apropiado (Gresse von Wangenheim, y otros, 2016), (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Flexibilidad y eficiencia de uso: la carga rápida del sistema y la facilidad de aprendizaje son especialmente importantes durante la interacción con teléfonos móviles (Gresse von Wangenheim, y otros, 2016), (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Diseño estético y minimalista: la información presentada debe ser consistente, dividido en pequeñas porciones y fácil de encontrar con el fin de minimizar la carga cognitiva de los usuarios (Gresse von Wangenheim, y otros, 2016), (Joyce, Lilley, Barker, & Jefferies, 2015). La aplicación debe ser simple, enfocada a una tarea, también tiene que ser visualmente agradable, fácil de aprender e intuitiva (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Ayudar a los usuarios a reconocer, diagnosticar y recuperarse de errores: en la interacción con teléfonos móviles es necesario proveer una retroalimentación significativa, mensajes de errores concisos, opciones de retorno y fácil recuperación de errores (Gresse von Wangenheim, y otros, 2016), (Joyce, Lilley, Barker, & Jefferies, 2015).</p> <p>Ayuda y documentación: La información debe ser fácil de encontrar enfocada y no extensos (Gresse von Wangenheim, y otros, 2016), (Joyce, Lilley, Barker, & Jefferies, 2015).</p>
<p>HEURISTICAS PROPUESTAS por (Gresse von Wangenheim, y otros, 2016).</p>	<p>En el EP1, proponen un modelo donde se utilizan las mismas heurísticas tradicionales de Nielsen (2004), pero adaptadas al contexto móvil. Los autores deciden eliminar dos de las heurísticas que consideraron no importantes en el contexto móvil son: prevención de errores, (Continua)</p>

MODELOS DE USABILIDAD	ATRIBUTOS Y MÉTRICAS
	<p>ayuda y documentación, y agregan adicionalmente las siguientes heurísticas:</p> <p>Compatibilidad entre diferentes plataformas: Debido a la constante innovación de los dispositivos móviles en el mercado, las aplicaciones deben ser flexibles para adaptarse a las distintas plataformas y dispositivos (Gresse von Wangenheim, y otros, 2016).</p> <p>Poca Interacción Humana / dispositivos: debido al tamaño de los dispositivos móviles y su contexto de uso, es importante que el esfuerzo de interacción sea reducido. Por ejemplo para reducir la entrada de datos puede preverse características de autocompletado, cajas de textos con sugerencias, etc. (Gresse von Wangenheim, y otros, 2016).</p> <p>Interacción física y ergonómica: los controles de navegación más comunes deben estar al alcance de los usuarios y ser fácilmente presionados (Gresse von Wangenheim, y otros, 2016).</p> <p>Legibilidad y diseño: El diseño de la interfaz para el dispositivo móvil debe asegurarse de que los cuadros de texto caben en la pantalla y que el diseño se ajusta múltiples dispositivos (Gresse von Wangenheim, y otros, 2016).</p>
<p>Modelo de Norman & Nielsen 2010 (Faliagka, Rigou, & Sirmakessis, 2014).</p>	<p>Visibilidad: debe ser obvio para el usuario donde tocar, deslizar o realizar otro tipo de interacción con el dispositivo (Faliagka, Rigou, & Sirmakessis, 2014).</p> <p>Retroalimentación: los usuarios deben estar en control de la interacción (Faliagka, Rigou, & Sirmakessis, 2014).</p> <p>Consistencia: los elementos visuales deben ser estandarizados (Faliagka, Rigou, & Sirmakessis, 2014).</p> <p>Operaciones no destructivas: siempre se debe permitir a los usuarios deshacer o prevenir antes de cualquier acción no reversible que pueda ser ejecutada (Faliagka, Rigou, & Sirmakessis, 2014).</p> <p>Descubrimiento: un principio conectado con la visibilidad y se refiere a la necesidad de dar al usuario fácil acceso a todas las funcionalidades disponibles (Faliagka, Rigou, & Sirmakessis, 2014).</p> <p>Escalabilidad: las aplicaciones que se ejecutan en plataformas deben adaptarse fácilmente a diferentes tamaños de pantalla sin causar problemas de eficiencia del usuario (los controles deben tener un tamaño (Continua)</p>

MODELOS DE USABILIDAD	ATRIBUTOS Y MÉTRICAS
	<p>adecuado para golpear ligeramente el dedo y en una distancia adecuada de los controles cercanos) (Faliagka, Rigou, & Sirmakessis, 2014).</p> <p>Confiabilidad: debido al hecho que los gestos son invisibles, es decir el usuario no ve donde toca, muchas veces se activan accidentalmente otros controles cercanos y el usuario pierde el sentido de control (Faliagka, Rigou, & Sirmakessis, 2014).</p> <p>Eficiencia: En este estudio miden el promedio de tiempo por cada tarea de la aplicación, o también miden el porcentaje de usuarios que completaron esas tareas (Faliagka, Rigou, & Sirmakessis, 2014).</p>
<p>PACMAD (Saleh, Binti Isamil, & Binti Fabil, 2015)</p>	<p>Efectividad: habilidad del usuario para realizar una tarea en un contexto específico, puede ser medido con un conjunto de tareas propuestas para el usuario (Saleh, Binti Isamil, & Binti Fabil, 2015).</p> <p>Eficiencia: habilidad del usuario para cumplir sus tareas de manera exacta y rápida, se mide por el tiempo invertido en realizar la tarea y el número de toques requeridos para la completitud de la tarea (Saleh, Binti Isamil, & Binti Fabil, 2015).</p> <p>Satisfacción: es el nivel de confort percibido por el usuario al utilizar el sistema, es medido por cuestionarios y varias técnicas de colección de datos (Saleh, Binti Isamil, & Binti Fabil, 2015).</p> <p>Aprendizaje: es el nivel de facilidad con el que el usuario puede ser capaz de usar la aplicación, el tiempo el cual le toma al usuario acostumbrarse a la aplicación eficientemente. Se mide cuanto tiempo un usuario tarda en alcanzar un nivel de competencia, durante la realización de una tarea que el usuario nunca había realizado antes (Saleh, Binti Isamil, & Binti Fabil, 2015).</p> <p>Facilidad de recordar o memorizar: Se trata de la capacidad de los usuarios para mantener la forma de utilizar la aplicación eficazmente, sin necesidad de volver a aprender a usarlo después de un largo período de inactividad. Se puede medir en un estudio longitudinal o con dos observaciones repetidas (Saleh, Binti Isamil, & Binti Fabil, 2015).</p> <p>Errores: Este componente es acerca de cuan libre de errores es la utilización del software tanto para el usuario y el desarrollador; y también sobre lo fácil y rápido que el software se recupera de errores eventuales. Su evaluación debe incluir el registro de la naturaleza del error, y su (Continua)</p>

MODELOS DE USABILIDAD	ATRIBUTOS Y MÉTRICAS
	<p>frecuencia de ocurrencia. Es decir, que se puede medir a través de una evaluación de la tasa de error. (Saleh, Binti Isamil, & Binti Fabil, 2015).</p> <p>Carga cognitiva: Se refiere al procesamiento cognitivo que el usuario necesita para utilizar la aplicación. Se puede medir a través de una herramienta de evaluación TLX (Task load index) de índice de carga de tareas de la NASA para medir la carga cognitiva de trabajo del usuario por la interacción de un sistema (Saleh, Binti Isamil, & Binti Fabil, 2015). Por ejemplo, un usuario del dispositivo móvil puede desear enviar un mensaje instantáneo o al caminar, (Saleh, Binti Isamil, & Binti Fabil, 2015).</p>
<p>Atributos y Métricas de usabilidad definidas en el estudio de (Hsiu Ching, 2015) basados en Nielsen.</p>	<p>Facilidad de aprendizaje: es la rapidez y facilidad con la que los usuarios pueden alcanzar un nivel de excelencia en el uso del sistema (Hsiu Ching, 2015).</p> <p>Eficiencia: es el número de tareas por unidad de tiempo que el usuario puede realizar utilizando el sistema (Hsiu Ching, 2015).</p> <p>La satisfacción: es la opinión subjetiva de los usuarios sobre el sistema o sobre algunas partes de él. Es uno de los atributos de usabilidad más difícil de alcanzar, ya que es completamente dependiente de la opinión subjetiva de los usuarios (Hsiu Ching, 2015).</p> <p>Tasa de Error: Esto se refiere a los errores cometidos durante el uso del sistema y lo fácil que es recuperarse de ellos (Hsiu Ching, 2015).</p> <p>Facilidad de recordar o memorizar: Esto se refiere a la facilidad de recordar la forma en la que debe funcionar un sistema. Nielsen describe como la característica de un sistema que permite al usuario para volver al sistema después de algún período de no haberlo utilizado, sin tener que aprender todo de nuevo (Hsiu Ching, 2015).</p>
<p>Modelo ISO 9241-11 (Tsiaousis & Giaglis, 2014).</p>	<p>Eficiencia: es directamente medida a través de la grabación de los tiempos de finalización de tareas (Tsiaousis & Giaglis, 2014).</p> <p>Efectividad: es medida a través del registro de errores y porcentaje de completitud de la tarea dentro del tiempo asignado.</p> <p>Satisfacción: es medida a través de cuestionarios como por ejemplo SUMI, que es proporcionado a cada participante después del experimento (Tsiaousis & Giaglis, 2014). (Continua)</p>

MODELOS DE USABILIDAD	ATRIBUTOS Y MÉTRICAS
Modelo GQM (Hussain, Hashim, & Nordin, 2014)	Medidas objetivas / medidas subjetivas (sentimientos de la persona): <ul style="list-style-type: none"> • Tiempo tomado en ingreso de datos / satisfacción con teclado físico o virtual. • Número de errores mientras ingresan datos / satisfacción con la salida de datos. • Tiempo tomado para instalar / satisfacción con el proceso de instalación. • Número de interacciones mientras instala la aplicación / satisfacción con la optimización del tamaño de la pantalla. • Tiempo tomado para aprender / satisfacción con la ayuda de la aplicación. • Número de errores mientras aprende / satisfacción con los contenidos. • Número de errores / disfrute. • Tiempo tomado para completar tarea / satisfacción con la interfaz. • Número de tareas exitosas en el primer intento / seguridad mientras se maneja la aplicación. • Número de tareas exitosas en un tiempo dado / facilidad de encontrar ayuda. • Tiempo tomado en iniciar la app / estrés. • Tiempo tomado en responder / satisfacción con indicador de señal. • Tiempo tomado en conectar con la red / satisfacción con joystick virtual. • Número de veces que la asistencia de voz es proporcionada en una tarea / satisfacción mientras aprende a usar la aplicación. • Número de recursos del sistema mostrados / satisfacción con el texto. • Número de solicitudes de actualización de la aplicación / satisfacción con el sistema de navegación. • Porcentaje de batería usada durante instalación / satisfacción con la pantalla táctil. • Porcentaje de batería usada / satisfacción con botón de menú.
Modelo de Aceptación de Tecnología (TAM2)	Navegación: se refiere a los problemas que los participantes experimentaron cuando se utiliza el diseño de navegación estructural de la aplicación, por ejemplo, botón de inicio en el pie de página (Biederer, Arguel, Liu, & Lau, 2014). (Continua)

