

ANDRÉS FERNANDO SOLANO ALEGRÍA

**METODOLOGÍA PARA LA EVALUACIÓN
COLABORATIVA DE LA USABILIDAD DE SISTEMAS
SOFTWARE INTERACTIVOS**

**Tesis presentada a la Facultad de Ingeniería
Electrónica y Telecomunicaciones de la
Universidad del Cauca para la obtención del
Título de**

**Doctor en:
Ciencias de la Electrónica**

**Director:
PhD. César Alberto Collazos Ordóñez
Universidad del Cauca, Colombia**

**Co-Director:
PhD. Cristian Alexandru Rusu
Pontificia Universidad Católica de Valparaíso, Chile**

<p>Comisión evaluadora: PhD. William Giraldo, Universidad del Quindío, Colombia PhD. Natalia Padilla, Universidad de Granada, España PhD. Alejandro Fernández, Universidad Nacional de La Plata, Argentina PhD. Jaime Muñoz, Universidad Autónoma de Aguascalientes, México</p>
--

**Popayán
2015**

A toda mi familia, por su apoyo incondicional. En especial le dedico este trabajo a mis padres, quienes siempre me han brindado todo su amor, cariño, buen ejemplo y me han apoyado en todas las metas que he querido alcanzar. Cualquier meta fijada, con sacrificio y dedicación constante es posible cumplirla, sin importar las condiciones que la vida nos ha impuesto.

Agradecimientos

A mi familia, por su apoyo incondicional. En especial agradezco a mis padres Liliana y Elvio, por su ejemplo, por su amor, por su tiempo, por su empuje y fortaleza, por ser mi mayor inspiración y estímulo.

A César Collazos, mi director y amigo, quien ha sido el gestor primario de todo este proceso, quien me ha guiado y apoyado incondicionalmente para alcanzar esta gran meta.

A Cristian Rusu, quien me ha colaborado desinteresadamente en la realización de mi trabajo. Gracias por su apoyo constante aún en la distancia física que separa a Colombia de Chile. Gracias por la confianza.

A Toni Granollers, le quedo infinitamente agradecido por su amistad, apoyo y asesorías en el momento preciso.

A mi amiga Yenny Méndez, por acompañar mi esfuerzo y por sus voces de aliento.

A mis colegas del grupo IDIS, GIT, UseCV, GRIHO y SINFOCI, que me alentaron constantemente. Gracias por haber compartido momentos de trabajo y compañerismo.

A la Universidad del Cauca, al Comité de Doctorado y a COLCIENCIAS, por el soporte económico brindado que hicieron posible en buena parte este estudio.

Resumen

La usabilidad es una característica de calidad fundamental para el éxito de un sistema interactivo, la cual corresponde a un concepto que incluye una serie de métricas y métodos con el objetivo de obtener sistemas fáciles de usar y aprender. Los métodos de evaluación de usabilidad son bastante diversos y su realización depende de variables tales como: costos, disponibilidad de tiempo, recursos humanos, etc. Varios métodos de evaluación pueden ser ejecutados sobre sistemas software interactivos, pero la incógnita está en el nivel de la información que brinda cada uno de ellos y/o la combinación de unos con otros. Esta investigación está enfocada en el estudio de un conjunto de métodos de evaluación de usabilidad en diferentes sistemas software interactivos, proponiendo una metodología de evaluación constituida por varias combinaciones de métodos de evaluación. Así, la metodología de evaluación podrá brindar información de la usabilidad de sistemas software interactivos de forma más completa que aquellos métodos de evaluación ejecutados de forma independiente. Adicionalmente, los métodos de evaluación que constituyen la metodología propuesta, han sido especificados de forma colaborativa siguiendo los principios definidos en la Ingeniería de Colaboración. De esta forma, se espera contribuir a la rigurosidad y consistencia en el análisis de resultados, así como también generar resultados más adecuados respecto a las ventajas del trabajo colaborativo.

Abstract

Usability is a fundamental quality characteristic for the success of an interactive system. It is a concept that includes a set of metrics and methods in order for users to obtain easy-to-use and easy-to-learn systems. Usability evaluation methods are quite diverse; its realization depends on variables such as costs, time availability, human resources, among others. A large number of usability evaluation methods can be employed to assess interactive software systems, but questions arise when deciding which method and/or combination of methods gives more (relevant) information. This research is aimed at analyzing a set of usability evaluation methods in different interactive software systems, proposing an assessment methodology made up of various combinations of evaluation methods. Thus, this assessment methodology will provide more complete and comprehensive information about the usability of interactive software systems than those evaluation methods conducted independently. Moreover, the evaluation methods which made up the proposed methodology have been specified collaboratively following the principles defined by the Collaboration Engineering. This has been done in order to contribute to the stringency and consistency of the analysis of results, as well as to generate more appropriate results, considering the advantages of the collaborative work.

Tabla de contenido

INTRODUCCIÓN	1
1.1 PLANTEAMIENTO DEL PROBLEMA.....	2
1.2 OBJETIVOS.....	4
1.2.1 <i>Objetivo General</i>	4
1.2.2 <i>Objetivos Específicos</i>	4
1.3 HIPÓTESIS.....	4
1.4 ORGANIZACIÓN DEL DOCUMENTO.....	5
BASE CONCEPTUAL	7
2.1 COMPONENTE USABILIDAD.....	7
2.1.1 <i>Sistemas interactivos</i>	7
2.1.2 <i>Experiencia de usuario</i>	8
2.1.2.1 Usabilidad.....	8
2.1.2.1.1 Evaluación de usabilidad.....	9
2.1.2.1.2 Métodos de evaluación de usabilidad – MEU.....	10
2.2 COMPONENTE COLABORATIVO.....	13
2.2.1 <i>Trabajo colaborativo</i>	13
2.2.1.1 Beneficios del trabajo colaborativo.....	13
2.2.1.2 Desventajas del trabajo colaborativo.....	13
2.2.1.3 Principios básicos del trabajo colaborativo.....	14
2.2.2 <i>Ingeniería de Colaboración</i>	14
2.2.2.1 Patrones de colaboración.....	15
2.2.2.2 Thinklets.....	16
2.2.3 <i>Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software</i>	16
ESPECIFICACIÓN COLABORATIVA DE LOS MEU	19
3.1 ¿POR QUÉ INTEGRAR TRABAJO COLABORATIVO AL PROCESO DE EVALUACIÓN DE USABILIDAD?.....	20
3.2 SELECCIÓN DE LOS MEU OBJETO DE ESTUDIO.....	21
3.2.1 <i>Selección de los métodos de inspección</i>	22
3.2.2 <i>Selección de los métodos de prueba</i>	23
3.3 PROPUESTA DE NOTACIÓN PARA EL MODELADO DE PROCESOS COLABORATIVOS EXTENDIENDO LA NOTACIÓN HAMSTERS.....	24
3.3.1 <i>Análisis de notaciones existentes</i>	25
3.3.2 HAMSTERS	27
3.3.3 <i>Propuesta: extensión de la notación HAMSTERS para utilizar en el MFP</i>	27
3.3.4 <i>Evaluación preliminar de la notación propuesta</i>	31
3.3.4.1 Diseño del caso de estudio.....	31
3.3.4.2 Resultados obtenidos.....	31
3.3.4.3 Análisis de resultados.....	35
3.3.5 <i>Beneficios de la notación extendida: MFP+HAMSTERS</i>	36

3.4 ESPECIFICACIÓN COLABORATIVA DE LOS MEU OBJETO DE ESTUDIO	36
3.4.1 <i>Actividades recurrentes destacadas</i>	36
3.4.2 <i>Especificación colaborativa del método: evaluación heurística</i>	37
3.4.2.1 Fase 1 – Diagnóstico de la técnica.....	37
3.4.2.2 Fase 2 – Descomposición de la técnica.....	38
3.4.2.3 Fase 3 – Establecer tareas colaborativas.....	38
3.4.2.4 Fase 4 – Relación con proceso colaborativo.....	42
3.4.2.5 Fase 5 – Documentación de la técnica.....	44
3.4.1.5.1 Documentación del diseño - etapa de planeación	44
3.4.1.5.2 Documentación del diseño - etapa de ejecución.....	49
3.4.1.5.3 Documentación del diseño - etapa de análisis de resultados.....	50
3.4.2.6 Fase 6 – Validación	55
3.4.1.6.1 Simulación.....	55
3.4.1.6.2 Prueba piloto	55
EJECUCIÓN DE LOS MEU COLABORATIVOS.....	57
4.1 DEFINICIÓN DE MÉTRICAS.....	58
4.1.1 <i>Selección de métricas</i>	59
4.1.2 <i>Normalización de las métricas</i>	60
4.2 SELECCIÓN DE LOS SISTEMAS SOFTWARE INTERACTIVOS	61
4.3 EJECUCIÓN DE LOS MEU COLABORATIVOS Y ANÁLISIS DE RESULTADOS	62
4.3.1 <i>Área: Televisión Digital Interactiva</i>	63
4.3.1.1 Aplicaciones de TDi objeto de estudio.....	64
4.3.1.2 Evaluaciones realizadas.....	66
4.3.1.2.1 Condiciones de evaluación	66
4.3.1.2.2 Métodos de inspección.....	67
4.3.1.2.3 Métodos de prueba.....	68
4.3.1.2.4 Resumen ejecución de los MEU.....	69
4.3.1.3 Análisis de resultados.....	69
4.3.1.3.1 Característica: detección de problemas de usabilidad.....	69
4.3.1.3.2 Característica: tiempo	72
4.3.2 <i>Área: Web Transaccional</i>	74
4.3.2.1 Sitio web transaccional objeto de estudio.....	75
4.3.2.2 Evaluaciones realizadas.....	75
4.3.3 <i>Área: Aplicaciones móviles</i>	75
4.3.3.1 Aplicación móvil objeto de estudio.....	76
4.3.3.2 Evaluaciones realizadas.....	77
4.4 <i>Análisis comparativo</i>	77
4.4.1 Detección de problemas de usabilidad.....	77
4.4.2 Tiempo.....	79
4.4.3 Consideraciones adicionales.....	84
4.4.4.3.1 Ejecución de los MEU.....	84
4.4.4.3.2 Herramientas software de soporte.....	85
METODOLOGÍA PARA LA EVALUACIÓN COLABORATIVA DE LA USABILIDAD DE SISTEMAS SOFTWARE INTERACTIVOS.....	87
5.1 CARACTERÍSTICAS SOBRESALIENTES DE LOS MEU.....	87
5.2 METODOLOGÍA PARA LA EVALUACIÓN COLABORATIVA DE LA USABILIDAD	88
5.2.1 <i>Discusión</i>	88
5.2.2 <i>Metodología propuesta</i>	90
5.2.2.1 Evaluación global: alta detección de problemas.....	91
5.2.2.2 Evaluación específica: reducción de tiempo	92
5.2.2.3 Evaluación enfocada a tareas específicas: sin restricciones de tiempo	93
5.2.3 <i>Aplicabilidad de la metodología</i>	94