MODELOS DE USABILIDAD	ATRIBUTOS Y MÉTRICAS
(Biederer, Arguel, Liu, & Lau, 2014).	<p>Ubicación: se utiliza para identificar los problemas específicamente en relación con la posición de los campos y botones en la pantalla (Biederer, Arguel, Liu, & Lau, 2014).</p> <p>La visibilidad: se refiere a los problemas de los participantes al intentar hallar las funciones específicas en la pantalla, por ejemplo, tamaño de los botones (Biederer, Arguel, Liu, & Lau, 2014).</p> <p>Usabilidad: determina si el participante fue capaz de utilizar una función específica, por ejemplo, si los participantes fueron capaces de utilizar el software, para completar una tarea designada (Biederer, Arguel, Liu, & Lau, 2014).</p> <p>El color: es una categoría en la que los participantes informan acerca de gustos y disgustos de los esquemas de color utilizados en la versión móvil de la app (Biederer, Arguel, Liu, & Lau, 2014).</p>
Modelo propuesto por Ardito, 2014 (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).	<p>Este conjunto de métricas propuestas son específicamente creadas para evaluar aplicaciones de aprendizaje (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).</p> <p>La dimensión de la presentación: se ocupa sólo de los aspectos que se relacionan con las herramientas de diseño visual de la aplicación de aprendizaje (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).</p> <p>La dimensión Hipermedialidad: es un aspecto importante para los usuarios (profesores y estudiantes), ya que permite la estructuración adecuada de los materiales que permite a los usuarios seleccionar una ruta personalizada para el aprendizaje (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).</p> <p>Proactividad de la aplicación: considera la capacidad del sistema para apoyar las actividades de los usuarios. Los errores en el sistema deben ser lo más bajo posible (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).</p> <p>La dimensión actividad del usuario: se centra en la creciente necesidad del usuario y cómo la aplicación hace frente a ella (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).</p>

A continuación en la Tabla 5 se puede observar una síntesis de los atributos más evaluados en los estudios de evaluación de usabilidad de aplicaciones móviles. Estos atributos pertenecen a los modelos: Nielsen, Bertini, ISO y PACMAD.

Tabla 5

Atributos y métricas usados para medir la usabilidad en los estudios primarios

Atributo	# Estudios	Métrica
Visibilidad del estado del sistema	9	% de visibilidad, es medido por cuestionarios y varias técnicas de colección de datos
Control y libertad para el usuario	11	posibilidad del usuario de reiniciar sus acciones, es medido por cuestionarios
Consistencia y estándares	12	% o nivel de cumplimiento con las normas y estándares
Errores	17	% de tasa de errores, número de errores cometidos por tarea
Reconocimiento en lugar de Recuerdo	6	% intuitividad de la aplicación, puede ser medido con estudios longitudinales
Flexibilidad y eficiencia de uso	7	Tiempo que tarda en cargar el sistema
Diseño estético y minimalista	3	Cumplimiento de guías de interfaz de usuario, por ejemplo tamaño de letra ajustable, colores, etc., es medido por cuestionarios
Compatibilidad entre diferentes plataformas	3	Nivel o porcentaje de flexibilidad de un sistema para adaptarse a las distintas plataformas
Confiabilidad	2	Número de veces que el usuario toca controles cercanos a su objetivo accidentalmente
Eficiencia	19	Tiempo invertido en realizar la tarea, Número de toques requeridos para la completitud de la tarea
Efectividad	7	Tareas resueltas por el usuario en un tiempo limitado, % de tareas completadas con éxito al primer intento
Satisfacción	13	Nivel de confort percibido por el usuario al utilizar el sistema, es medido por cuestionarios y varias técnicas de colección de datos
Facilidad de Aprendizaje	7	Tiempo el cual le toma al usuario acostumbrarse a la aplicación eficientemente, Tiempo usado para terminar una tarea la primera vez
Carga cognitiva	5	% procesamiento cognitivo que el usuario necesita para utilizar la aplicación. Se puede medir a través de una herramienta de evaluación TLX (Task load index) de índice de carga de tareas de la NASA

Como se puede observar los tres atributos más evaluados son: Eficiencia, Errores y Satisfacción. La eficiencia es especialmente importante durante la interacción con teléfonos móviles ya que al medirla se puede saber la habilidad del usuario para cumplir sus tareas de manera exacta y rápida. Es también imprescindible conocer cuan libre de errores es la utilización del software tanto para el usuario y el desarrollador, para saber si el sistema requiere mejoras, para este fin se mide el atributo Errores. Finalmente la satisfacción juega un papel importante dentro de la usabilidad ya que es la opinión de los usuarios y su nivel de confort al utilizar el sistema, si un usuario no está satisfecho simplemente no utilizara la aplicación.

Algunos de los atributos menos evaluados son muy influyentes en la usabilidad de las aplicaciones de tipo móvil como por ejemplo: carga cognitiva y compatibilidad entre plataformas. Estos atributos son específicos de las aplicaciones móviles multiplataforma, la carga cognitiva porque los usuarios usan las aplicaciones desde cualquier sitio y se enfrentan a distintos tipos de distracciones que tienden a influir en su comportamiento y recursos cognitivos (Tsiaousis & Giaglis, 2014). Al igual la compatibilidad entre las distintas plataformas móviles porque gracias a la constante innovación de los dispositivos móviles en el mercado, las aplicaciones deben ser más flexibles para adaptarse a las distintas plataformas y dispositivos (Gresse von Wangenheim, y otros, 2016).

RQ4: ¿Qué tipo de aplicaciones móviles son evaluadas y a través de qué actividades?

Dentro de los estudios primarios se encontró que diferentes tipos de aplicaciones son evaluadas, la mayoría con un 41% de estudios primarios son aplicaciones con respecto a la salud para el monitoreo de pacientes, registros médicos, control de enfermedades como la diabetes, obesidad, etc. Un 23% corresponde a aplicaciones móviles varias por ejemplo: una aplicación de información de museos, aplicaciones de reconocimiento de plantas, aplicaciones espirituales, entre otras. El 8% pertenece a las aplicaciones de aprendizaje o enseñanza y aplicaciones de tipo red social. Finalmente el 3% de aplicaciones evaluadas pertenecen a la categoría de: aplicaciones multimodales y aplicaciones de privacidad; se encontró apenas 2% de estudios primarios evalúan el siguiente tipo de aplicaciones: nativas, portales web, aplicaciones de banca móvil. A continuación se presenta la Figura 11, sobre el tipo de aplicaciones evaluadas en los estudios primarios, considerando que se tomó en cuenta 39 de los 42 estudios primarios, porque los tres restantes no mostraban información sobre las aplicaciones evaluadas o simplemente son estudios referentes a la teoría existente sobre determinados método de evaluación de usabilidad de aplicaciones móviles.

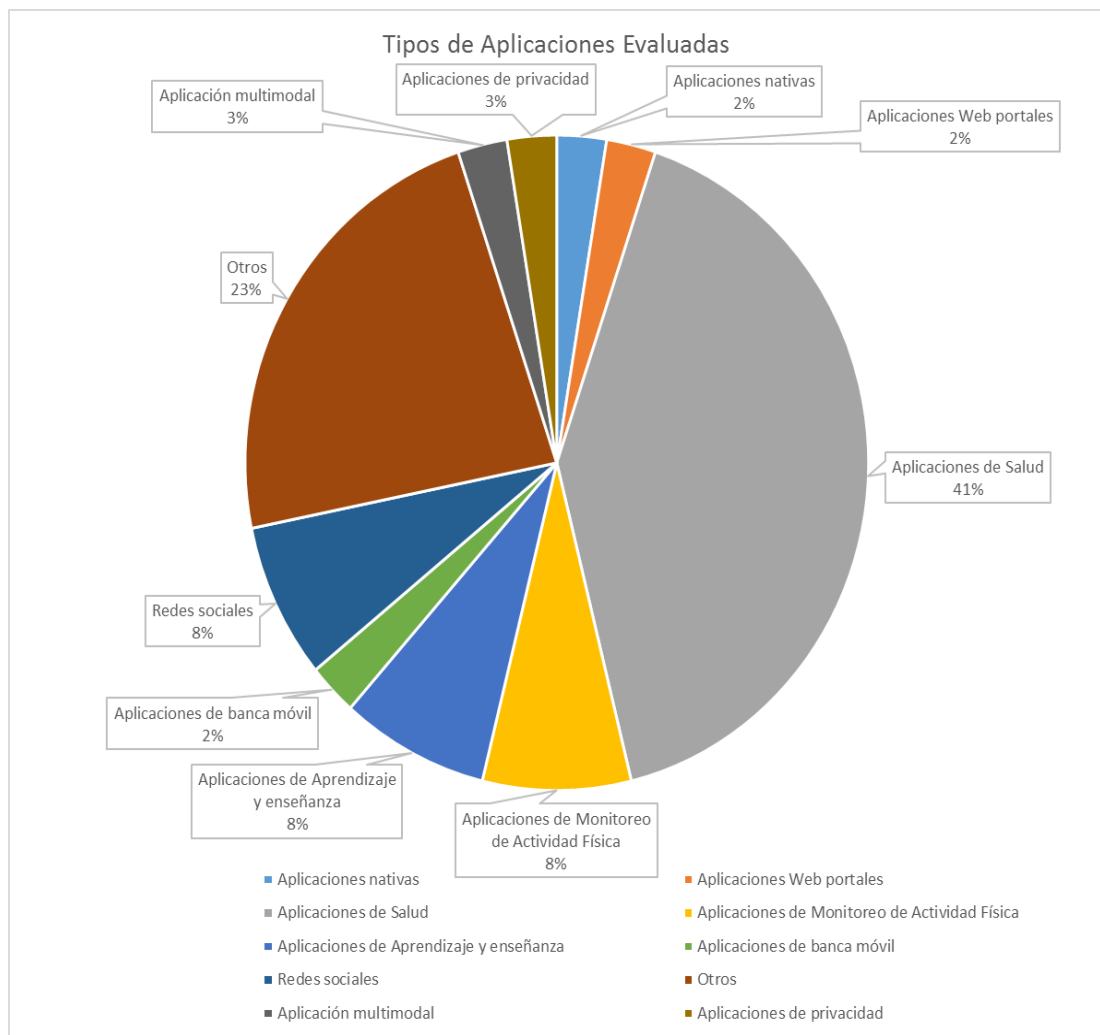


Figura 11. Tipos de Aplicaciones evaluadas

Dependiendo de la metodología para evaluar usabilidad de las aplicaciones móviles, los autores de los estudios establecen diferentes tareas concretas que varían su nivel de complejidad, las mismas que son realizadas por usuarios reales, aleatorios o expertos en usabilidad. Algunos estudios no especifican las tareas realizadas para evaluar la usabilidad de las aplicaciones.

RQ5: ¿Qué ventajas e inconvenientes se generan al momento de evaluar la usabilidad de las aplicaciones móviles en los estudios?

Ventajas:

- Las heurísticas de Nielsen (1994) son un importante punto de partida cuando se desea evaluar la usabilidad de las aplicaciones, pero en el contexto móvil

estas heurísticas necesitan ser adaptadas a las características específicas de los dispositivos móviles (Gresse von Wangenheim, y otros, 2016).

- La evaluación heurística tiene un bajo costo, es relativamente efectiva y rápida para evaluar la usabilidad de aplicaciones móviles, es el método más popularmente usado desde 1990 (Joyce, Lilley, Barker, & Jefferies, 2015).
- La Evaluación heurística no requiere planeación previa (Yáñez Gómez, Cascado Caballero, & Sevillano, 2014). Además no es mandatorio para los evaluadores que sean expertos en usabilidad (Yáñez Gómez, Cascado Caballero, & Sevillano, 2014).
- Los métodos como Recorrido cognitivo y evaluación heurística han hecho posible identificar problemas de usabilidad a los que se enfrentan los usuarios novatos y experimentados, mientras se centra en tareas específicas y también en el diseño de interfaz (Saleh, Binti Isamil, & Binti Fabil, 2015).
- Las evaluaciones de campo, en un ambiente real y usuarios finales, facilita la investigación sobre usabilidad de cualquier tipo de aplicación móvil (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015).
- El marco de diseño CRUISER que integra el uso de desarrollo ágil de aplicaciones con la Interacción Humano-Máquina, mejora la facilidad de uso de las aplicaciones. Esto se realiza mediante el uso de los primeros prototipos, diseños que luego son refinados en etapas posteriores del ciclo de vida del producto software. La facilidad de uso de la aplicación es aún mayor con la iteración entre todas las etapas del desarrollo del sistema desde los requerimientos iniciales del producto software (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).
- La combinación de pruebas de laboratorio con las de campo, y las pruebas "basadas en el usuario" con las que son "basadas en expertos", revelan una amplia gama de problemas de usabilidad. La evaluación heurística y "think aloud" revelaron información más valiosa y detallada sobre las interfaces y textos, mientras que la evaluación de campo en la vida real reveló cuestiones prácticas, tales como, problemas de conectividad y facilidad de uso general. En la evaluación de usabilidad de aplicaciones móviles es útil dar la importancia a un enfoque de métodos mixtos, ya que diferentes cuestiones se dieron a

conocer en el laboratorio en comparación con la prueba de campo. Las interfaces pueden ser muy eficaces, eficientes y deseables en una situación de laboratorio, pero si falla la comunicación entre los diferentes componentes de la herramienta en una situación de la vida real, la satisfacción será baja (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014).

- El protocolo para evaluación de usabilidad “think aloud” provee las opiniones honestas de los usuarios durante el uso de la aplicación (Skorin-Kapov, Dobrijevic, & Piplica, 2014).
- La evaluación automatizada de usabilidad AUT genera automáticamente registros de interacciones y da una visión en conjunto de todos los problemas por cada patrón con un gráfico de interacción apropiado. Esta evaluación no necesariamente requiere expertos o sujetos, de manera que reduce costo, tiempo y gastos de personal (Kluth, Krempels, & Samsel, 2014).
- El protocolo “think aloud”, donde los usuarios informan su comprensión e interpretación sobre cada función del sistema, puede proporcionar una explicación más detallada del comportamiento y la comprensión de los usuarios (Lyles, Sarkar, & Osborn, 2014).
- Los cuestionarios son un punto de partida potencialmente útil para la comprensión de las percepciones de los usuarios sobre la facilidad de uso y su satisfacción con la aplicación (Lyles, Sarkar, & Osborn, 2014).
- Las evaluaciones de usabilidad basadas en expertos y basadas en modelos son menos costosas, con estos tipos de evaluaciones muchas interfaces de usuario pueden ser testeadas (Hussain, Hashim, & Nordin, 2014).
- A partir de la prueba de la usabilidad de aplicaciones con usuarios potenciales, se obtienen datos significativos que podrían ser utilizados por los desarrolladores para refinar el contenido y la presentación de la aplicación antes de su despliegue (Reynoldson, y otros, 2014).
- En un escenario de contexto real, es decir, en una evaluación de campo es posible detectar problemas difíciles o imposibles de descubrir y simular en un laboratorio ya que los parámetros contextuales influyen el rendimiento de los participantes (López-Gil, Urretavizcaya, Fernández-Castro, & Losada, 2014).

- La incorporación una evaluación de usabilidad de campo en un proceso de desarrollo ágil como Intermod ha producido un impacto significativo, ya que se ha demostrado que se pueden integrar de manera coherente (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).
- Los resultados de evaluar un primer prototipo de una aplicación móvil se utilizan para ajustar el manual de usuario y desarrollar un segundo prototipo de la aplicación mejorada (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014), es decir, mediante la implementación de las recomendaciones de los expertos y los usuarios, los problemas de usabilidad se pueden eliminar para que en la próxima fase de ejecución de la aplicación pueda producir una aplicación móvil que maximice la experiencia del usuario (Luhur & Widjaja, 2014).

Desventajas:

- En algunos estudios de usabilidad, los problemas más frecuentes se relacionan a que las aplicaciones no cumplen los principios de diseño de Norman y Nielsen, como la visibilidad y retroalimentación (Faliagka, Rigou, & Sirmakessis, 2014).
- Sin modificar el método de evaluación heurística podría no ser efectivo al momento de evaluar la usabilidad de aplicaciones móviles (Joyce, Lilley, Barker, & Jefferies, 2015).
- El modelo QUIM para evaluar usabilidad de aplicaciones móviles no es óptimo aún necesita validación (Saleh, Binti Isamil, & Binti Fabil, 2015).
- El modelo Music para evaluar usabilidad de aplicaciones móviles, mide el rendimiento en función del usuario, pero no puede indicar con precisión los atributos de usabilidad, como la facilidad de aprendizaje y satisfacción (Saleh, Binti Isamil, & Binti Fabil, 2015).
- Los modelos de evaluación de usabilidad todavía son limitados y restringidos debido a que los diseñadores y desarrolladores muchas veces no se interesan en el área de la interacción humano-computador y producen inconvenientes a los usuarios de las aplicaciones al no considerar sus necesidades y preferencias (Saleh, Binti Isamil, & Binti Fabil, 2015). Los modelos de evaluación de la

usabilidad existentes deben ser personalizados para abordar el enfoque móvil (Angulo Cevallos, 2014).

- El modelo ISO 9126-1 propone el atributo atractivo de la aplicación, una característica que es la menos evaluada en los estudios, a pesar de que no debería ser ignorada porque causa un efecto positivo en la usabilidad y afecta en la decisión de empezar a usar una aplicación (Cruz Zapata, Fernández, Idri, & Toval, 2015).
- Un problema a considerar importante es la falta de procesos iterativos de diseño-evaluación, lo mismo que causa que la calidad y usabilidad de las aplicaciones disminuya (Cruz Zapata, Fernández, Idri, & Toval, 2015).
- Al evaluar la usabilidad de aplicaciones para usuarios adultos mayores, las actividades como entrevistas son muy complejas de realizar, por lo que es necesario un estudio que establezca una forma diferente y apropiada de trabajar con este tipo de usuarios (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015).
- El método de evaluación automatizada AUT, es únicamente para aplicaciones de la plataforma iOS (Kluth, Krempels, & Samsel, 2014), no existe en la literatura un enfoque automatizado de evaluación de usabilidad multiplataforma.
- Es posible que la usabilidad sea calificada como más positiva en las pruebas de laboratorio y entrevistas que durante la prueba de campo, debido al sesgo de deseabilidad social (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014). Se observa que las respuestas socialmente deseables se dan con más frecuencia en las entrevistas cara a cara. Los usuarios pueden haber querido demostrar que tenían las capacidades necesarias para utilizar el sistema o querían satisfacer al investigador (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014).
- En la literatura de HCI, se reconoce el impacto del contexto ambiental como atributo en la usabilidad móvil, sin embargo falla al dar un nivel medible de la influencia de cada tipo de distracción, este tema es muy amplio y complejo para ser preciso (Tsiaousis & Giaglis, 2014). Las distracciones del ambiente tienden

a reducir visibilidad del usuario, memoria y recursos cognitivos (Tsiaousis & Giaglis, 2014).

- En varios estudios los autores coinciden como desventaja el número pequeño de participantes en sus evaluaciones de usabilidad (Mirkovic, Kaufman, & Ruland, 2014).
- La mayoría de violaciones realizadas por los usuarios son: la facilidad de ingreso, legibilidad de pantalla y comprensibilidad (Mirkovic, Kaufman, & Ruland, 2014).
- En el clásico método de observación de usuarios, se hace difícil capturar todos los ingresos de datos en la interfaz de la pantalla táctil, mientras que la obtención de los datos es tedioso, porque hay que clasificar manualmente los problemas dentro de los patrones de diseño de interacción definidos (Kluth, Krempels, & Samsel, 2014).
- En la era digital, especialmente cuando se trabaja en aplicaciones móviles, la puntualidad de los datos es un problema general. El número y la calidad de las aplicaciones están cambiando constantemente (Kalz, y otros, 2014).
- Los teléfonos inteligentes tienen un impacto en la apariencia y la usabilidad de sus aplicaciones, por tanto, pueden afectar a las experiencias de los participantes (Kalz, y otros, 2014).
- En un cuestionario basado en directrices de diseño, respondido por expertos puede pasar que los resultados de usabilidad estén sesgados en la dirección del interés general de los investigadores, que no necesariamente corresponden a la opinión de usuarios reales (Cruz Zapata, Hernandez Niñirola, Idri, Fernandez-Aleman, & Toval, 2014).
- Las pruebas de laboratorio parecen ser incapaces de asegurar por completo la facilidad de uso en el paradigma móvil que es el contexto de uso (Yáñez Gómez, Cascado Caballero, & Sevillano, 2014).
- En los cuestionarios las percepciones de los usuarios pueden no proporcionar una perspectiva completa de la facilidad de uso (Lyles, Sarkar, & Osborn, 2014).
- La evaluación con usuarios es costosa y consume tiempo (Hussain, Hashim, & Nordin, 2014).

- La eficiencia relativa de usuario de la aplicación se comparó con la de un usuario experto y un novato, el tiempo necesario para que los participantes novatos completen las tareas era de una a tres veces mayor que la del usuario experto, característica importante que se debería tomar en cuenta en las evaluaciones de usabilidad de aplicaciones móviles (O'Malley, Dowdall, Burls, Perry, & Curran, 2014).
- En la evaluación de campo, algunas características específicas de teléfonos móviles pueden ser mejor controladas en un laboratorio, como por ejemplo el cambio de orientación, el movimiento físico de los usuarios que tiene un impacto significativo, etc.; los cambios voluntarios del usuario pueden ser fácilmente controlados en un laboratorio (López-Gil, Urretavizcaya, Fernández-Castro, & Losada, 2014).

RQ6: ¿Con que frecuencia se consideran los factores de usabilidad (usuarios, tareas, contexto de uso) en los estudios existentes?