EVALUACIÓN DE LA METODOLOGÍA	97
6.1 ÁREA: REDES SOCIALES	98
6.1.1 <i>Sistema interactivo objeto de estudio.....</i>	98
6.2 ENCUESTA A EXPERTOS	98
6.3 EJECUCIÓN DE LOS MEU SUGERIDOS POR EXPERTOS	99
6.3.1 <i>Evaluaciones realizadas.....</i>	100
6.4 EJECUCIÓN DE LA COMBINACIÓN: <i>EVALUACIÓN ESPECÍFICA.....</i>	100
6.4.1 <i>Evaluaciones realizadas.....</i>	100
6.5 RESUMEN EJECUCIÓN DE LOS MEU	100
6.6 ANÁLISIS COMPARATIVO DE RESULTADOS	101
6.6.1 <i>Característica: detección de problemas de usabilidad.....</i>	101
6.6.2 <i>Característica: tiempo.....</i>	103
6.7 CONCLUSIONES DE LA METODOLOGÍA PROPUESTA.....	105
EXPERIENCIAS Y RECOMENDACIONES.....	109
CONCLUSIONES Y TRABAJO FUTURO	115
8.1 CONCLUSIONES	115
8.2 TRABAJO FUTURO.....	118
8.3 PUBLICACIONES.....	120
REFERENCIAS BIBLIOGRÁFICAS.....	123

Lista de tablas

TABLA 1. PATRONES DE COLABORACIÓN.	15
TABLA 2. COMPARACIÓN DEL SOPORTE QUE CADA NOTACIÓN PUEDE OFRECER A LOS REQUISITOS NECESARIOS PARA REPRESENTAR EL MFP.	26
TABLA 3. SÍMBOLOS DE LOS PARTICIPANTES.	28
TABLA 4. REPRESENTACIÓN GRÁFICA DE ACTIVIDADES COLABORATIVAS.	30
TABLA 5. RESULTADOS OBTENIDOS PARA EL MFP TRADICIONAL.	33
TABLA 6. RESULTADOS OBTENIDOS PARA EL MFP QUE INCLUYE ELEMENTOS DE HAMSTERS.	34
TABLA 7. ACTIVIDADES GENERALES DE LA EVALUACIÓN HEURÍSTICA.	38
TABLA 8. ACTIVIDADES DE LA ETAPA DE PLANEACIÓN.	39
TABLA 9. ACTIVIDADES DE LA ETAPA DE EJECUCIÓN.	40
TABLA 10. ACTIVIDADES DE LA ETAPA DE ANÁLISIS DE RESULTADOS.	40
TABLA 11. ASOCIACIÓN DE PATRONES A LA ACTIVIDAD Nº 11: CREACIÓN DE UNA LISTA INTEGRADA DE PROBLEMAS.	41
TABLA 12. ASOCIACIÓN DE PATRONES A LA ACTIVIDAD Nº 15: ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS.	41
TABLA 13. ASOCIACIÓN DE PATRONES A LA ACTIVIDAD Nº 16: REALIZAR RECOMENDACIONES PARA SOLUCIONAR LOS PROBLEMAS DE USABILIDAD.	42
TABLA 14. ASOCIACIÓN DE PATRONES A LA ACTIVIDAD Nº 17: IDENTIFICAR ELEMENTOS POSITIVOS DEL SISTEMA.	42
TABLA 15. RELACIÓN DE THINKLETS A LA ACTIVIDAD Nº 11: CREACIÓN DE UNA LISTA INTEGRADA DE PROBLEMAS.	43
TABLA 16. RELACIÓN DE THINKLETS A LA ACTIVIDAD Nº 15: ANÁLISIS E INTERPRETACIÓN DE LOS RESULTADOS.	43
TABLA 17. RELACIÓN DE THINKLETS A LA ACTIVIDAD Nº 16: REALIZAR RECOMENDACIONES PARA SOLUCIONAR LOS PROBLEMAS DE USABILIDAD.	43
TABLA 18. RELACIÓN DE THINKLETS A LA ACTIVIDAD Nº 17: IDENTIFICAR ELEMENTOS POSITIVOS DEL SISTEMA.	44
TABLA 19. AGENDA DETALLADA CON LAS ACTIVIDADES QUE CONFORMAN LA ETAPA DE PLANEACIÓN.	47
TABLA 20. AGENDA DETALLADA CON LAS ACTIVIDADES QUE CONFORMAN LA ETAPA DE EJECUCIÓN.	49
TABLA 21. AGENDA DETALLADA CON LAS ACTIVIDADES QUE CONFORMAN LA ETAPA DE ANÁLISIS DE RESULTADOS.	53
TABLA 22. DESCRIPCIÓN DE LAS MÉTRICAS A CONSIDERAR EN EL ANÁLISIS DE RESULTADOS DE LOS MEU.	60
TABLA 23. EJEMPLO DE ESQUEMA DE TABLA DE NORMALIZACIÓN.	61
TABLA 24. RESUMEN DE LA EJECUCIÓN DE LOS MEU EN EL ÁREA DE TDI.	69
TABLA 25. TABLA DE NORMALIZACIÓN PARA LA MÉTRICA CTP.	70
TABLA 26. MEDIDAS DE LA MÉTRICA CTP.	70
TABLA 27. TABLA DE NORMALIZACIÓN PARA LA MÉTRICA CPC.	70
TABLA 28. MEDIDAS DE LA MÉTRICA CPC.	71
TABLA 29. TABLA DE NORMALIZACIÓN PARA LA MÉTRICA CPF.	71
TABLA 30. MEDIDAS DE LA MÉTRICA CPF.	71
TABLA 31. TABLA DE NORMALIZACIÓN PARA LAS MÉTRICAS CORRESPONDIENTES A LA CARACTERÍSTICA TIEMPO.	72
TABLA 32. MEDIDAS DE LA MÉTRICA TEP.	72
TABLA 33. MEDIDAS DE LA MÉTRICA TEE.	73
TABLA 34. MEDIDAS DE LA MÉTRICA TEA.	73
TABLA 35. MEDIDAS DE LA MÉTRICA CTP POR ÁREA DE APLICACIÓN.	77
TABLA 36. MEDIDAS DE LA MÉTRICA CPC POR ÁREA DE APLICACIÓN.	78

TABLA 37. MEDIDAS DE LA MÉTRICA CPF POR ÁREA DE APLICACIÓN.....	79
TABLA 38. MEDIDAS DE LA MÉTRICA TEP POR ÁREA DE APLICACIÓN.....	80
TABLA 39. MEDIDAS DE LA MÉTRICA TEE POR ÁREA DE APLICACIÓN.....	80
TABLA 40. MEDIDAS DE LA MÉTRICA TEA POR ÁREA DE APLICACIÓN.....	81
TABLA 41. EXPERIENCIA (EN AÑOS) DE EXPERTOS/EVALUADORES POR ÁREA DE APLICACIÓN.	82
TABLA 42. CARACTERÍSTICAS SOBRESALIENTES DE LOS MEU.....	88
TABLA 43. RESUMEN DE LA COMBINACIÓN: EVALUACIÓN GLOBAL.	92
TABLA 44. RESUMEN DE LA COMBINACIÓN: EVALUACIÓN ESPECÍFICA.....	93
TABLA 45. RESUMEN DE LA COMBINACIÓN: EVALUACIÓN ENFOCADA A TAREAS ESPECÍFICAS (SIN RESTRICCIONES DE TIEMPO).	94
TABLA 46. COMPONENTES SOFTWARE DE UN SISTEMA INTERACTIVO.	95
TABLA 47. COMPONENTES HARDWARE DE UN SISTEMA INTERACTIVO.	95
TABLA 48. PORCENTAJES DE VOTACIÓN.	99
TABLA 49. RESUMEN DE LA EJECUCIÓN DE LOS MEU EN EL ÁREA DE REDES SOCIALES.....	101
TABLA 50. MEDIDAS DE LAS MÉTRICAS ASOCIADAS A LA CARACTERÍSTICA: <i>DETECCIÓN DE PROBLEMAS DE USABILIDAD</i>	101
TABLA 51. MEDIDAS DE LAS MÉTRICAS ASOCIADAS A LA CARACTERÍSTICA: <i>TIEMPO</i>	103

Lista de figuras

FIGURA 1. ESTRUCTURA DEL CAPÍTULO 2.....	7
FIGURA 2. FRAGMENTO DEL MAPA DE SELECCIÓN.....	18
FIGURA 3. ESTRUCTURA DEL CAPÍTULO 3. PARTE 1.....	19
FIGURA 4. ESTRUCTURA DEL CAPÍTULO 3. PARTE 2.....	20
FIGURA 5. REPRESENTACIÓN DE UNA ACTIVIDAD EN EL MFP, USANDO ELEMENTOS DE HAMSTERS.....	28
FIGURA 6. REPRESENTACIÓN DE LAS ACTIVIDADES EN EL MFP TRADICIONALMENTE DEFINIDO.....	29
FIGURA 7. REPRESENTACIÓN DE LAS ACTIVIDADES USANDO ÚNICAMENTE ELEMENTOS DE HAMSTERS.....	29
FIGURA 8. MFP EXTENDIENDO LA NOTACIÓN HAMSTERS.....	30
FIGURA 9. PORCIÓN DEL MFP TRADICIONAL – ETAPA DE EJECUCIÓN DE LA EVALUACIÓN HEURÍSTICA.....	32
FIGURA 10. PORCIÓN DEL MFP QUE INCLUYE ELEMENTOS DE HAMSTERS – ETAPA DE EJECUCIÓN DE LA EVALUACIÓN HEURÍSTICA.....	34
FIGURA 11. MFP DE LA ETAPA DE PLANEACIÓN DE LA EVALUACIÓN HEURÍSTICA. PARTE 1.....	45
FIGURA 12. MFP DE LA ETAPA DE PLANEACIÓN DE LA EVALUACIÓN HEURÍSTICA. PARTE 2.....	46
FIGURA 13. MFP DE LA ETAPA DE EJECUCIÓN DE LA EVALUACIÓN HEURÍSTICA.....	49
FIGURA 14. MFP DE LA ETAPA DE ANÁLISIS DE RESULTADOS DE LA EVALUACIÓN HEURÍSTICA. PARTE 1.....	51
FIGURA 15. MFP DE LA ETAPA DE ANÁLISIS DE RESULTADOS DE LA EVALUACIÓN HEURÍSTICA. PARTE 2.....	52
FIGURA 16. ESTRUCTURA DEL CAPÍTULO 4.....	57
FIGURA 17. MENÚ PRINCIPAL DE LA EPG.....	65
FIGURA 18. OPCIÓN "SEMANA" DE LA EPG.....	65
FIGURA 19. MENSAJES PUBLICADOS EN EL TABLÓN.....	66
FIGURA 20. MENSAJES ENVIADOS A TRAVÉS DEL CHAT.....	66
FIGURA 21. DISPOSITIVOS HARDWARE UTILIZADOS EN LAS EVALUACIONES DE USABILIDAD.....	67
FIGURA 22. ESTRUCTURA DEL CAPÍTULO 5.....	87
FIGURA 23. EVALUACIÓN GLOBAL DE LA USABILIDAD DE SISTEMAS INTERACTIVOS.....	91
FIGURA 24. EVALUACIÓN ESPECÍFICA DE LA USABILIDAD DE SISTEMAS INTERACTIVOS.....	92
FIGURA 25. EVALUACIÓN ENFOCADA A TAREAS ESPECÍFICAS.....	94
FIGURA 26. ESTRUCTURA DEL CAPÍTULO 6.....	97
FIGURA 27. SECUENCIA DE MÉTODOS SUGERIDA POR EXPERTOS.....	100

Capítulo 1

Lo importante es no dejar de hacerse preguntas. Albert Einstein

Introducción

Los sistemas interactivos¹ están creciendo en popularidad, y actualmente con las innovaciones tecnológicas, ocupan un lugar de importancia en la sociedad; esto incrementa el potencial de dichos sistemas para que los usuarios puedan entretenerse, informarse, comunicarse o usarlos en diferentes áreas de aplicación. Lo anterior implica un constante desafío, como es mejorar la calidad de los sistemas interactivos.