Los factores de usabilidad que son considerados en cada estudio se refiere a la importancia cada autor dio a los factores que definen un producto software como usable: usuarios, tareas y contexto de uso (Saleh, Binti Isamil, & Binti Fabil, 2015), en el proceso de evaluación, estos factores se detallan a continuación:

- **Usuarios:** son las personas, quienes van a usar la aplicación.
- **Tareas:** son las funcionalidades de la aplicación.
- **Contexto de uso:** se refiere a los factores del entorno donde se va a usar la aplicación móvil (Olsina, Santos, & Lew, 2014). Se consideran varias sub-entidades y propiedades de contexto de uso que incluyen: **Actividad** (qué están haciendo los usuarios en el momento de uso, tienen una influencia significativa en la capacidad de atención del usuario), **Día / hora del día** (el día y la hora pueden afectar lo que un usuario está haciendo, y el nivel de luz natural también. A diferencia de las aplicaciones de escritorio que normalmente se accede en el interior, el uso de aplicaciones móviles es particularmente sensible a este factor contextual que influye en la visibilidad), **Localización** (la ubicación del usuario influye en muchos elementos; por ejemplo en el interior, al aire libre, en un auto o en un ascensor, características que también pueden estar relacionadas con la actividad de los usuarios), **Rendimiento de la Red**

(la velocidad a la que un archivo y datos de las aplicaciones se descargan va a tener un gran impacto debido a la disminución de la capacidad de atención de los usuarios), **Perfil de usuario** (la creciente complejidad del software combinado con una demografía de usuarios que son ancianos tiene un efecto importante en la usabilidad de las aplicaciones móviles) y **Dispositivo** (el tamaño, el tipo del dispositivo y sus características físicas influyen en lo que el usuario puede ver o no ver, así como la colocación y el número de controles y widgets en las pantallas son reducidos) (Olsina, Santos, & Lew, 2014).

En la Tabla 6, se puede observar cada estudio primario y cuáles son los factores que consideran los autores al momento de evaluar la usabilidad.

El factor tareas es considerado en 5 estudios para ser exactos. Los autores al medir la usabilidad se ayudan de expertos o investigadores, utilizan una lista de comprobación y mientras realizan las tareas específicas verifican que la aplicación cumpla con las heurísticas o características de la lista (Gresse von Wangenheim, y otros, 2016), (Tomé Klock & Gasparini, 2015), (Joyce, Lilley, Barker, & Jefferies, 2015) o también evalúan la usabilidad con un cuestionario basado en directrices de diseño de interfaz (Cruz Zapata, Hernandez Niñirola, Idri, Fernandez-Aleman, & Toval, 2014), (Arnhold, Quade, & Kirch, 2014).

Mientras que en 17 estudios primarios, se consideran los factores de usabilidad: tareas y usuarios, y dejan un poco de lado el contexto de uso. Esto se debe a que las evaluaciones son realizadas con usuarios finales a través de tareas concretas y bajo un ambiente controlado sin distracción alguna. En estos estudios los métodos de evaluación de usabilidad usados son: pruebas de laboratorio (Faliagka, Rigou, & Sirmakessis, 2014), (Skorin-Kapov, Dobrijevic, & Piplica, 2014), el protocolo “Think Aloud”(Choi, Suk Lee, & Ho Park, 2015), los cuestionarios de usabilidad (Nawi, Isa Hamzah, Chy Ren, & Tamuri, 2015), (Mirkovic, Kaufman, & Ruland, 2014), (Foo, y otros, 2015), (Abdul, Hussain, & Zainol, 2014) y recorrido cognitivo con usuarios reales como el estudio de (Lira, de Souza, Ferreira, & Carvalho, 2014). Generalmente las evaluaciones heurísticas no toman en cuenta estos dos factores, sin embargo en un estudio se mostró una evaluación heurística híbrida (Vélez, Boakye, Kanter, & Bakken, 2014) donde se planea un proceso de evaluación con expertos, usuarios y tomando en cuenta las tareas y contexto de uso.

Algunos estudios no capturan o controlan el contexto de uso. De los cuarenta y dos estudios primarios en 17 toman en cuenta los tres factores de usabilidad al evaluar las diferentes aplicaciones móviles. Por ejemplo el modelo PACMAD considera los tres factores y propone el atributo carga cognitiva, que se refiere al procesamiento cognitivo que el usuario necesita al utilizar la aplicación (Saleh, Binti Isamil, & Binti Fabil, 2015). Se puede decir que, al considerar los tres factores de usabilidad: tareas, usuarios y contexto de uso en el proceso de evaluación se podría reconocer diferentes cuestiones respecto a la usabilidad (Van Der Weegen, Verwey, Tange, Spreuwwenberg, & de Witte, 2014), (Hsiu Ching, 2015) y contribuye a identificar problemas del mundo real que podrían existir al usar aplicaciones, tales como, señales pobres de celular, mala iluminación, y las interrupciones del flujo de trabajo frecuentes (Vélez, Boakye, Kanter, & Bakken, 2014). Las evaluaciones de campo que son realizadas en un ambiente más natural, en el que los usuarios habitualmente van a usar la aplicación, toman en cuenta también estos tres factores de usabilidad (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015).

Tabla 6

Factores de usabilidad considerados en cada estudio primario.

FACTOR DE ESTUDIO USABILIDAD		TOTAL
TAREAS	EP1 - EP4 - EP8 - EP24 - EP30	5
USUARIOS Y TAREAS	EP3 - EP6 - EP11 - EP12 - EP13 - EP17 - EP19 - EP21 - EP22 - EP29 - EP31 - EP33 - EP34 - EP35 - EP37 - EP38 - EP42	17
USUARIOS, TAREAS Y CONTEXTO DE USO	EP5 - EP7 - EP9 - EP10 - EP14 - EP15 - EP16- EP20 - EP23 - EP25 -EP27 - EP28 - EP32 - EP36 - EP39 - EP40 - EP41	17
NO ESPECIFICA	EP2 - EP18 - EP26	3
TOTAL DE ESTUDIOS		42

Finalmente en 3 estudios primarios no se encontró información alguna respecto a los factores de usabilidad tomados en cuenta en las evaluaciones, algunos de ellos porque solo proponen guías, heurísticas y métodos de evaluación de usabilidad que no

son probados en una evaluación real, o también porque son estudios sistemáticos de mapeo referentes a otros estudios que no mencionan detalles sobre los factores de usabilidad.

RQ7: ¿Con que frecuencia se considera el diseño de interacción respecto a los factores de usabilidad en el desarrollo de aplicaciones móviles?

En los estudios primarios se pudo observar que existieron problemas de diseño de interacción porque los desarrolladores de las aplicaciones móviles evaluadas en los estudios, cuando crean el diseño de interacción muchas veces no toman en cuenta los factores que determinan la usabilidad de un producto software como: usuarios (para quien es la aplicación), tareas (objetivos que debe cumplir la aplicación) y contexto de uso (donde se va a usar la aplicación), esto produce una afectación en la facilidad de uso de una aplicación móvil. En la Tabla 7, se puede observar la frecuencia con la que los factores usabilidad son tomados en cuenta en los estudios primarios al desarrollar el diseño de interacción de sus aplicaciones móviles.

La mayoría de estudios primarios 26, muestran solo el proceso de evaluación de usabilidad de determinadas aplicaciones móviles que ya están en el mercado y no son desarrolladas por ellos, así que, no mencionan detalles específicos del proceso de desarrollo y diseño de interacción. Sin embargo en estos estudios al evaluar la usabilidad se dan problemas que tienen que ver con el diseño de interacción, y muestran la falta de atención de los desarrolladores hacia los usuarios, tareas y contexto de uso de sus aplicaciones.

Por ejemplo el factor usuarios, al momento de desarrollar una aplicación móvil, es importante tomar en cuenta a quien va dirigida la aplicación, este es el caso de aplicaciones para usuarios ancianos en las que se pudo observar que en ciertos casos no toman en cuenta las capacidades y necesidades de este tipo de usuario. Los ancianos son un amplio grupo de usuarios que podrían beneficiarse de aplicaciones por ejemplo para la diabetes. Sin embargo, varios estudios han demostrado una falta de aceptación y el uso insatisfactorio de las tecnologías móviles innovadoras entre este grupo de personas (Arnhold, Quade, & Kirch, 2014). Una potencial razón es la escasa consideración de los requisitos y preferencias de usabilidad de las personas mayores, en cuanto a interfaz de la aplicación, en general, sus experiencias en el manejo de los dispositivos móviles y aplicaciones son frecuentemente limitadas (Arnhold, Quade, &

Kirch, 2014). También es importante mencionar que, la noción de lo que constituye una tarea más simple que otra, puede depender de varios parámetros contextuales o del tipo de usuario, como por ejemplo, una persona con artritis puede percibir el ingreso manual del peso en un dispositivo móvil como una tarea más difícil que la realización de la medición de la presión arterial donde las mediciones se registran automáticamente (Skorin-Kapov, Dobrijevic, & Piplica, 2014). Por lo tanto, se requiere un examen de los factores de influencia relacionados con los usuarios, por ejemplo sus capacidades físicas (Skorin-Kapov, Dobrijevic, & Piplica, 2014).

En los estudios de (Cruz Zapata, Fernández, Idri, & Toval, 2015) y (Akma Ahmad, Zainal, Abdul Razak, Wan Adnan, & Osman, 2015) se aconseja que los desarrolladores deben crear aplicaciones móviles más útiles, no sólo para las personas ancianas, sino también para los usuarios novatos.

Otro ejemplo de esta deficiencia de crear el diseño de interacción sin pensar en los usuarios, son las aplicaciones de privacidad, las cuales no son intuitivas, su interfaz es inconsistente, ciertas características de la aplicación no están visibles claramente a los usuarios y hablan un lenguaje muy técnico que incluso para usuarios expertos no es entendible (Assal, Hurtado, Imran, & Chiasson, 2015).

En 5 estudios primarios si tomaron en cuenta los factores usuarios y tareas al momento de desarrollar el diseño de interacción. En un estudio los investigadores crearon un primer prototipo de la aplicación, evaluaron su usabilidad e interfaz y tomando en cuenta los problemas o recomendaciones dadas por los usuarios y expertos crearon un nuevo prototipo de la aplicación con mejoras (Luhur & Widjaja, 2014), en el segundo estudio los investigadores crean la aplicación móvil de una manera iterativa donde se iba evaluando su facilidad de uso con usuarios y tareas preestablecidas (Biederer, Arguel, Liu, & Lau, 2014), y en los demás estudios los investigadores añaden tareas específicas según el tipo de aplicación con usuarios puntuales para predecir sus necesidades y comportamientos (Maren Witt, 2014), (Beul-Leusmann, y otros, 2014), (O'Malley, Dowdall, Burls, Perry, & Curran, 2014).

Finalmente en 11 estudios primarios los investigadores si tomaron en cuenta los tres factores de usabilidad: usuarios, tareas y contexto de uso, al crear el diseño de interacción de sus aplicaciones. Los autores controlaron estos factores, considerando el diseño de interacción en las etapas de desarrollo de software, usando prototipos y

primeros diseños con usuarios reales en un ambiente natural, para que en etapas posteriores sean refinados (Hussain, Saleh, Taher, Ahmed, & Lammasha, 2015).