Uno de los factores más importantes al evaluar la calidad de un sistema interactivo corresponde al éxito o fracaso del usuario en las tareas que realiza. La satisfacción del usuario durante la realización de estas tareas, y el entendimiento real de la estructura del sistema con el cual interactúa también son factores a considerar [1].

En el entorno actual, en el que los sistemas interactivos están dirigidos a un público cada vez más amplio, a usuarios cada vez menos expertos en el manejo de los mismos, la *Experiencia de Usuario* (UX, por sus siglas en inglés *User eXperience*) [2] es un aspecto fundamental para el éxito de dichos sistemas. Entonces, ¿Cómo puede determinarse que los usuarios van a interactuar satisfactoriamente con los sistemas software? La respuesta parece ser simple: mejorando la calidad de la experiencia del usuario.

La UX se refiere a “cómo se sienten las personas acerca de un producto y su satisfacción cuando lo usan, lo miran, lo sostienen, lo abren o cierran” [2]. La UX abarca diferentes facetas relacionadas a la calidad de un producto software como [3]: usabilidad, accesibilidad, emotividad, multiculturalidad, jugabilidad, entre otras. Actualmente, el término UX está siendo ampliamente utilizado, tanto así que estándares como la ISO relacionada a calidad de uso y SQuaRE (Systems and Software Quality Requirements and Evaluation) [4], ya se refieren a esta terminología. Así, *la presente investigación se enfoca exclusivamente en la faceta usabilidad de la UX*, concretamente en el estudio de la usabilidad de sistemas interactivos, siguiendo el enfoque de una *usabilidad formativa* [5].

“La usabilidad es un atributo de calidad del software” [6] que conlleva una serie de métricas y métodos con el objetivo de obtener sistemas fáciles de usar y de aprender, además influye directamente en el éxito de cualquier nueva aplicación o sistema interactivo. La usabilidad es la característica de calidad más visible, puesto que determina la satisfacción del usuario con el sistema, lo que a su vez determina que éste quiera volver a utilizarlo en el futuro [7]. En muchos casos, la realización de pequeñas mejoras en la usabilidad de un sistema contribuye a un significativo aumento de la calidad de la experiencia del usuario con el sistema.

La usabilidad reduce los errores ocasionados por los usuarios y lleva a que estos realicen las tareas de manera más eficiente y efectiva, aumentando así su satisfacción y mejorando su experiencia global con la aplicación o sistema con el cual interactúan [5].

¹ Sistema interactivo se le denomina a aquel sistema que se interrelaciona y depende de las acciones de un usuario para realizar una tarea, es decir, todo sistema en el que interactúan persona y máquina.

Ahora bien, una variedad de investigadores han ejecutado varios métodos que intentan evaluar el grado de satisfacción de la usabilidad de diferentes sistemas interactivos, sin embargo, en escasos trabajos se considera la posibilidad de involucrar a un grupo de personas que trabajen de forma colaborativa en el proceso de evaluación. Además, en dichos trabajos la información detallada del proceso, como: entregables, requerimientos, roles, entre otra, no es lo suficientemente bien definida. En ese sentido, este trabajo propone la especificación colaborativa de un conjunto de métodos de evaluación (elaborada siguiendo los principios definidos en la Ingeniería de Colaboración [8]), la cual provee una secuencia de actividades bien definidas, especificación de entregables, descripción de los diferentes participantes del proceso de evaluación y especificación del proceso de comunicación entre los participantes. Esto con el fin de ofrecer documentación acerca de cómo ejecutar evaluaciones colaborativas de usabilidad de sistemas interactivos.

En general, y teniendo en cuenta lo anterior, *esta investigación intenta establecer una metodología de evaluación a través de la especificación de procesos colaborativos que soportan actividades de los métodos de evaluación que la conforman*, de tal manera que dichos procesos puedan ser aplicados por personas involucradas en la evaluación de la usabilidad de *sistemas software interactivos*. Dado que los métodos de evaluación de usabilidad están enfocados en recolectar datos relacionados con la interacción del usuario y un sistema software, la presente investigación pretende solamente abordar los sistemas interactivos para los cuales los métodos estudiados tienen alcance.

1.1 Planteamiento del problema

La evaluación de la usabilidad de un sistema interactivo es una de las etapas más importantes dentro del Diseño Centrado en el Usuario (DCU), ya que permite obtener las características de la usabilidad del sistema y la medida en que los atributos, paradigmas y principios de usabilidad se están aplicando en éste [9]. Los Métodos de Evaluación de Usabilidad (MEU), que permiten medir la aplicación de éste atributo en cierto sistema y bajo ciertos factores, son bastante diversos; su realización depende de variables tales como [10]: costos, disponibilidad de tiempo, recursos humanos que interpreten los resultados, entre otras. Así, la elección de un método para evaluar la usabilidad de un sistema interactivo no resulta sencilla [11]. Teniendo en cuenta lo anterior, y una serie de trabajos relacionados (ver Anexo A), *se hace necesario establecer una metodología de evaluación que podría dar lugar a sistemas interactivos con un alto nivel de usabilidad; una metodología que incluya combinaciones de MEU apropiadas para evaluar sistemas interactivos, y que además, promueva el trabajo colaborativo entre los participantes del proceso de evaluación*.

Los MEU se agrupan generalmente en dos grupos [9]: inspección y prueba, los cuales pueden ser usados iterativamente y en etapas distintas del desarrollo de un sistema. El camino más común para hacer las evaluaciones de usabilidad es realizar por lo menos uno de los métodos de inspección y uno de los métodos de prueba, dependiendo de la situación que se presenta. Esto porque los métodos de inspección son realizados por analistas y expertos en usabilidad con el objetivo de identificar errores y problemas de diseño del sistema en cuestión, mientras que los métodos de prueba permiten evaluar el sistema realizando pruebas empíricas del diseño de la interfaz con usuarios representativos [9][12]. Así, es posible definir los problemas de usabilidad en una primera etapa, para luego evaluar empíricamente la influencia de esos problemas. Ahora bien, la variedad de MEU, tanto de inspección como de prueba, ofrece distintos factores que

influyen en su realización, con características que determinan el tipo de información que puede obtenerse, las etapas en que pueden ser realizados, entre otras. Además, existen diferencias en los tiempos necesarios para llevar a cabo cierto método de evaluación, los equipos/materiales necesarios, cantidad de expertos, cantidad de usuarios, entre otras.

Ahora bien, los MEU tienen fortalezas y debilidades y están enfocados a evaluar aspectos específicos de usabilidad, por lo que expertos en usabilidad recomiendan combinarlos en una evaluación para complementarlos entre sí [10][13]. Varios MEU pueden ser realizados sobre sistemas interactivos, pero la incógnita está en el nivel de la información que entrega cada uno de ellos y/o la combinación de unos con otros. El problema surge al momento de decidir cuál de los métodos de evaluación existentes (métodos de inspección y de prueba) o cuál de sus combinaciones es apropiada para evaluar la usabilidad de sistemas software interactivos, de manera que pueda obtenerse la mayor cantidad de información relevante, considerando tiempos razonables, obtención de problemas importantes, entre otros factores; por tal razón surge la pregunta de investigación: **¿Cuáles de los MEU existentes son apropiados para ejecutar sobre sistemas software interactivos en distintos escenarios?**

Actualmente, la tendencia cada vez más común es la de trabajar colaborativamente entre personas para alcanzar un objetivo común. El trabajo se organiza en equipos y cada integrante –a menudo distante y distribuido por cualquier lugar del planeta– interactúa con el resto del grupo para obtener una mejor productividad [7][14]. Al integrar aspectos de trabajo colaborativo a un proceso determinado, el objetivo no es sólo la mejora de la comunicación, sino también lograr mayor participación, compromiso, entre los integrantes de un grupo que trabajan en torno a una actividad común, lo que conlleva a una mejor calidad del producto elaborado.

Centrándonos en el contexto del diseño y evaluación de interfaces de usuario, el proceso de evaluación de usabilidad no es ajeno a esta tendencia de trabajar colaborativamente. La disciplina Interacción Humano-Computador (HCI, por sus siglas en inglés *Human-Computer Interaction*) reconoce la necesidad de equipos multidisciplinares que permitan realizar una evaluación de forma más adecuada. Así, con el objetivo de contribuir al proceso de evaluación de usabilidad tradicionalmente definido, una serie de actividades de los MEU que van a conformar la metodología de evaluación serán especificadas de tal forma que integren procesos colaborativos (en los que participan varias personas de diferentes áreas de conocimiento, las cuales pueden estar distribuidas geográficamente), definan los roles de los miembros del grupo, el proceso de comunicación a establecer, entregables a generar, entre otra información relevante. Con base en lo anterior, surge una segunda pregunta de investigación, y es: **¿Cómo poder estructurar dichas actividades de forma colaborativa, de tal manera que permitan obtener efectivamente información sobre la usabilidad de sistemas software interactivos?** Considerando esta pregunta, en la presente investigación será utilizado el *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software* [15], el cual está basado en la *Metodología para el desarrollo de procesos colaborativos* [8] propuesta por investigadores en Ingeniería de Colaboración, con el objetivo de especificar la forma de realización colaborativa de un conjunto de actividades que conforman los MEU incluidos en la metodología de evaluación.