Pocos estudios de HCI estudian los efectos de las distracciones ambientales en la usabilidad de aplicaciones móviles (Tsiaousis & Giaglis, 2014). En el desarrollo es imposible cubrir todos los escenarios y ambientes donde van a ser usadas las aplicaciones (Joyce, Lilley, Barker, & Jefferies, 2015). Todavía un conjunto de heurísticas dedicadas a este dominio deben incluir una forma de contexto de uso así esto será menos considerado por evaluadores de usabilidad (Joyce, Lilley, Barker, & Jefferies, 2015).

Tabla 7

Factores de usabilidad tomados en cuenta en el desarrollo de aplicaciones móviles de los estudios primarios

FACTOR DE USABILIDAD	ESTUDIO	TOTAL
USUARIOS Y TAREAS	EP27 - EP32 – EP33 - EP35 - EP38	5
USUARIOS, TAREAS Y CONTEXTO DE USO	EP5 - EP9 - EP10 - EP19 - EP20 –EP25 – EP28 - EP36 - EP37 - EP40 - EP41	11
NO ESPECIFICA	EP1 - EP2 - EP3 - EP4 - EP6 - EP7 - EP8 - EP11 - EP12 - EP13 - EP14 - EP15 - EP16 - EP17 - EP18 - EP21 - EP22 - EP23 - EP24 - EP26 - EP28 - EP29 - EP30 -EP31 - EP34 - EP39 - EP42	26
TOTAL DE ESTUDIOS		42

Como desarrolladores debemos prestar más atención a las interfaces de usuario, tomando en cuenta los siguientes factores: usuarios, permitir a los mismos que intervengan dentro de la etapa de desarrollo de las aplicaciones, por ejemplo con prototipos de las pantallas, para estudiar que sugerencias y necesidades podrían tener,

el segundo factor a considerar es el contexto de uso, donde se va a usar la aplicación, que características de la interfaz deben tener las aplicaciones para que se ajusten a los diferentes ambientes donde será usada. Una interfaz confusa puede producir dificultades en el aprendizaje de la aplicación (Enriquez & Casas, 2013).

RQ8: ¿En qué medida difiere el diseño de interacción de las app móviles entre las plataformas existentes?

El diseño de interacción entre las aplicaciones móviles difiere en gran medida entre las plataformas existentes, a pesar de que cada dispositivo es necesario que cumpla con los principios de usabilidad independientes de la plataforma, en las últimas décadas se ha podido observar un efecto heterogéneo en las interfaces de sus aplicaciones móviles (Faliagka, Rigou, & Sirmakessis, 2014).

En general los estándares de las plataformas e industrias pueden ser usados como guías para aplicar una interacción de usuario consistente (Joyce, Lilley, Barker, & Jefferies, 2015), sin embargo los desarrolladores no saben si utilizar la filosofía de la interacción propia de cada plataforma móvil o la creación de diseños de interacción multiplataforma (Barea, Ferré, & Villarroel, 2013). A menudo las directrices específicas de interfaz de usuario se aplican en el desarrollo de las aplicaciones móviles, según la plataforma móvil (Cruz Zapata, Hernandez Niñirola, Idri, Fernandez-Aleman, & Toval, 2014). IOS, con su lenguaje de programación Objective-C y su entorno de desarrollo Xcode, siempre ha estado muy centrado en la experiencia del usuario lo cual facilita la creación de interfaces de usuario atractivas y dinámicas (Cruz Zapata, Hernandez Niñirola, Idri, Fernandez-Aleman, & Toval, 2014). Por el contrario, el constructor de interfaces más comúnmente utilizado para Android (Android Eclipse) y el vínculo entre el código y la interfaz de usuario precisa mejoras si se desea ser tan bueno como Xcode y su enfoque del guion gráfico (Cruz Zapata, Hernandez Niñirola, Idri, Fernandez-Aleman, & Toval, 2014).

En los estudios primarios, se observó la influencia que provoca la diferencia del diseño de interacción entre una misma aplicación en distintas plataformas, ya que al utilizar una aplicación móvil los participantes que son dueños del mismo tipo de dispositivo móvil y tienen previa experiencia con el mismo, muestran un incremento del nivel de autoconfianza y menos peticiones de ayuda mientras realizan las tareas (Mirkovic, Kaufman, & Ruland, 2014). En cambio quienes no tenían o era limitada su

experiencia con este tipo de dispositivos, necesitaron la orientación de rutina antes de usar una aplicación (Mirkovic, Kaufman, & Ruland, 2014). Es decir la experiencia de los usuarios con la aplicación móvil dependerá en cierto punto de la plataforma y dispositivo con el que los usuarios estén acostumbrados a interactuar.

Esta experiencia del usuario se adaptaría mejor en un enfoque de aplicaciones móviles multiplataforma, donde la aplicación conserva una apariencia similar pero con un diseño de interacción que sigue las convenciones de la plataforma, porque con respecto al equipo de desarrollo de software hay un esfuerzo ligeramente menor de mantenimiento en comparación con el enfoque dependiente de la plataforma, debido a que los dos diseños comparten el mismo esquema básico de diseño de interacción y también porque favorece en la usabilidad ya que provoca una reacción positiva por parte del usuario sin dejar de ofrecer cierta coherencia entre las distintas plataformas (Barea, Ferré, & Villarroel, 2013).

En los últimos tiempos por la constante innovación de los diferentes dispositivos móviles en el mercado, las aplicaciones cada vez deberían ser más flexibles para adaptarse a las distintas plataformas y dispositivos (Van Der Weegen, Verwey, Tange, Spreeuwenberg, & de Witte, 2014). Para este fin se tiene que incrementar: las funcionalidades y la integración e interactividad de las aplicaciones móviles, garantizando así una atención especial en la comprensión de las diferencias entre aplicaciones de otras plataformas, la experiencia de la facilidad de uso y del punto de vista usuario, porque los requisitos del usuario, expectativas y comportamientos son algo diferente dependiendo de la plataforma móvil con la que interactúen (Olsina, Santos, & Lew, 2014).

En resumen, se ha extraído los datos que responden a las preguntas de investigación por cada estudio primario. En general, sintetizando algunos de los hallazgos obtenidos de las preguntas de investigación están que: la mayoría de técnicas usadas en la evaluación de aplicaciones móviles son modificaciones de las metodologías para evaluar aplicaciones de escritorio; muchas de las aplicaciones evaluadas son respecto a la salud; tan solo existen dos estudios que utilicen un framework automatizado para la evaluación de usabilidad; no todos los autores de los estudios primarios toman en cuenta el factor contexto de uso dentro del proceso de evaluación de usabilidad; en cuanto al diseño de interacción se pudo conocer que no

existe una guía de diseño de interfaz común entre las plataformas, lo cual provoca que las interfaces sean distintas y confundan al usuario, por tanto disminuye la usabilidad de la aplicación móvil. El análisis de estos resultados más detalladamente se presentan en la siguiente sección de síntesis de resultados de la SLR sobre repositorios.

5.3 Síntesis de Resultados de la SLR

Este análisis es una síntesis de las preguntas de investigación RQ1, RQ2 y RQ5, que se consideró necesario para resumir y comparar los resultados más importantes hallados en la Revisión sistemática de literatura.

Los resultados de la SLR sobre los repositorios, han sido resumidos en un modelo conceptual, que parte de la evaluación de usabilidad de aplicaciones móviles multiplataforma, que es el tema principal de la investigación. Este mapa conceptual consta de: los tipos de metodologías y modelos usados para evaluar este tipo de aplicaciones, sus características, ventajas y desventajas identificadas en las investigaciones. En caso de los modelos se puede observar los atributos específicos que cada uno denota. A continuación en la Figura 12, se puede observar en resumen la revisión.