Teniendo en cuenta lo anterior, una vez identificados los MEU apropiados para ejecutar sobre sistemas software interactivos en distintos escenarios e integrados aspectos de trabajo colaborativo en actividades que conforman dichos métodos, surge la siguiente

pregunta de investigación: **¿Cómo combinar (integrar) los MEU para maximizar el potencial de identificación de información relevante?** Con base en este interrogante, un conjunto de MEU van a constituir la metodología de evaluación para que esta pueda brindar la información de la usabilidad de sistemas software interactivos de forma más completa que aquellos métodos de evaluación ejecutados de forma independiente. En ese sentido, la investigación está enfocada en estudiar un conjunto de MEU en distintos sistemas interactivos, siendo la televisión digital interactiva, web transaccional, aplicaciones móviles y redes sociales, las áreas de aplicación a utilizar como base experimental. Sin embargo, se espera que la metodología pueda ser utilizada para evaluar la usabilidad de sistemas software interactivos de diferentes áreas de aplicación, tales como: videojuegos, entornos virtuales de aprendizaje, entre otros.

Es importante destacar que los usuarios y sus necesidades toman la mayor importancia en este trabajo de investigación y alrededor de ellos gira la metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos, ya que finalmente son ellos quienes utilizan los sistemas para alcanzar un objetivo determinado. Lo buscado es que ellos alcancen dichos objetivos de manera fácil y eficiente, lo que contribuye directamente a su satisfacción.

1.2 Objetivos

1.2.1 Objetivo General

- Establecer una metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos.

1.2.2 Objetivos Específicos

- Identificar los métodos de evaluación de usabilidad de sistemas software interactivos apropiados para distintos escenarios.
- Especificar las actividades que requieren trabajo colaborativo para su ejecución, que conforman los métodos de evaluación seleccionados.
- Diseñar una metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos.
- Evaluar la metodología propuesta.

1.3 Hipótesis

El planteamiento de la hipótesis del presente trabajo de investigación está en torno a la aplicación de la metodología para la evaluación de usabilidad y su efectividad al momento de obtener información sobre la usabilidad de sistemas software interactivos. Así, se define la hipótesis de investigación y la hipótesis nula de la siguiente manera:

Hipótesis de investigación: la aplicación de una metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos, conformada por un conjunto de MEU, brinda la información de la usabilidad de forma más completa que aquellos métodos de evaluación ejecutados de forma independiente.

Hipótesis nula: la aplicación de una metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos, conformada por un conjunto de MEU, no brinda la información de la usabilidad de forma más completa que aquellos métodos de evaluación ejecutados de forma independiente.

De esta forma, se espera que al finalizar el trabajo la hipótesis sea satisfecha, a base de los procedimientos empíricos a lo largo de la presente investigación.

1.4 Organización del documento

La organización del documento de trabajo de grado está dividida en 8 capítulos, los cuales se describen brevemente a continuación:

El **Capítulo 1**, es el capítulo actual referente a la introducción, el cual se ha dividido en el planteamiento del problema, los objetivos del trabajo de grado, la hipótesis de solución y la estructura del documento.

El **Capítulo 2**, presenta los referentes teóricos necesarios para comprender la información presentada en el documento. Los referentes teóricos se dividen en los grupos: *componente usabilidad* y *componente colaborativo*.

El **Capítulo 3**, presenta la especificación colaborativa de los métodos de evaluación objeto de estudio. Esta información corresponde a la realización de cada una de las fases que conforman el *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software*. Adicionalmente, este capítulo presenta información acerca de una propuesta de notación para el diseño de procesos colaborativos.

El **Capítulo 4**, presenta la ejecución y análisis de resultados de los MEU objeto de estudio sobre sistemas que pertenecen a las áreas de aplicación: televisión digital interactiva, web transaccional y aplicaciones móviles. El análisis de resultados fue realizado con base en una serie de métricas descritas en este capítulo.

El **Capítulo 5**, presenta las distintas combinaciones de MEU que conforman la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos*, las cuales fueron propuestas con base en el análisis descrito en el Capítulo 4. Este capítulo presenta además información sobre la aplicabilidad de la metodología.

El **Capítulo 6**, presenta la evaluación de la metodología propuesta considerando un escenario específico de evaluación en el área de las redes sociales. Este capítulo incluye el análisis comparativo entre los resultados obtenidos en una combinación de MEU propuesta por expertos y una que hace parte de la metodología propuesta. También son presentadas algunas conclusiones acerca de la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos*.

El **Capítulo 7**, presenta una serie de experiencias y recomendaciones obtenidas gracias a la realización de la presente investigación.

Finalmente, el **Capítulo 8** presenta las conclusiones de los resultados obtenidos y su articulación con los objetivos planteados. Adicionalmente, son presentadas las actividades futuras para fortalecer el presente trabajo.

Capítulo 2

*Las palabras son como monedas,
que una vale por muchas como
muchas no valen por una.
Francisco de Quevedo*

Base conceptual

Este capítulo presenta los referentes teóricos que forman parte de la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos* según la Figura 1.

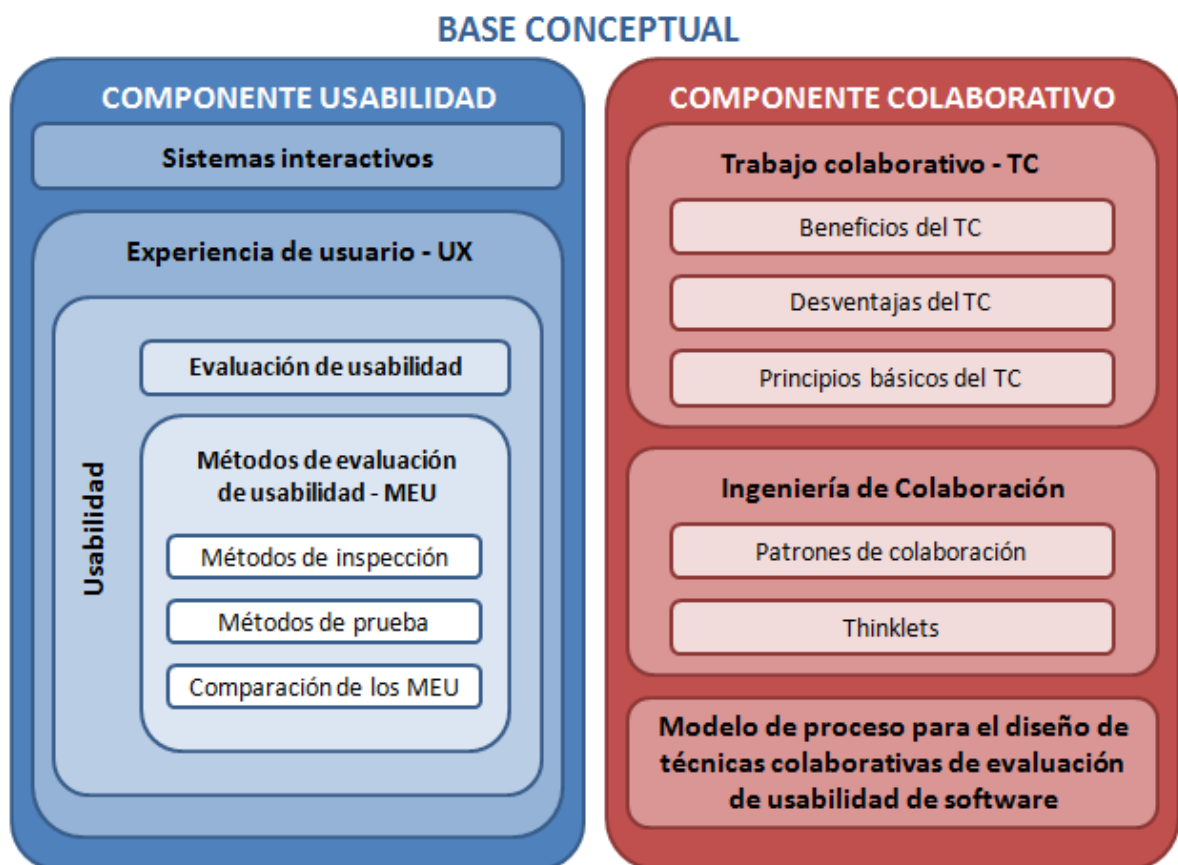


Figura 1. Estructura del capítulo 2.

2.1 Componente usabilidad

2.1.1 Sistemas interactivos

¿En qué consiste la “interacción”? Según la R.A.E. (Real Academia de la Lengua Española)², es la “acción que se ejerce recíprocamente entre dos o más objetos, agentes, fuerzas, funciones, etc.”. En el campo de las Tecnologías de la Información y de la Comunicación (TIC) se tiene la particularidad de que las personas interactúan entre ellas y con diversos dispositivos hardware y software para llevar a cabo una serie de tareas. Esos sistemas que

² Disponible en: <http://lema.rae.es/drae/?val=interacción>. Consultado en Julio de 2014.

actúan de interconexión entre personas, y que favorecen la realización de las tareas y el alcance de los objetivos propuestos son los denominados Sistemas Interactivos (SI).

Los SI se dividen en dos grupos [16][17]: SI software y SI hardware. Los SI software son los sistemas que incluyen algún tipo de software, tales como sitios web, aplicaciones de escritorio, interfaces de telefonía móvil, entre otros, algunos ejemplos son: formularios, tablas, imágenes, iconos, valores numéricos, entre otros. Los SI hardware incluyen dispositivos físicos (interfaces hardware) con los que se puede interactuar, algunos ejemplos son: teclado, pantalla, impresora, lector de tarjetas, entre otros. Es importante mencionar que todos los SI hardware incluyen partes de software (las partes de software son una pequeña parte de un SI hardware grande) [16].

En la presente investigación los conceptos *Sistema Interactivo* y *Sistema Software Interactivo* serán utilizados indistintamente, puesto que al hacer referencia a este último, implícitamente se tienen en cuenta aspectos relacionados al hardware del sistema.

2.1.2 Experiencia de usuario

El término Experiencia de Usuario (o UX, por sus siglas en inglés *User eXperience*) se refiere a “cómo se sienten las personas acerca de un producto y su satisfacción cuando lo usan, lo miran, lo sostienen, lo abren o cierran” [2]. Actualmente, existen diferentes definiciones de la UX utilizadas por profesionales en el área de HCI, siendo una de las más destacadas la presentada en el estándar ISO 9241-210 [18]: “*Percepciones y respuestas de una persona que resultan de la utilización y/o uso anticipado de un producto, sistema o servicio*”.

En [19], se recopilan una serie de definiciones de UX, sin embargo, estas son válidas en contextos de uso específicos y no abarcan todos los aspectos que deberían considerarse para evaluar la UX que se obtiene al interactuar con un sistema interactivo. Así, en [20] se propone una definición más general tal que cubre todos los aspectos requeridos: “*La experiencia de usuario atiende a todos los factores, tanto internos como externos del usuario y del sistema interactivo, que causen alguna sensación a quien esté utilizando un sistema interactivo concreto en un determinado contexto de uso*”.

La UX abarca diferentes facetas relacionadas a la calidad de un producto software. El estándar ISO/IEC 25010 [4] considera de forma general las siguientes facetas de la UX: accesibilidad, dependabilidad, emotividad, jugabilidad, usabilidad, entre otras. Así pues, la presente investigación está enfocada exclusivamente en la faceta *usabilidad* de la UX, concretamente en el estudio de la usabilidad de sistemas software interactivos. La selección de la faceta usabilidad obedece a que es la más directamente relacionada con la calidad en uso. Además, porque al momento de su evaluación, esta permite obtener información indirectamente sobre otras de las facetas.

2.1.2.1 Usabilidad

El término *usabilidad* coloquialmente es definido como “facilidad de uso” ya sea de una página web, una aplicación informática o cualquier otro sistema que interactúe con un usuario [21]. En la actualidad existe una serie de definiciones formales, aunque no existe una definición consensuada de este término. Por esta razón, serán presentadas una serie de definiciones formales y dadas por autores prominentes del área con el objetivo de establecer la idea general del concepto.

Según la norma ISO 9241 [6], “*Usabilidad es el grado en el que un producto software puede*

ser utilizado por usuarios específicos para alcanzar objetivos específicos con efectividad, eficiencia y satisfacción en un contexto de uso específico". Es una definición centrada en el concepto de calidad en uso, es decir, se refiere a cómo el usuario realiza tareas específicas en escenarios específicos con efectividad. La efectividad se refiere al nivel de exactitud con que el usuario cumple los objetivos; la eficiencia se refiere a los recursos usados para la concreción de estos objetivos por parte del usuario, mientras que la satisfacción está relacionada con la comodidad y postura del usuario durante la interacción con el producto.

La definición de usabilidad dada por Bevan es [22]: *"La facilidad de uso y aceptabilidad de un sistema o producto para una clase particular de usuarios que llevan a cabo tareas específicas en un ambiente específico, donde la facilidad de uso afecta el rendimiento y satisfacción del usuario y la aceptabilidad afecta si el producto es utilizado o no"*. Esta definición es bastante completa y pretende integrar las dos definiciones dadas por la ISO.

Por otro lado, Jakob Nielsen, define la usabilidad con base en el siguiente comentario [5]: *"La utilidad de un sistema en cuanto a medio para conseguir un objetivo, tiene un componente de funcionalidad (utilidad funcional) y otro basado en el modo en que los usuarios pueden usar esta funcionalidad"*. Asimismo, define la usabilidad en torno a componentes medibles: facilidad de aprendizaje, eficiencia, facilidad de recordar, errores y satisfacción subjetiva.

Finalmente, para el presente trabajo de investigación será adoptada la definición dada por la ISO 9241, puesto que es ampliamente conocida y representa completamente el sentido del término en cuestión, tal como se menciona en el trabajo relacionado [1].

2.1.2.1.1 Evaluación de usabilidad

La evaluación de usabilidad se ha determinado como la actividad que comprende un conjunto de métodos que analizan la calidad de uso de un sistema interactivo, en diferentes etapas del ciclo de vida del desarrollo [23]. Este proceso puede ser llevado a cabo por personas con diferentes habilidades y conocimientos, involucrando usuarios potenciales y actuales, expertos en usabilidad, diseñadores de sistemas, entre otros. Es necesario realizar la evaluación de usabilidad para validar que el producto final cumple con los requerimientos y es usable [7]. La evaluación tiene tres objetivos principales [24]:

- Evaluar el alcance y la accesibilidad de la funcionalidad de los sistemas. La evaluación de la funcionalidad del sistema debe estar orientada a identificar que las funcionalidades cumplan con los requerimientos del usuario.
- Evaluar la experiencia del usuario en su interacción con el sistema.
- Identificar problemas específicos del sistema. El objetivo final de la evaluación es identificar problemas específicos en el diseño.

Como consecuencias directas de la evaluación se tiene [13]:

- Mejoramiento en la calidad de los procesos: derivada de una cultura de desarrollo organizada y consciente de la importancia de la evaluación.
- Mejoramiento en la calidad en los productos: validación consciente y temprana de los diferentes módulos que conforman el sistema.
- Manejo eficiente de los recursos tiempo y dinero: consecuencia derivada de la corrección temprana de fallas.
- Posibilidad de reproducir éxitos en otros proyectos: cada módulo desarrollado se

convierte en una fuente confiable de código reutilizable, además de evolucionar a procedimientos ágiles y óptimos para la evaluación del sistema.

- Dominación de los riesgos del proceso: entre más rápido se detecten fallas, las estrategias de contingencia de riesgos serán más efectivas.
- Confianza y satisfacción del cliente: la validación conjunta con el usuario, evita sorpresas desagradables en etapas críticas del desarrollo.

La evaluación de usabilidad es una parte fundamental del enfoque iterativo del desarrollo de software, porque las actividades de evaluación pueden producir soluciones de diseño para su aplicación en el próximo ciclo de desarrollo o, al menos, un mayor conocimiento sobre la naturaleza del problema de interacción detectado. Por tanto, la evaluación de usabilidad es parte inherente del proceso de desarrollo [25].

2.1.2.1.2 Métodos de evaluación de usabilidad – MEU

La evaluación de la usabilidad de un sistema software es una de las etapas más importantes dentro del diseño centrado en el usuario, ya que permite obtener las características de la usabilidad de un sistema y la medida en que los atributos, paradigmas y principios de usabilidad se están aplicando en éste [1]. Es por esto que los Métodos de Evaluación de Usabilidad (MEU) se han convertido en una fuente interesante de estudio por parte de los investigadores de la usabilidad, sus características de aplicación, la variedad de métodos existentes y los resultados que generan.

Un método de evaluación es un procedimiento sistemático para recolectar datos relacionados con la interacción del usuario final y un producto software o sistema [26]. Cada dato recolectado, según el modelo planteado, permitirá establecer el grado de satisfacción de los indicadores asociados con cada atributo, métrica y criterio; los indicadores satisfechos son el reflejo de la usabilidad del sistema [27].

Existen diversos MEU, cada uno con sus características propias, los cuales se pueden desarrollar sobre diferentes representaciones del sistema, sean prototipos en papel, prototipos funcionales, sistemas terminados, etc. Estos métodos permiten establecer una comunicación entre el usuario y los desarrolladores, ya que estos últimos identifican los objetivos, percepciones, problemas y cuestionamientos de los usuarios [1]. Adicionalmente, los MEU permiten validar las decisiones de diseño, descubriendo los problemas y los aciertos asociados al sistema, marcando las diferencias y pensamientos comunes, respectivamente, entre los desarrolladores y los usuarios [1].

Los MEU se clasifican en dos grupos [1][28]: (i) *métodos de inspección*: consisten en inspecciones del diseño de la interfaz realizadas por expertos en usabilidad y (ii) *métodos de prueba*: consisten en pruebas empíricas del diseño de la interfaz con usuarios representativos. Estas clasificaciones y los principales métodos involucrados en cada uno de los grupos son presentados a continuación.

Métodos de inspección

Los métodos de inspección corresponden a MEU en los que son realizadas inspecciones del diseño de la interfaz usando métodos heurísticos. Son realizados por expertos en usabilidad, y están basados en el recorrido y análisis del sistema en evaluación, identificando errores y problemas de diseño. Se basan en el conocimiento y experiencia de los expertos [5]. Dentro de este grupo de métodos, según autores de estudios como [1][7][10], los más relevantes

son: evaluación heurística, recorrido cognitivo, recorrido pluralista, inspección de estándares, análisis de acciones. La descripción de estos métodos es presentada en el Anexo B.1, destacando sus características más relevantes.

Métodos de prueba

Los métodos de prueba corresponden a MEU que realizan pruebas empíricas del diseño de la interfaz con usuarios representativos, es decir, son pruebas basadas en la experiencia real de los usuarios. Estos métodos se basan en la premisa de que no puede asegurarse cuán usable es un determinado prototipo o producto software, sin antes probarlo con usuarios representativos del sistema [10]. Los usuarios realizan tareas concretas en el sistema, los evaluadores analizan los resultados para determinar si la interfaz brinda a los usuarios el soporte que requieren para la realización de dichas tareas [5].

Dentro de las características más importantes de los métodos de prueba está el hecho de que la evaluación es realizada bajo condiciones controladas y de la forma más representativa posible de la realidad (lugar real de trabajo con el sistema o laboratorios de usabilidad), en donde los usuarios realizan ciertas acciones asignadas por los evaluadores con el fin de obtener información de su experiencia con el sistema en evaluación, de manera que puedan detectarse problemas de usabilidad [5].

Las ventajas de realización de los métodos de prueba son la representatividad del mundo real del entorno de la prueba y la buena capacidad de encontrar problemas importantes y recurrentes de la usabilidad del sistema en evaluación. Ahora bien, la desventaja más importante está en el costo de establecimiento de las instalaciones necesarias para un ambiente lo más representativo posible [1].

El lugar físico de desarrollo de los métodos de prueba es un aspecto muy importante ya que, al ser pruebas empíricas, los usuarios deben sentirse cómodos en el lugar donde realizan las pruebas y debe ser un ambiente lo más representativo posible de la realidad, sino el mismo lugar de interacción con el sistema. Para este objetivo, han sido construidos los laboratorios de usabilidad en los cuales son realizados métodos de prueba de usabilidad. Estos permiten tener un mayor control sobre el usuario y las acciones que realiza, así las medidas se tornan más precisas [1]. De esta manera, para los evaluadores el laboratorio de usabilidad resulta muy provechoso, por las ventajas y características que ofrece; para los participantes de las pruebas, sin embargo, el hecho de tener que trasladarse a estos laboratorios puede ser un obstáculo en ciertas ocasiones.

Los métodos de prueba siguen un mismo proceso, en donde es necesaria la preparación de la prueba a realizar, el entorno de la prueba, definición de los perfiles de usuarios, selección de los usuarios representativos, entre otros aspectos. Para realizar los métodos es necesario seguir algunos pasos, los cuales se presentan a continuación [1]:

- Definición de la prueba: en este paso se definen aspectos básicos para realizar la prueba como: propósito de la prueba, perfiles de usuarios, lista de tareas, entorno de la prueba, datos a recolectar, entre otros. Esto con el fin de obtener una clara visión del plan y los objetivos de la prueba.
- Selección de participantes: los participantes de la prueba deben corresponder a usuarios representativos que se ajusten al perfil de usuario definido.
- Preparación de los materiales: preparación de la documentación necesaria para

realizar la prueba, como los acuerdos de confidencialidad, posibles documentos de orientación sobre la prueba, apuntes para evaluar el progreso de la prueba, cuestionarios pre y pos prueba y cualquier documento que se deba disponer.