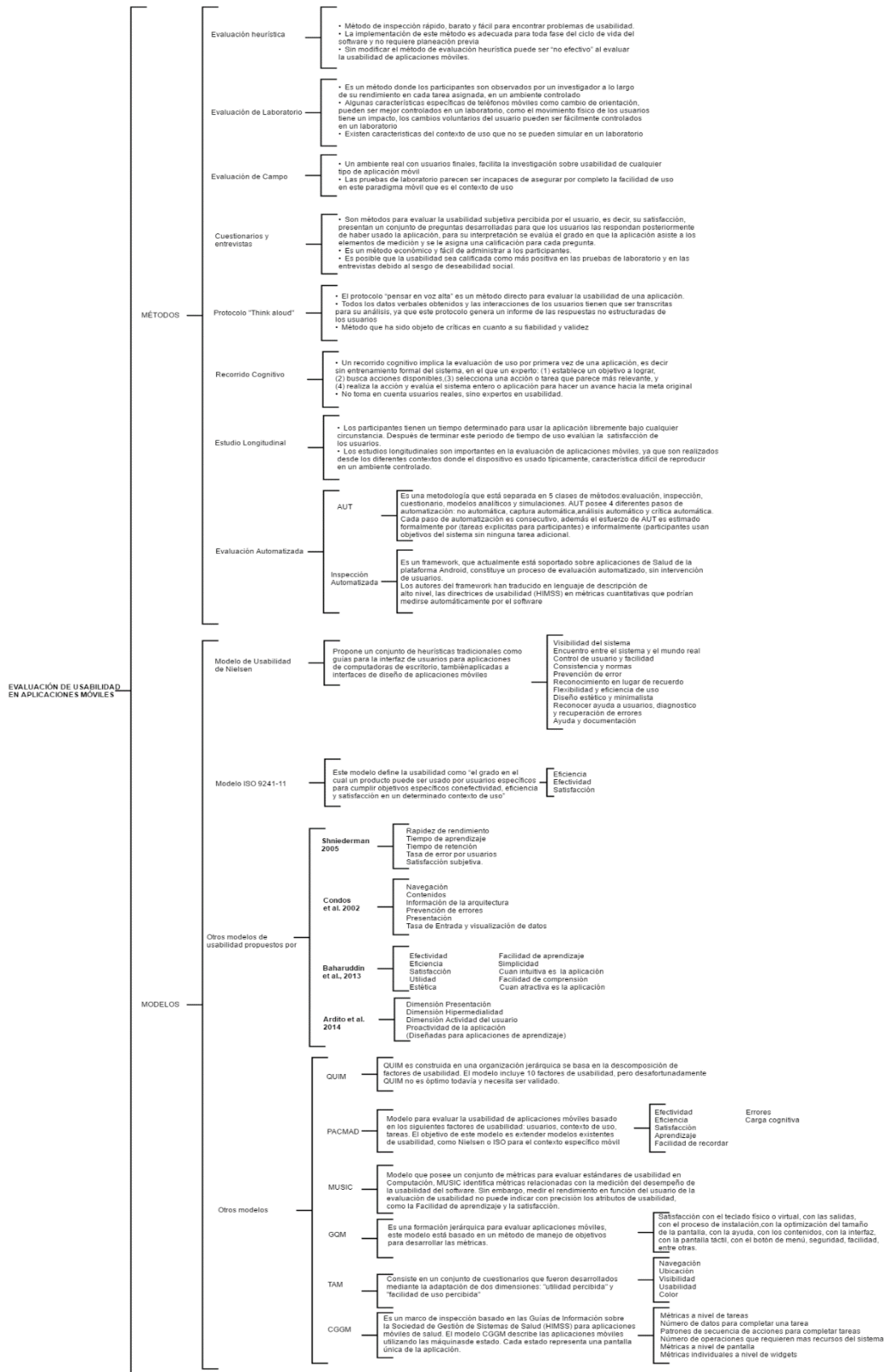


Figura 12. Síntesis de Resultados de los Estudios Primarios

Dentro de la evaluación de usabilidad de aplicaciones móviles existen muchas metodologías y modelos para ser aplicados, sin embargo cada una tiene sus pros y contras. Comenzando con la evaluación heurística si bien es cierto es un método de inspección rápido, barato y fácil que requiere un grupo de expertos que evalúen la conformidad del diseño de la aplicación con un conjunto de heurísticas, sin embargo no toman en cuenta los usuarios que sería lo ideal si se desea obtener datos más específicos. A parte las heurísticas más usadas son las de Nielsen y Molich, las mismas necesitan ser adaptadas al contexto móvil, ya que no son aplicables directamente. En el método de evaluación de usabilidad de laboratorio, donde los participantes son observados mientras realizan las tareas asignadas existen elementos del ambiente que no pueden ser simulados, por tanto, los usuarios pueden realizar las tareas de forma más rápida y eficiente al no tener distracciones. En cambio en una evaluación de campo si se evalúa a los participantes en un contexto de uso real, pero en este proceso de evaluación es difícil para el investigador observar las expresiones faciales, las veces que comete errores al tocar la pantalla, o el cambio de orientación de la pantalla, factores que en un laboratorio pueden ser mejor controlados. En los métodos de evaluación observacionales: de campo y de laboratorio se hace difícil para los evaluadores capturar todos los ingresos de datos en la interfaz de la pantalla táctil, mientras que la obtención de los datos es tediosa, porque hay que clasificar manualmente los problemas dentro de los patrones de diseño de interacción especificados.

Existen también métodos para evaluar la usabilidad subjetiva, es decir, la percibida por el usuario, a través de un conjunto de preguntas desarrolladas para los usuarios, estos métodos son las encuestas y entrevistas, las mismas que son evaluadas a través de escalas. Estos métodos son importantes para obtener opiniones significativas de los usuarios finales de las aplicaciones móviles, pero no capturan ciertos elementos que definen la usabilidad de una aplicación como por ejemplo eficiencia, eficacia, etc. Una desventaja de la aplicación de estas técnicas es que puede ser posible que la usabilidad sea calificada como más positiva, debido al sesgo de deseabilidad social. En otros campos, se observa que las respuestas socialmente deseables se dan con más frecuencia en las entrevistas cara a cara. En el método de estudio longitudinal los participantes tienen un tiempo determinado para usar la aplicación libremente, y

después evalúan la satisfacción de los usuarios, al igual que los cuestionarios este método solo captura datos subjetivos del usuario, pero en un contexto de uso real.

Existe un protocolo de evaluación de usabilidad llamado “Think aloud”, que es un método directo para evaluar la usabilidad, en el que los datos verbales de los usuarios son transcritos de manera no estructurada, razón por la cual este protocolo ha sido objeto de críticas en cuanto a su fiabilidad. Generalmente este protocolo requiere ser combinado con otras técnicas de evaluación como con la de laboratorio y campo.

Otro método para evaluar la usabilidad de aplicaciones móviles es el método de Recorrido Cognitivo, que evalúa el uso de una aplicación por primera vez, el experto establece el objetivo a lograr, busca las acciones o caminos disponibles, selecciona y realiza una acción que lo conducirá hacia la meta.

Dentro de la literatura existen escasos métodos de evaluaciones automatizadas, entre estas tenemos AUT, que es una metodología que está separada en 5 clases de métodos: evaluación, inspección, cuestionario, modelos analíticos y simulaciones. AUT posee 4 diferentes pasos de automatización: no automática, captura automática, análisis automático y crítica automática. Otro método automatizado de evaluación de usabilidad encontrado en la literatura se basa en un framework de inspección automatizada, que es soportado para Android y constituye un proceso de evaluación automatizado de aplicaciones médicas o de salud.

Los métodos existentes han sido combinados por la mayoría de autores de los estudios primarios, para la obtención de datos más relevantes. En caso de las evaluaciones heurísticas los autores han tenido que adaptar las heurísticas para el contexto de aplicaciones móviles.

Por otro lado las evaluaciones de usabilidad basadas en modelos de usabilidad existentes deben ser personalizadas para abordar el enfoque móvil. Los modelos más usados son: el modelo propuesto por Nielsen, con heurísticas tradicionales para computadoras de escritorio, y el modelo ISO 9241 que mide la eficiencia, efectividad y satisfacción de una aplicación. El modelo PACMAD es propuesto para evaluar la usabilidad y extiende los modelos existentes de usabilidad, como Nielsen o ISO pero para el contexto específico de aplicaciones móviles, PACMAD implementa el atributo carga cognitiva que se refiere al procesamiento cognitivo que el usuario necesita utilizar la aplicación. Existen también muchos modelos propuestos por algunos

investigadores, cada uno con diferentes aportes dentro del contexto móvil, pero no abarcan todos los factores de usabilidad, en la mayoría dejan de lado el contexto de uso de la aplicación.

En cuanto a los factores que determinan usabilidad: usuarios, tareas y contexto de uso se ha podido observar que no todos los autores consideran estos factores tanto al desarrollar como al evaluar la usabilidad de las aplicaciones móviles, quizás si estos factores fueran tomados en cuenta desde etapas tempranas del ciclo de vida de software, esta característica sería menos considerada en el proceso de evaluación. Por ejemplo si al desarrollar una aplicación móvil se consideraría más a sus usuarios puntuales, se podría comprender de mejor manera sus necesidades y capacidades, haciendo participes a los usuarios de una forma iterativa diseño-evaluación, se podrían capturar diferentes características que influyen para que la creación de un producto software eficiente, usable y con el que el usuario se sienta satisfecho. Al igual se requiere investigar más sobre la combinación de evaluaciones de usabilidad con de metodologías de ágiles de desarrollo, que ha demostrado su eficiencia y buenos resultados en los escasos estudios de este tipo realizados.

Finalmente en lo que respecta al diseño de interacción de las aplicaciones móviles y el efecto heterogéneo entre las plataformas existentes, se puede decir que en la literatura no se ha encontrado una guía de diseño de interacción independiente de la plataforma, problema que influye en la usabilidad de una aplicación multiplataforma, ya que las expectativas y forma de usar una aplicación por los usuarios depende en gran medida de la plataforma a la que este acostumbrado a interactuar. Si se crea una guía de interfaz de usuario multiplataforma esta característica podría ser menos estimada en las evaluaciones y causaría menos confusión al usuario.

Con los resultados y análisis de la presente Revisión sistemática de literatura, se procede a desarrollar las conclusiones, recomendaciones y posibles investigaciones futuras.

CAPÍTULO VI

CONCLUSIONES, RECOMENDACIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

6.1 Conclusiones

Existen distintas técnicas para medir la usabilidad de las aplicaciones móviles multiplataforma, entre ellas se encuentran las evaluaciones heurísticas, evaluaciones observacionales como de campo y laboratorio, evaluaciones de usabilidad subjetiva como cuestionarios y entrevistas, evaluaciones con el protocolo “Think aloud”, evaluaciones de recorrido cognitivo y estudios longitudinales. Se puede decir que los métodos más usados son: la evaluación heurística y los cuestionarios, ya que son considerados como técnicas más rápidas y de bajo costo. Además los diferentes autores de los estudios optan también por la combinación de distintos métodos para capturar los factores contextuales de los dispositivos móviles, que un solo método de evaluación no cubre, por lo que, se demuestra una falta de métodos integrales que envuelvan los factores del contexto móvil.

En las evaluaciones de usabilidad basadas en modelos o estándares, los que son más comúnmente utilizados son: el Modelo de Usabilidad propuesto por Nielsen que propone un conjunto de heurísticas como guías de interfaz de usuario y el Modelo ISO 9241-11 que mide un conjunto de atributos para considerar a un producto software usable, ambos modelos requieren ser adaptados para abarcar el contexto móvil, ya que fueron creados únicamente para evaluar aplicaciones tradicionales de escritorio. La combinación y personalización de los modelos existentes de usabilidad muestra la falta de modelos formalizados o que estos no se han definido de forma consistente en el proceso de evaluación de usabilidad de aplicaciones móviles.

Los atributos considerados para la evaluación de usabilidad de aplicaciones móviles dependen del modelo de usabilidad y en otros casos depende de los autores de los estudios ya que no siguen un modelo en específico y miden atributos que ellos consideran importantes. Entre los más estimados están: la visibilidad del sistema,

encuentro entre el sistema y el mundo real, control y libertad de usuario, consistencia y estándares, prevención de errores, reconocimiento en lugar de recuerdo, flexibilidad y eficiencia de uso, diseño estético y minimalista, ayuda a los usuarios a recuperarse de los errores, ayuda y documentación, compatibilidad entre diferentes plataformas, poca interacción humana/dispositivo, interacción física y ergonómica, legibilidad y diseño, efectividad, eficiencia, satisfacción, facilidad de aprendizaje, facilidad de recordar, errores y carga cognitiva. Las métricas usadas para medir estos atributos van desde medidas subjetivas como puntuaciones, escalas en cuestionarios y listas de verificación, hasta medidas objetivas por ejemplo para medir eficiencia se toma el tiempo que tarda un usuario en finalizar una tarea, estas métricas dependen en gran medida del tipo de atributo que se desea medir.

El tipo de aplicaciones móviles más evaluadas son aplicaciones respecto a la salud, esto se debe a que una aplicación va dirigida a cierto tipo de usuarios y en el caso de estas aplicaciones van enfocadas a personas con diferentes capacidades y enfermedades, por ello se debería estudiar más las limitaciones de los futuros usuarios a través de pruebas de usabilidad, ya que en la literatura se han observado varios problemas de usabilidad porque los desarrolladores de este tipo de aplicaciones no hacen partícipes a los usuarios finales en etapas de desarrollo de la aplicación. En lo que respecta a las actividades que se realizan en las pruebas de usabilidad, se puede observar que varían su nivel de complejidad y son cuidadosamente establecidas con el fin de que se evalúe la totalidad de cada función de la aplicación.

En el proceso de evaluación de usabilidad de aplicaciones móviles multiplataforma existen aciertos y desaciertos, entre los aciertos más relevantes se puede mencionar que: las heurísticas y modelos existentes sirven como un punto de partida para adaptar las características del contexto móvil, la evaluación heurística es de bajo costo y rápida, las evaluaciones de campo capturan factores del contexto ambiental que no se pueden simular en un laboratorio, la usabilidad es mayor si existe una iteración en todas las etapas del ciclo de vida del producto software, la combinación de métodos permite capturar una amplia y detallada gama de problemas de usabilidad, los cuestionarios son útiles para comprender las percepciones y deseos

de los usuarios, la incorporación de un método de evaluación con un proceso de desarrollo ágil podría producir aplicaciones más fáciles de usar. Entre los inconvenientes o desaciertos están que: sin modificar las evaluaciones heurísticas y modelos podrían no ser efectivas, los modelos de evaluación son limitados ya que muchas veces los desarrolladores no prestan atención al área de interacción humano-dispositivo, la falta de procesos iterativos desarrollo/evaluación, las entrevistas, cuestionarios realizados por expertos y pruebas de laboratorio podrían ser sesgadas positivamente a la dirección de los investigadores, las distracciones del ambiente tienden a reducir recursos cognitivos de la persona, aspecto que no siempre se toma en cuenta en las evaluaciones, la evaluación con usuarios es costosa y consume tiempo, el tipo de teléfono y plataforma tiene un impacto en la apariencia, usabilidad y podría afectar las experiencias de los usuarios.