- Ejecución de prueba piloto: antes de realizar la prueba misma es necesario ejecutar una prueba piloto, que asegure una correcta definición de la prueba y de los pasos que la conforman. Sirve para evaluar el correcto desempeño de los implementos, para estimar la duración de las tareas y del proceso. Debe realizarse previamente a la ejecución de las pruebas reales para tomar medidas correctivas si es necesario.
- Realización de la prueba: los participantes realizan la prueba de usabilidad. Debe existir una persona encargada de la comunicación con el participante, mientras los evaluadores se encargan de observar el proceso para obtener la información necesaria. Es importante recalcar al participante que la prueba pretende evaluar la usabilidad del producto, no el nivel de eficiencia o inteligencia de la persona. Toda la información es almacenada para su posterior análisis.
- Análisis de la prueba: en este paso son estructuradas las interpretaciones y descripciones de la experiencia del usuario. Son analizados los problemas de usabilidad encontrados y se postula un conjunto de posibles soluciones a esos problemas. En este paso se realizan dos conjuntos de acciones importantes: (i) *síntesis de los datos*: recolección y síntesis de los datos relevantes para su posterior análisis, tales como tiempos de las tareas, información estadística (promedios, mínimos, máximos, etc.), porcentajes de completitud de tareas, etc., (ii) *análisis de los datos*: se realiza un estudio de los datos recolectados para identificar y formalizar los problemas de usabilidad.
- Finalmente, es elaborado un informe final de la prueba, el cual contiene toda la información del proceso.

Todo este proceso puede ser aplicado a cualquiera de los métodos de prueba. Al final del proceso han sido encontrados los problemas de usabilidad del sistema en evaluación y han sido propuestas algunas posibles soluciones, por lo que sería factible y recomendable hacer los cambios en el sistema, con el fin de mejorar la experiencia del usuario durante la interacción. Así, se reconoce un proceso iterativo, en el que los métodos de prueba detectan problemas de usabilidad para la posterior modificación del sistema, luego es evaluado nuevamente el sistema mediante las pruebas y son realizados los cambios requeridos, y así sucesivamente, hasta que sea necesario o se cuente con la factibilidad técnica y económica para realizar el proceso [1].

Dentro de este grupo de métodos, según autores de estudios como [1][7][10], los más relevantes son: pensando en voz alta, interacción constructiva, experimentos formales, métodos de interrogación (cuestionarios y entrevistas), grabación del uso, medida de prestaciones, test retrospectivo, método conductor. La descripción de estos métodos es presentada en el Anexo B.2, destacando sus características más relevantes.

Comparación de los MEU

Cada MEU tiene sus propias características, ventajas y desventajas que lo diferencian de los otros, lo que permite captar las fortalezas y debilidades que determinarán cuándo y bajo qué condiciones es conveniente realizar uno de los métodos sobre otro, ya sea para los métodos de inspección como para los métodos de prueba. En el Anexo B.3 se presentan las ventajas y desventajas de los MEU mencionados anteriormente, las cuales corresponden a una recopilación de distintas fuentes bibliográficas relacionadas con el tema tales como

[1][5][7][10][13][27][29]. Esto con el fin de obtener una visión más completa de sus características, de tal manera que pueda realizarse un análisis comparativo entre los MEU.

2.2 Componente colaborativo

2.2.1 Trabajo colaborativo

El trabajo colaborativo es definido como “la nominación general y neutral de múltiples personas que trabajan juntas para producir un producto o servicio” [30]. También, el trabajo colaborativo es definido como un conjunto de procesos de trabajo que se relacionan; estos procesos generan tareas para ser desarrolladas por los integrantes del equipo de trabajo, buscando alcanzar objetivos comunes [15]. El trabajo colaborativo persigue “el desarrollo de conocimiento compartido entre los participantes del grupo, la aceleración de los flujos de información, la coordinación de los flujos de recursos para producir economías de costos y tiempos” [14]. El aspecto más importante de trabajar en grupo es el hecho de contar con un objetivo común que canaliza los esfuerzos individuales y ofrece un sentido de pertenencia que fomente la unión [14]. Los miembros del grupo tienen la oportunidad de aprender tomando en consideración otros puntos de vista, maneras distintas de hacer las cosas, interpretaciones diferentes de conceptos, experiencia de otros y la forma como dan solución a problemas similares [31].

El trabajo colaborativo permite compartir conocimiento y recursos a varias personas dispersas o no geográficamente. Dado que la colaboración tiene especial importancia, con ella se busca coordinar toda una serie de actividades que propendan por “la comunicación, la coordinación y la negociación con el fin de aumentar la productividad” [14], de esta manera, el trabajo colaborativo tiende a “maximizar los resultados y minimizar la pérdida de tiempo y de información” [14].

2.2.1.1 Beneficios del trabajo colaborativo

Fomentar el trabajo colaborativo durante la ejecución de una actividad entre los diferentes integrantes de un grupo, genera ciertas ventajas tales como [32]:

- Un grupo entiende mejor un problema que una sola persona.
- Existe responsabilidad compartida.
- Facilita la detección de errores.
- Un grupo posee más información (conocimiento) que un solo miembro. Hay más alternativas para la resolución de problemas.
- Se produce sinergia: la efectividad y calidad de la producción de un grupo es mayor que la suma de lo que pueda producir cada miembro en forma individual. La ejecución de las decisiones es asumida de mejor manera.
- Acceso a un mayor volumen de información útil y filtrada gracias a las contribuciones de otras personas.
- Agilización de procesos de aprendizaje ante la posibilidad de recurrir a miembros experimentados del grupo.

2.2.1.2 Desventajas del trabajo colaborativo

Es necesario especificar que no siempre el trabajo grupal es beneficioso, esta forma de trabajo puede presentar algunas desventajas, entre las cuales se destacan [32]:

- Procesos lentos, consumo de tiempo (ejemplo: tendencia a repetir ideas que ya

- fueron mencionadas).
- Falta de coordinación del trabajo hecho por el grupo.
- Influencias inapropiadas de las dinámicas de grupo (ejemplo: control del tiempo, participación no equilibrada).
- Tendencia a “descansar en otros” (los otros harán el trabajo).
- Incapacidad de completar tareas (responsabilidades no claras).
- Pérdida de tiempo productivo (esperas, interacción social).
- La pérdida del interés en trabajar activamente debido a los logros y protagonismo de otros miembros del grupo.

2.2.1.3 Principios básicos del trabajo colaborativo

En [33] se han propuesto unos principios que son la base para que un grupo pueda desempeñarse bajo el enfoque del trabajo colaborativo:

- Es necesario que cada uno de los integrantes contribuya de un modo particular al logro de las metas del grupo. Nadie gana méritos “a costa” del trabajo de los demás.
- Con el objetivo de alcanzar metas comunes, es necesario que todos los integrantes brinden apoyo y ayuda en el cumplimiento de las tareas y el trabajo.
- La responsabilidad debe ser igual para cada integrante del equipo.
- Son indispensables las habilidades interpersonales tales como: confianza mutua, comunicación clara y sin ambigüedades, apoyo mutuo y resolución constructiva de conflictos, durante el trabajo en equipo.
- Debe propenderse por el respeto de puntos de vista diferentes entre los integrantes.
- La formación de grupos es intencional y basada en la heterogeneidad. Los grupos se constituyen con base a las diferencias de habilidades, así como de características de personalidad y género de los integrantes.

2.2.2 Ingeniería de Colaboración

En primera instancia se podría definir el término colaboración, como la “unión de esfuerzos intelectuales por los miembros de un grupo de trabajo que buscan alcanzar una meta común” [34]. De aquí parte la Ingeniería de Colaboración la cual está enfocada al diseño de procesos colaborativos reutilizables. Los procesos colaborativos necesitan ser explícitamente diseñados, estructurados y manejados. Este es el eje central de la Ingeniería de Colaboración, en la cual “se diseñan procesos repetitivos colaborativos, los cuales se pueden transferir a grupos, usando técnicas y tecnología de colaboración” [35].

La Ingeniería de Colaboración se define como “un enfoque para el diseño de prácticas de trabajo colaborativo para tareas recurrentes destacadas y transferencia de esos diseños a practicantes para que los ejecuten por ellos mismos, sin la intervención continua de un facilitador profesional” [34]. De esta definición, se destacan los siguientes términos:

- Una *práctica de trabajo colaborativo* es un “proceso recurrente que solamente puede completarse por la combinación del esfuerzo mental de varias personas” [34]. Diseñar un proceso colaborativo en Ingeniería de Colaboración significa “crear y documentar una prescripción”, entendiéndose prescripción como “un documento escrito que define un conjunto estructurado de pasos para lograr objetivos y las condiciones bajo las cuales esos pasos serán ejecutados” [34].
- Se dice que una tarea es *destacada* “si la organización deriva beneficios sustanciales, o por la cual se anticipa una pérdida por no alcanzarla con éxito” [34].

- Una tarea es *recurrente* si “la tarea debe ser manejada repetidamente y se puede completar usando un diseño de proceso similar cada vez que se ejecuta” [34].

La Ingeniería de Colaboración se enfoca en las tareas recurrentes por dos razones, primero porque “el beneficio de diseñar una tarea *ad hoc*³ se incrementa solamente una vez, mientras que para tareas recurrentes se incrementa cada vez que la tarea es utilizada” [35]. En segundo lugar, para las personas que trabajan en tareas *ad hoc* “no es indispensable aprender sobre su experiencia dado que es poco probable que el proceso se repita; a diferencia de los que participan en actividades repetitivas, quienes necesitan adquirir cierta habilidad y brindarle conocimiento a los demás participantes” [35]. En la Ingeniería de Colaboración se destacan los elementos: *patrones de colaboración* y *thinklets*, que son la base para obtener la especificación colaborativa de un proceso/actividad determinada. Dichos elementos son descritos a continuación.

2.2.2.1 Patrones de colaboración

En las investigaciones realizadas alrededor de la Ingeniería de Colaboración, se han establecido una serie de patrones, relacionados con la forma en la cual un grupo trabaja colaborativamente hacia sus metas. Estos patrones, denominados *patrones de colaboración* se definen en términos del “movimiento del grupo desde su estado inicial hasta su estado final” [36]. Cada patrón tiene subpatrones que se pueden relacionar con las actividades, en la descripción del proceso genérico. La Tabla 1 presenta los *patrones de colaboración* con sus respectivos subpatrones:

Tabla 1. Patrones de colaboración.

PATRONES DE COLABORACIÓN	
Patrón: Generación	Es un patrón a partir del cual el grupo crea contenido. Consiste en pasar de tener pocos a muchos conceptos que son compartidos por el grupo.
Subpatrones asociados:	
<ul style="list-style-type: none"> • <i>Reunir</i>: Coleccionar y compartir conceptos entre los miembros del grupo. • <i>Crear</i>: Producir y compartir nuevas ideas que no fueron previamente conocidas por los miembros del grupo. • <i>Elaborar</i>: Adicionar detalles a los conceptos que ya fueron compartidos por los miembros del grupo. 	
Patrón: Reducción	El objetivo de este patrón es mantener sólo la información que cumple con un determinado criterio o criterios. Consiste en pasar de tener muchos conceptos a unos pocos que el grupo considere que requieren mayor atención.
Subpatrones asociados:	
<ul style="list-style-type: none"> • <i>Seleccionar</i>: Escoger un subconjunto de conceptos existentes. • <i>Abstraer</i>: Derivar conceptos más generales desde instancias específicas, en el conjunto existente. • <i>Resumir</i>: Capturar la esencia de los conceptos, sin eliminar conceptos únicos. 	
Patrón: Clarificación	El objetivo de este patrón es lograr el entendimiento común de conceptos manejados por el grupo. Consiste en pasar de tener un menor a un mayor conocimiento compartido de los conceptos, las palabras y frases usadas para expresarlos.
Subpatrón asociado:	
<ul style="list-style-type: none"> • <i>Describir</i>: proponer explicaciones y descripciones alternativas de un concepto. 	
Patrón: Organización	Consiste en pasar de tener un menor a un mayor conocimiento de las relaciones entre los conceptos que el grupo este considerando.
Subpatrones asociados:	
<ul style="list-style-type: none"> • <i>Clasificar</i>: ordenar conceptos en categorías. • <i>Estructurar</i>: crear arreglos espaciales entre conceptos para representar sus relaciones conceptuales. 	
Patrón: Evaluación	Consiste en pasar de un menor a un mayor conocimiento del valor relativo de los conceptos bajo consideración. Este patrón tiene como efectos apoyar a la toma de decisiones y a la

³ Ad hoc es una expresión que significa literalmente “para esto”, una tarea ad hoc hace referencia a una tarea diseñada específicamente para una ocasión o contexto determinado.

PATRONES DE COLABORACIÓN	
	comunicación del grupo.
Subpatrones asociados:	
<ul style="list-style-type: none"> • <i>Votar</i>: evaluar la opinión del grupo, respecto a los conceptos. • <i>Categorizar</i>: identificar un orden de preferencia entre los conceptos. • <i>Valorar</i>: Especificar y detallar en el valor de los conceptos. 	
Patrón: Construcción de Consenso	Consiste en moverse de tener pocos a muchos miembros del grupo quienes estarán dispuestos a comprometerse para un objetivo.
Subpatrones asociados:	
<ul style="list-style-type: none"> • <i>Medir</i>: Evaluar el grado en que los participantes están dispuestos a comprometerse con una propuesta. • <i>Diagnosticar</i>: Buscar la comprensión de las causas de desacuerdos. • <i>Apoyar</i>: Tratar de convencer a otros a aceptar y adoptar una posición. • <i>Resolver</i>: Buscar la forma de superar las causas de desacuerdos. 	

2.2.2.2 Thinklets

Como fue indicado anteriormente, los *patrones de colaboración* son una guía de cómo se ejecutará un proceso. Sin embargo, la descripción de los patrones no presenta una forma detallada para guiar a un equipo de forma precisa a través de la ejecución de un proceso. Así, debe ser posible seleccionar bloques de construcción existentes y unirlos para especificar cómo un determinado patrón debe realizarse. Los *thinklets* proveen estas capacidades, son “técnicas de facilitación repetibles, transferibles y predecibles para asistir a un grupo en alcanzar su objetivo acordado” [37]. Un *thinklet* “constituye la unidad más pequeña del capital intelectual necesario para crear un *patrón de la colaboración* repetible y predecible entre las personas que trabajan hacia un objetivo” [38].

Los *thinklets* se detallan completamente y son modificables. Pueden usarse para construir nuevos grupos de procesos; son recetas creadas para ser utilizadas por novatos, ya que son fácilmente aprendidos, recordados y pueden adaptarse fácilmente a un diseño de proceso [35]. Una de las grandes ventajas de los *thinklets* es que los diseñadores de procesos colaborativos pueden emplearlos para escoger soluciones conocidas y no invertir esfuerzos en inventar y probar nuevas. Esto puede reducir tanto el esfuerzo como el riesgo durante el desarrollo de procesos grupales [38].

La documentación que provee un *thinklet*, cuenta con la siguiente información [39]:

- Un nombre metafórico o representativo relacionado con los patrones que crea el *thinklet*.
- Criterios para decidir cuándo escoger o no el *thinklet*.
- Información general del *thinklet*, como entradas y salidas. Las entradas son variables que deben ser instanciadas al tiempo que el *thinklet* es usado y las salidas son los entregables que deben generarse.
- Cómo usar el *thinklet*. Aquí son descritos los pasos que conforman el *thinklet*.
- Una historia exitosa que ayuda a clarificar las circunstancias bajo las cuales el *thinklet* es útil.
- Una explicación del nombre, esto hace más fácil recordar el nombre del *thinklet*.

El Anexo C presenta los *thinklets* utilizados en la presente investigación, los cuales han sido propuestos por investigadores en Ingeniería de Colaboración.

2.2.3 Modelo de proceso para el diseño de técnicas colaborativas de evaluación

de usabilidad de software

El *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software* [15], el cual está basado en la *Metodología para el desarrollo de procesos colaborativos* [8] propuesta por investigadores en Ingeniería de Colaboración, permite obtener la especificación colaborativa de un método de evaluación de usabilidad, a partir de la identificación de tareas/actividades recurrentes y/o destacadas. De esta manera, las actividades especificadas de forma colaborativa promueven la comunicación, la coordinación y la negociación, con el fin de aumentar la productividad en la realización de dichas actividades. El modelo se compone de las siguientes fases [15]: diagnóstico de la técnica, descomposición de la actividad, establecer tareas colaborativas, relación con proceso colaborativo, documentación de la técnica y validación. El procedimiento a seguir en cada fase es el siguiente:

Fase 1 – Diagnóstico de la técnica: en esta fase se realiza una descripción detallada del método de evaluación objeto de estudio. Esta descripción incluye información como: entregables, requerimientos, participantes y demás características relevantes del método.

Fase 2 – Descomposición de la técnica: para el método de evaluación de usabilidad objeto de estudio, deben identificarse las actividades generales que lo componen y determinar la secuencia entre ellas.

Fase 3 – Establecer tareas colaborativas: en esta fase se describen las subactividades que componen cada una de las actividades generales identificadas en la fase anterior (Fase 2). La división de las subactividades permite identificar cuáles se realizarían de forma colaborativa, a las cuales se les asocia uno (ó más) *patrones de colaboración*.

Fase 4 – Relación con proceso colaborativo: en esta fase se relacionan los *thinklets* a las subactividades definidas como colaborativas, considerando los *patrones de colaboración* asociados. Los *thinklets* identificados deben adecuarse a los recursos, al grupo y hasta las propias habilidades de las personas involucradas en la ejecución de los procesos colaborativos [34]. Para seleccionar los *thinklets* fueron considerados los siguientes aspectos:

- *Criterios para decidir cuándo usar o no el thinklet*, los cuales se encuentran en la documentación que provee cada *thinklet*.
- *Pasos que conforman el thinklet*, estos se encuentran en la documentación del *thinklet*. Este aspecto es útil ya que permite determinar si los pasos que ofrece el *thinklet* se ajustan de manera adecuada al proceso que conforma la actividad.
- *Mapa de selección* [40]: indica el tipo de relación entre los *thinklets*. Los tipos de relación se representan mediante colores: el verde indica si la relación es excelente; naranja si la relación es posible pero difícil y rojo si la relación es imposible. Este aspecto es útil ya que permite identificar los *thinklets* que tienen una excelente relación, lo cual da mayor certeza en que se obtendrán los resultados adecuados con su ejecución, además sirve para descartar las relaciones que se han establecido como imposibles. La Figura 2 presenta un fragmento del mapa de selección.

La relación de estos thinklets es:	FreeBrainstorm	OnePage	Comparative Brainstorm	LeafHopper	DealersChoice	PlusMinusInteresting
Excelente Posible pero difícil Imposible						
Punto de partida?						
FreeBrainstorm						
OnePage						
Comparative Brainstorm						
LeafHopper						
DealersChoice						
PlusMinusInteresting						

Figura 2. Fragmento del mapa de selección.

Fase 5 – Documentación de la técnica: a partir de la información obtenida en las fases anteriores, en esta fase deben generarse los siguientes elementos definidos en Ingeniería de Colaboración [8]: modelo de facilitación del proceso y agenda detallada.

1. *Modelo de Facilitación del Proceso (MFP):* es utilizado para mostrar el flujo del proceso y los elementos tales como *thinklet*, número de secuencia, *patrón de colaboración* y nombre, relacionados con cada una de las actividades que conforman el proceso.
2. *Agenda detallada:* documento que presenta de manera clara y detallada la información de las actividades que forman parte del proceso diseñado. La agenda detallada consta de la siguiente información: número de secuencia de la actividad, descripción de la actividad, entregables, *thinklet* y *patrón de colaboración* que está siendo utilizado, descripción detallada de los pasos requeridos para ejecutar la actividad (proceso colaborativo) y participantes de la actividad.

Fase 6 – Validación: en esta fase se valida la especificación colaborativa del método de evaluación a partir de las siguientes formas de validación propuestas en Ingeniería de Colaboración [8]:

- *Prueba piloto:* se ejecuta el método de evaluación colaborativo diseñado; como resultado de esta ejecución debe evaluarse la efectividad del proceso. El objetivo es verificar si la ejecución del método puede llevarse a cabo con los recursos definidos.
- *Recorrido:* con algunos de los participantes en la ejecución del método de evaluación colaborativo, es realizada la evaluación del mismo, para identificar falencias y dificultades durante la ejecución.
- *Simulación:* el equipo de trabajo que diseñó el método de evaluación colaborativo da respuesta a una serie de preguntas, como por ejemplo: ¿Estos pasos son suficientes?, ¿Está toda la información disponible?, entre otras. Esta validación prueba la lógica del diseño y si a cada paso se le creará el entregable requerido. Las respuestas generadas en esta forma de validación servirán de referente para realizar las mejoras respectivas.
- *Revisión:* se genera discusión sobre el diseño realizado, entre el equipo de trabajo y los potenciales usuarios que ejecutarán el método de evaluación colaborativo. Discutir el diseño revelará diferentes perspectivas, de esta forma se identifican partes del diseño ineficientes.

Capítulo 3

La gran pregunta sigue siendo: ¿es esta la manera más rápida y útil de hacerlo? Jakob Nielsen

Especificación colaborativa de los MEU

Este capítulo presenta la especificación colaborativa de los métodos de evaluación de usabilidad objeto de estudio. Esta información corresponde a la realización de las fases que conforman el *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software* [15]. Las Figuras 3 y 4 presentan la estructura del presente capítulo.