En la etapa de evaluación de usabilidad de aplicaciones móviles no todos los autores toman en cuenta los tres factores importantes de usabilidad: usuarios, tareas y contexto de uso, por ejemplo en la evaluación heurística que es realizada por expertos dejan de lado los usuarios y el contexto de uso, en los cuestionarios en cambio solo toman en cuenta usuarios y tareas, casi no consideran el factor contexto de uso, el mismo que es muy importante ya que las distracciones tienden influir en el comportamiento y en los recursos cognitivos del usuario.

En el proceso de desarrollo del diseño de interacción de las aplicaciones móviles, los factores que determinan usabilidad: usuarios, tareas y contexto de uso, no siempre son considerados, lo que produce problemas de usabilidad en etapas posteriores del ciclo de vida del producto software. El factor usuarios es muy importante ya que hay que tomar en cuenta sus necesidades y capacidades al crear una aplicación, este factor a veces no es muy considerado porque los investigadores prefieren realizar evaluaciones con expertos. El factor tareas también es significativo ya que la aplicación debe cumplir con las tareas requeridas por el usuario, este factor en cambio es el más considerado en los estudios ya que generalmente los desarrolladores se enfocan más en que el sistema cumpla con los requisitos funcionales y no funcionales y dejan un poco de lado la parte del diseño de interacción. Finalmente el factor

contexto de uso últimamente está siendo más considerado porque se ha descubierto la influencia que este factor causa al usar una aplicación móvil, sin embargo aún requiere un fuerte estudio. En la mayoría de estudios primarios 26 no especifican como se creó el diseño de interacción de las aplicaciones, simplemente se limitan a evaluarlas, en 11 de los estudios si toman en cuenta los tres factores al crear sus aplicaciones lo cual es adecuado y produce buenos resultados en las evaluaciones de usabilidad, se puede notar que si se consideran estos factores desde etapas tempranas del ciclo de vida de un producto software estos podrían ser menos considerados en las evaluaciones y por último en 5 estudios se consideran usuarios y tareas, dejando de lado el contexto de uso

El diseño de interacción de las aplicaciones móviles evidentemente difiere entre las plataformas existentes, ya que al momento de desarrollar una aplicación móvil, cada plataforma sigue sus directrices y filosofía de interfaz de usuario. Esto ocasiona confusión al usuario al utilizar sistema y podría provocar que disminuya la usabilidad de la aplicación porque en estudios previos se ha demostrado que los requisitos del usuario, expectativas y comportamientos son algo diferente según la plataforma móvil. Por ejemplo los usuarios de determinada plataforma móvil al usar aplicaciones de la misma plataforma consideran al sistema más fácil de usar y requieren menos peticiones de ayuda al realizar las tareas en la evaluación. Al contrario un usuario que use una aplicación en una plataforma diferente a la que este acostumbrado a interactuar, se le dificulta más usar el sistema. Entonces, las percepciones de los usuarios cambian dependiendo de la plataforma móvil, aspecto que debería considerarse por los desarrolladores, para que los sistemas guarden más coherencia en el diseño de interfaz, a través de la aplicación de guías de interfaz de usuario multiplataforma, de manera que este aspecto no influya tanto en las percepciones de los usuarios y por ende en la usabilidad de una aplicación.

Finalmente, con la combinación un método de desarrollo con un método de evaluación de usabilidad integral que capture los aspectos del contexto móvil, a través de un proceso iterativo de desarrollo-evaluación, se puede descubrir una amplia y detallada gama de problemas de usabilidad que pueden ser resueltos a tiempo, para la

producción de una aplicación fácil de utilizar. En las aplicaciones móviles multiplataforma existe un factor que produce un impacto significativo en su usabilidad, este factor es la plataforma móvil. Esta característica puede afectar las experiencias y comportamientos de los usuarios, por lo que al crear el diseño de interacción de la aplicación multiplataforma es importante la aplicación de guías de interfaz de usuario independientes de la plataforma móvil, para que el efecto heterogéneo de una misma aplicación en distintas plataformas se reduzca y así que este factor no influya negativamente en la facilidad de aprendizaje y usabilidad de una aplicación. Finalmente, la adopción de mecanismos automatizados para la evaluación de usabilidad requiere ser más investigada, para la combinación de más de un método y también para ofrecer un enfoque multiplataforma, porque los métodos existentes son solo para plataformas y aplicaciones muy específicas.

6.2 Recomendaciones

Es recomendable y muy importante tomar en cuenta el contexto ambiental dentro de la usabilidad móvil, ya que las distracciones tienden a reducir visibilidad del usuario, memoria y recursos cognitivos, factor que en ciertos métodos de evaluación no es tomado en cuenta. En una gran cantidad de estudios primarios de evaluación de usabilidad solo considera los factores: tareas y usuarios.

Evaluar las aplicaciones dentro del proceso de desarrollo de software tomando en cuenta todos los principales factores de usabilidad: tareas, usuarios, y contexto de uso.

La creación y aplicación de guías o directrices de interfaz de usuario independientes de la plataforma, que pueden ser utilizadas para el desarrollo y evaluación de usabilidad de las aplicaciones en las distintas plataformas, para que no difieran tan notablemente como en la actualidad, ya que cada dispositivo es necesario que cumpla con principios de usabilidad independientes de la plataforma, lo que en la realidad no se cumple porque en los últimos tiempos se ha observado un efecto heterogéneo entre los diseños de interfaz de las aplicaciones.

Estudiar a fondo los factores de usabilidad en el proceso de desarrollo del producto software, así esto podría ser menos considerado por los evaluadores en la etapa de evaluación de usabilidad.

La adopción de mecanismos automatizados para la evaluación de usabilidad requiere ser más investigada, para la combinación de más de un método.

6.3 Futuras Líneas de Investigación

En el proceso de realización de esta tesis, se cumplió con los objetivos planteados de la SLR y los resultados obtenidos han contribuido a exponer nuevas líneas de investigación, las cuales son las siguientes: (1) considerar un framework automatizado para evaluar la usabilidad de aplicaciones móviles, tomando en cuenta todos los factores de usabilidad: usuarios, tareas y contexto de uso. Pensar también en la Línea de Productos de Software SPL “Software product line” para el desarrollo de apps móviles para distintos usos. Una SPL sugiere una nueva manera de desarrollo de software, en donde ya no se desarrolla un software para cada producto diferente, sino una línea de productos. De esta manera, los productos que comparten algunas de sus características pueden ser desarrolladas a partir de artilugios comunes sin la necesidad de empezar a crearlos desde cero (Miranda, Casas, & Marcos, 2015) (2) crear un conjunto de guías y lineamientos de interfaz de usuario común entre aplicaciones multiplataforma, (3) estudiar el contexto de uso dentro del proceso de desarrollo y evaluación de usabilidad de aplicaciones móviles, (4) realizar evaluaciones de usabilidad de aplicaciones móviles ubicuas, web y multimodales, (5) la adopción de procesos de evaluación de usabilidad iterativos en la etapa de desarrollo de software, integrando las metodologías de desarrollo con las de evaluación de usabilidad.

Referencias

- Abdul, N., Hussain, A., & Zainol, A. (2014). *A Usability Testing On jFakih Learning Games for Hearing Impairment Children*. Universiti Utara Malaysia, Malaysia.
- Aizawa, K., Maeda, K., Ogawa, M., Sato, Y., Kasamatsu, M., Waki, K., & Takimoto, H. (2014). *Comparative Study of the Routine Daily Usability of FoodLog: A Smartphone based Food Recording Tool Assisted by Image Retrieval*. University of Tokyo, Tokyo, Japan.
- Akma Ahmad, N., Zainal, A., Abdul Razak, F. H., Wan Adnan, W. A., & Osman, S. (2015). *USER EXPERIENCE EVALUATION OF MOBILE SPIRITUAL APPLICATIONS FOR OLDER PEOPLE: AN INTERVIEW AND OBSERVATION STUDY*. Universiti Selangor, MARA University of Technology, Malaysia.
- Angulo Cevallos, E. (2014). *Case Study on Mobile Applications UX: Effect of the Usage of a Cross-Platform Development Framework*. Universidad Politécnica de Madrid, Madrid.
- Arnhold, M., Quade, M., & Kirch, W. (2014). *Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the*

- special requirements of diabetes patients age 50 years or older*. Technische Universität Dresden, Dresden, Germany.
- Assal, H., Hurtado, S., Imran, A., & Chiasson, S. (2015). *What's the deal with privacy apps? A comprehensive exploration of user perception and usability*. Carleton University, Linz, Austria.
- Baharuddin, R., Singh, D., & Razali, R. (2013). *Usability Dimensions for Mobile Applications-A Review*. Universiti Kebangsaan Malaysia, Selangor, Malaysia.
- Barea, A., Ferré, X., & Villarroel, L. (2013). *Android vs. iOS Interaction Design Study for a Student Multiplatform App*. Universidad Politécnica de Madrid Campus de Montegancedo, Boadilla del Monte (Madrid), Spain.
- Beul-Leusmann, S., Samsel, C., Wiederhold, M., Krempels, K.-H., Jakobs, E.-M., & Ziefle, M. (2014). *Usability Evaluation of Mobile Passenger Information Systems*. RWTH Aachen University, Germany.
- Biederer, M., Arguel, A., Liu, J., & Lau, A. (2014). *From Web-Based to Mobile: Experiences of Developing a Personally Controlled Health Management System*. University of Applied Sciences Weihenstephan, Triesdorf, Freising, Germany.
- Cañedo, R. (2000). *Ciencia y tecnología en la sociedad. Perspectiva histórico-conceptual*. Centro Nacional de Información de Ciencias Médicas, La Habana, Cuba.
- Cazau, P. (2006). *INTRODUCCIÓN A LA INVESTIGACIÓN EN CIENCIAS SOCIALES*. Buenos Aires.
- Charland, A., & Leroux, B. (2011). *Mobile application development: web vs. native*. Institute of Translational Health Sciences, New York.
- Choi, M., Suk Lee, H., & Ho Park, J. (2015). *Usability of Academic Electronic Medical Record Application for Nursing Students' Clinical Practicum*. Nursing Policy Research Institute, College of Nursing, Yonsei University, Korea.
- Chung, C. (2008, Abril 19). <http://carloschungr.blogspot.com/>. Obtenido de <http://carloschungr.blogspot.com/2008/04/enfoques-cuantitativo-deductivo-y.html>
- Coello, P. A., Ferreira González, I., & Urrútia, G. (2011, Agosto). Obtenido de <http://www.revespcardiol.org/es/revisiones-sistematicas-metaanalisis-bases-conceptuales/articulo/90024424/>
- Cruz Zapata, B., Fernández, J. L., Idri, A., & Toval, A. (2015). *Empirical Studies on Usability of mHealth Apps: A Systematic Literature Review*. University of Murcia, Murcia, Spain.
- Cruz Zapata, B., Hernandez Niñirola, A., Idri, A., Fernandez-Aleman, J., & Toval, A. (2014). *Mobile PHRs Compliance with Android and iOS Usability Guidelines*. University of Murcia, Murcia, Spain.
- de Lima Salgado, A., & Pimenta Freire, A. (2014). *Heuristic Evaluation of Mobile Usability: A Mapping Study*. Federal University of Lavras, Lavras, MG, Brazil.
- Di Giovanni, P., Romano, M., Sebillo, M., Tortora, G., Vitiello, G., Ginige, T., . . . Ginige, A. (2012). *User centered scenario based approach for developing mobile interfaces for Social Life Networks*. University of Zurich, Zurich.

- Dieste, O., Grimán, A., Juristo, N., & López, M. (2008). *Revisiones Sistemáticas: Recomendaciones para un Proceso*. Universidad Politécnica de Madrid, Madrid.
- Donaldson, S., & Siegel, S. (2000). *Successful Software Development 2nd Edition*. Prentice Hall PTR.
- Enriquez, J. G., & Casas, S. I. (2013). *USABILIDAD EN APLICACIONES MÓVILES*. Santa Cruz, Argentina.
- Faliagka, E., Rigou, M., & Sirmakessis, S. (2014). *A usability study of iPhone built-in applications*. Institution of Western Greece & University of Patras,.
- Ferré, X. (2000). *Principios Básicos de Usabilidad para Ingenieros Software*. Universidad Politécnica de Madrid, Madrid.
- Ferré, X. (2005). *Marco de Integración de la Usabilidad en el Proceso de Desarrollo de Software*. Madrid.
- Ferré, X., Juristo, N., Windl, H., & Constantine, L. (2001). Usability Basics for Software Developers. *Journal IEEE Software*, 18(1), 22-29.
- Fonseca, R. (2014). *Conceptualización e Infraestructura para la Investigación Experimental en Ingeniería del Software*. Universidad Politécnica de Madrid, Madrid.
- Foo, E., McDonald, R., Savage, E., Floyd, R., Butler, A., Rumball-Smith, A., & Connor, S. (2015). *Mobile task management tool that improves workflow of an acute general surgical service*. University of Canterbury, University of Otago, New Zealand.
- Gresse von Wangenheim, C., Witt, T., Borgatto, A., Nunes, J., Lacerda, T., Krone, C., & Souza, L. (2016). *An Usability Score for Mobile Phone Applications based on Heuristics*. Federal University of Santa Catarina, Florianópolis – SC – Brazil.
- Hsiu Ching, L. H. (2015). *Developing Mobile Application Design of Virtual Pets for Caring for the Elderly*. National Yunlin University of Science and Technology, Douliu, Taiwan.
- Humayoun, S., Ehrhart, S., & Ebert, A. (2013). *Developing mobile apps using cross-platform frameworks: a case study*. University of Kaiserslautern.
- Hussain, A., Abdul, N., & Zainol, A. (2014). *A Usability Testing On jFakih Learning Games for Hearing Impairment Children*. Universiti Utara Malaysia, Malaysia.
- Hussain, A., Hashim, N. L., & Nordin, N. (2014). *mGQM: Evaluation Metric for Mobile and Human Interaction*. Universiti Utara Malaysia, Kedah-Malaysia.
- Hussain, A., Saleh, A., Taher, A., Ahmed, I., & Lammasha, M. (2015). *USABILITY EVALUATION METHOD FOR MOBILE LEARNING APPLICATION USING AGILE: A SYSTEMATIC REVIEW*. Universiti Utara Malaysia, Sintok. Kedah, Malaysia.
- IBM. (2012). *El desarrollo de aplicaciones móviles nativas, Web o híbridas*.
Obtenido de
ftp://ftp.software.ibm.com/la/documents/gb/commons/27754_IBM_WP_Native_Web_or_hybrid_2846853.pdf
- Ibrahim, H., Hussain, A., & Binti, N. (2014). *Evaluating Mobile Banking Application: Usability Dimensions and Measurements*. Putrajaya, Malaysia.

- Joyce, G., Lilley, M., Barker, T., & Jefferies, A. (2015). *Smartphone Application Usability Evaluation: The Applicability of Traditional Heuristics*. School of Computer Science, University of Hertfordshire College Lan, UK.
- Kalz, M., Lenssen, N., Felzen, M., Rossaint, R., Tabuenca, B., Specht, M., & Skorning, M. (2014). *Smartphone Apps for Cardiopulmonary Resuscitation Training and Real Incident Support: A Mixed-Methods Evaluation Study*. University of the Netherlands, Heerlen, Netherlands.
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Keele University, UK.
- Kitchenham, B., Wohlin, C., Runeson, P., Höst, M., Ohlsson, M., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering*. New York: Springer-Verlag.
- Kluth, W., Krempels, K. H., & Samsel, C. (2014). *Automated Usability Testing for Mobile Applications*. RWTH Aachen University, Aachen, Germany.
- Lira, W., de Souza, C., Ferreira, R., & Carvalho, S. (2014). *Experimenting on the Cognitive Walkthrough with Users*. Vale Institute of Technology, Belém, Brazil.
- López, A., Cabrera, C., & Luz, V. (2008). *INTRODUCCIÓN A LA CALIDAD DE SOFTWARE*. Universidad Tecnológica de Pereira, Pereira.
- López-Gil, J. M., Urretavizcaya, M., Fernández-Castro, I., & Losada, B. (2014). *Integrating Field Studies in Agile Development to Evaluate Usability on Context Dependant Mobile Applications*. Universidad del País Vasco, Donostia-San Sebastián.
- Luhur, H. S., & Widjaja, N. D. (2014). *Location-based social networking media for restaurant promotion and food review using mobile application*. Bina Nusantara University, 0 Jakarta, Indonesia.
- Lyles, C., Sarkar, U., & Osborn, C. (2014). *Getting a Technology-based Diabetes Intervention Ready for Primetime: A Review of Usability Testing Studies*. UCSF School of Medicine, San Francisco.
- Mandia Grossi, L., Torres Pisa, I., & Marin, H. (2014). *Oncoaudit: development and evaluation of an application for nurse auditors*. Escola Paulista de Enfermagem, Universidade Federal de São Paulo, São Paulo, Brazil.
- Maren Witt, S. (2014). *A set of quantitative user experience metrics for multi-modal dialog systems*. Speech Communication, Erlangen, Germany.
- Miranda, M., Casas, S., & Marcos, C. (2015). *Análisis de desarrollo de software orientado a Feature - Línea de producto de Software para Aplicaciones de TVDI*. Universidad Nacional de la Patagonia Austral, Río Gallegos.
- Mirkovic, J., Kaufman, D., & Ruland, C. (2014). *Supporting Cancer Patients in Illness Management: Usability Evaluation of a Mobile App*. Oslo University Hospital, Arizona State University, , Oslo, Norway, Scottsdale, AZ, United States.
- Nawi, A., Isa Hamzah, M., Chy Ren, C., & Tamuri, H. (2015). *Adoption of Mobile Technology for Teaching Preparation in Improving Teaching Quality of Teachers*. Universiti Brunei Darussalam, Universiti Kebangsaan Malaysia, International Graduate Studies College, elangor International Islamic University, Brunei, Malaysia.

- Olsina, L., Santos, L., & Lew, P. (2014). *Evaluating Mobileapp Usability: A Holistic Quality Approach*. Universidad Nacional de La Pampa, Beihang University, La Pampa, Argentina, China.
- O'Malley, G., Dowdall, G., Burls, A., Perry, I., & Curran, N. (2014). *Exploring the Usability of a Mobile App for Adolescent Obesity Management*. Temple Street Children's University Hospital, Dublin, Ireland.
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human Computer Interaction*. UK: Addison-Wesley .
- Reynoldson, C., Stones, C., Allsop, M., Gardner, P., Bennett, M., Closs, J., . . . Knapp, P. (2014). *Assessing the Quality and Usability of Smartphone Apps for Pain Self-Management*. University of Leeds, Leeds.
- Rodriguez, J., Sierra, E., & Jaramillo, L. (2015). Model for measuring usability of survey mobile apps, by analysis of usability evaluation methods and attributes. *10th Iberian Conference on Information Systems and Technologies (CISTI)*, 1-6.
- Saleh, A., Binti Isamil, R., & Binti Fabil, N. (2015). *EXTENSION OF PACMAD MODEL FOR USABILITY EVALUATION METRICS USING GOAL QUESTION METRICS*. Universiti Sains Isalm Malaysia, Malaysia.
- Scalone, F. (2006). *ESTUDIO COMPARATIVO DE LOS MODELOS Y ESTANDARES DE CALIDAD DEL SOFTWARE*. UNIVERSIDAD TECNOLOGICA NACIONAL, Buenos Aires.
- Scholtz, J. (2004). Usability Evaluation. *National Institute of Standards and Technology*.
- Skorin-Kapov, L., Dobrijevic, O., & Piplica, D. (2014). *Towards Evaluating the Quality of Experience of Remote Patient Monitoring Services: A Study Considering Usability Aspects*. University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia.
- Tomé Klock, A. C., & Gasparini, I. (2015). *A Usability Evaluation of Fitness-Tracking Apps for Initial Users*. Santa Catarina State University, Santa Catarina, Brazil.
- Tsiaousis, A., & Giaglis, G. (2014). *Mobile websites: usability evaluation and design*. Athens University of Economics and Business, Athens, Greece.
- Van Der Weegen, S., Verwey, R., Tange, H., Spreeuwenberg, M., & de Witte, L. (2014). *Usability testing of a monitoring and feedback tool to stimulate physical activity*. CAPHRI school for Public health and Primary care, Faculty of health Medicine and lifsciences, Maastricht University, Netherlands.
- Vélez, O., Boakye, P., Kanter, A., & Bakken, S. (2014). *A USABILITY STUDY OF A MOBILE HEALTH APPLICATION FOR RURAL GHANAIAN MIDWIVES*. Columbia University Medical Center, New York.
- Wasserman, T. (2010). *Software engineering issues for mobile application development*. Carnegie Mellon University, New York.
- Wein, L. (2014). *Visual Recognition in Museum Guide Apps: Do Visitors Want It?* Amsterdam University College, Amsterdam, Netherlands.
- Wohlin, C., Runeson, P., Höst, M., & Ohlsson, M. (2012). *Experimentation in Software Engineering*.
- Xu, J., Ding, X., Huang, K., & Chen, G. (2014). *A Pilot Study of an Inspection Framework for Automated Usability Guideline Reviews of Mobile Health*

- Applications*. University of Massachusetts Lowell, Nanjing University, Nanjing, China.
- Yáñez Gómez, R., Cascado Caballero, D., & Sevillano, J.-L. (2014). *Heuristic Evaluation on Mobile Interfaces: A New Checklist*. Universidad de Sevilla,, Seville, Spain.
- Zhang, D., & Adipat, B. (2005). Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. *International Journal of Human-Computer Interaction*, 18(3), 293-308.