ESPECIFICACIÓN COLABORATIVA DE LOS MEU

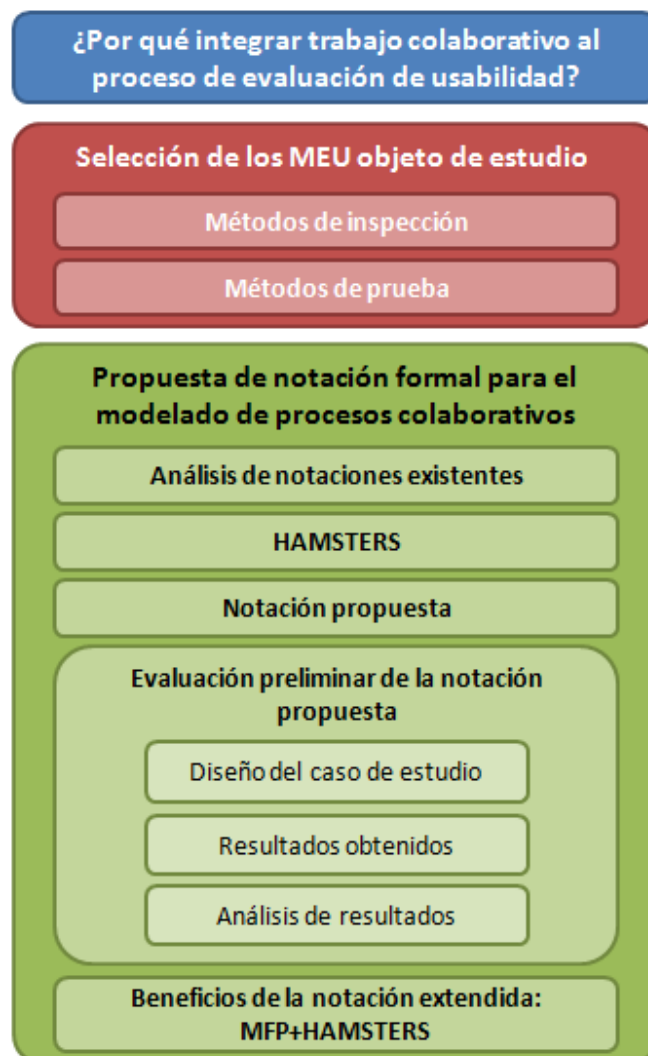


Figura 3. Estructura del capítulo 3. Parte 1.

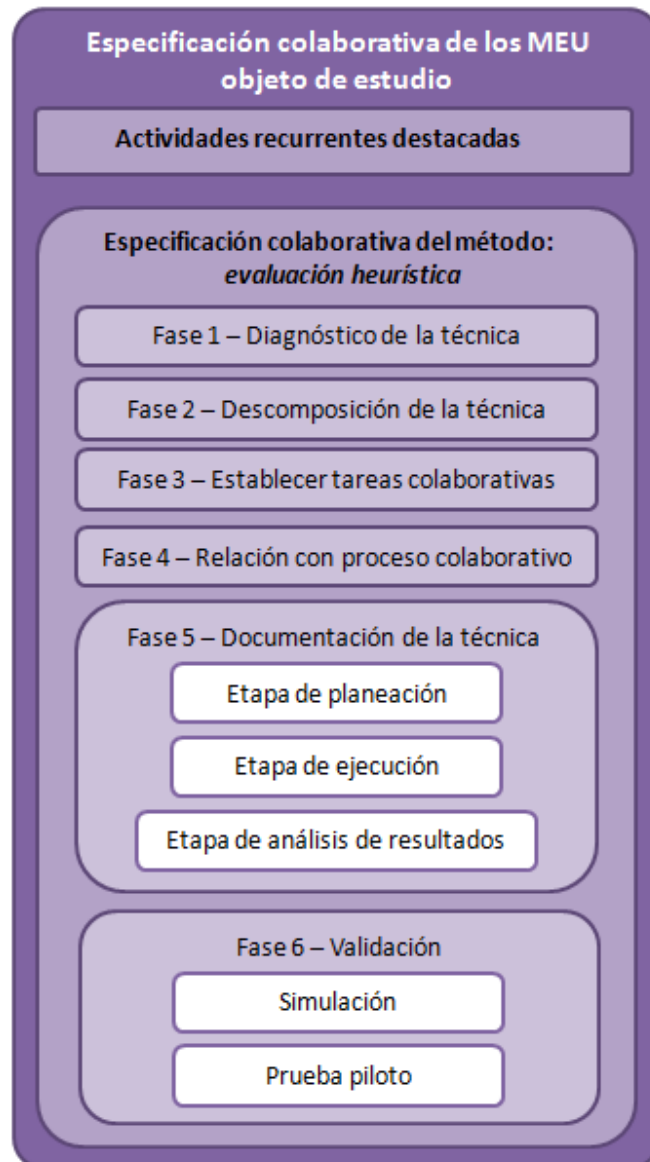


Figura 4. Estructura del capítulo 3. Parte 2.

3.1 ¿Por qué integrar trabajo colaborativo al proceso de evaluación de usabilidad?

Una buena parte de las actividades que conforman el proceso de evaluación de usabilidad dependen en gran medida del conocimiento y experiencia de los evaluadores. Ahora bien, en trabajos relacionados como [41][42][43] puede comprobarse que los MEU que integran procesos colaborativos permiten trabajar y compartir conocimiento y recursos a varias personas dispersas o no geográficamente durante el proceso de evaluación de usabilidad, esto con el fin de obtener sistemas más usables, aprovechando las ventajas del trabajo colaborativo. De esta manera, en el proceso de evaluación de usabilidad pueden participar personas con diferentes habilidades y conocimientos, involucrando usuarios potenciales y actuales, expertos en usabilidad, entre otros, formando equipos que trabajen conjuntamente

en la evaluación de sistemas software interactivos. Esto resulta muy adecuado ya que logra obtenerse información sobre los aspectos psicológicos del usuario, los aspectos sociales, temas del diseño del sistema, diseño gráfico, entre otros [7].

La investigación de la literatura muestra claramente la importancia de la colaboración entre los practicantes de la usabilidad durante el proceso de evaluación. La reunión de múltiples evaluadores en la identificación y análisis de problemas de usabilidad ha resultado útil para mejorar la rigurosidad y la fiabilidad de la identificación de problemas [41]. Esto es, el trabajo colaborativo entre evaluadores aumenta la probabilidad de que la gran mayoría de los problemas reales sean encontrados y se mejora la consistencia en el análisis de resultados. Por otro lado, la inclusión de trabajo colaborativo a procesos relacionados con la evaluación de usabilidad permite obtener algunos beneficios como [41]: obtener resultados más ricos en contenido, identificar una mayor cantidad de problemas de usabilidad, mejorar la fiabilidad y evitar que los resultados estén sesgados por la perspectiva de una sola persona, generar mejores propuestas de rediseño y el aseguramiento de la calidad. Además, la adición de múltiples perspectivas parece mejorar la eficacia de los métodos de inspección [41]. En ese sentido, vincular en la ejecución de los MEU a varias personas con diferentes conocimientos y experiencia resulta ser de gran utilidad.

Por otro lado, la literatura ofrece descripciones a un alto nivel sobre el proceso de evaluación de usabilidad. Estudios en los que se han llevado a cabo un conjunto de MEU (de inspección y de prueba), como es el caso de [44][45][46], por mencionar unos pocos, no proveen información detallada sobre cómo realizar el análisis de los datos recolectados. Con base en lo anterior, el presente trabajo intenta especificar de forma detallada e incluyendo aspectos de trabajo colaborativo, una serie de pasos a seguir (mediante el uso de *thinklets*) para llevar a cabo actividades que conforman MEU, actividades no sólo de la etapa de análisis de resultados, sino también de la etapa de preparación o planeación.

Dado que se tiene poca documentación sobre guías o lineamientos de cómo ejecutar evaluaciones colaborativas de usabilidad de sistemas interactivos, en la especificación colaborativa de un método de evaluación se cuenta con una secuencia de actividades bien definidas, entregables a generar, descripción de los participantes, especificación del proceso de comunicación entre los participantes de la evaluación, entre otra información relevante. Dicha especificación se espera que sea utilizada por personas (practicantes) responsables de estructurar el proceso de evaluación de usabilidad de sistemas interactivos.

3.2 Selección de los MEU objeto de estudio

Con base en las Tablas 3, 4 y 5 del Anexo B.3, es importante recalcar que no existe un "mejor método", todos tienen fortalezas y debilidades y están enfocados a evaluar aspectos específicos de usabilidad, por lo que la combinación de métodos es el procedimiento más adecuado para la evaluación de la usabilidad [27]. La selección de ellos depende de factores económicos y de tiempo, de simplicidad, de las fases en el ciclo de desarrollo, del tipo de resultados a obtener, de los recursos humanos que interpreten los resultados, de la cantidad de usuarios e implicados, entre otros aspectos.

Dado que existe un buen número de métodos para evaluar la usabilidad de sistemas interactivos, se hace necesario seleccionar un conjunto reducido de MEU objeto de estudio a considerar en esta investigación. Para ello, se ha tomado como referente el trabajo de tesis doctoral [10], en el cual ha sido realizada una *valoración total de utilidad* de los MEU

(asignando uno de los siguientes valores: *muy útil*, *útil* y *poco útil*) en base a los siguientes criterios: presencia de usuarios, necesidad de formación, aplicabilidad, cercanía a la Ingeniería de Software, aportación vs esfuerzo y representatividad. Así, fueron considerados los métodos de evaluación valorados como “muy útiles” y “útiles”. Adicionalmente, en el proceso de selección de los MEU objeto de estudio es considerada la información de las Tablas 3, 4 y 5 del Anexo B.3, así como también las experiencias y recomendaciones consignadas en trabajos de investigación relacionados como [1][15][47].

3.2.1 Selección de los métodos de inspección

Los métodos de inspección para la evaluación de usabilidad seleccionados como objeto de estudio en la presente investigación son los siguientes:

- *Evaluación heurística*: este método de evaluación de usabilidad fue seleccionado debido a que es el más popular de los métodos de inspección, básicamente por su facilidad y rapidez de realización [48]. Puede ser realizado en distintas etapas del desarrollo de un sistema software, incluso muy tempranamente, por lo que no es necesaria la disponibilidad de una versión terminada del sistema. Aunque el método es simple, los problemas no necesariamente son fáciles de encontrar, siempre dependiendo de la cantidad y nivel de los evaluadores. Es bastante eficiente, pero no es un método tan exhaustivo como para encontrar todos y cada uno de los problemas de usabilidad del sistema en evaluación [5].

La principal ventaja de la evaluación heurística es su bajo coste, que depende del número de evaluadores expertos que realicen el proceso. Un número mínimo de tres evaluadores permite detectar la mayoría de problemas de usabilidad del sistema. Por ello, los costes resultan mucho menores que en cualquier otro método de evaluación que requiera de la participación de usuarios finales para llevarlo a cabo [44]. Por otro lado, en cuanto al grado de aportación de usabilidad frente al esfuerzo de realizar el método, el valor es alto, puesto que el esfuerzo no es excesivo y las mejoras en usabilidad que pueden obtenerse son importantes.

Este método de evaluación tiene un alto nivel de aplicabilidad porque puede realizarse en un conjunto amplio de sistemas interactivos. Sin embargo, es importante mencionar que el conjunto seleccionado de heurísticas de usabilidad juega un papel clave en la realización de la evaluación heurística. Dependiendo de la experiencia de los evaluadores, podrían necesitar heurísticas más específicas, esto con el fin de minimizar la probabilidad de ignorar problemas específicos del dominio.

- *Recorrido cognitivo*: este método de inspección fue seleccionado puesto que se concentra en aspectos importantes como: (1) soporte del sistema para que el usuario cumpla su objetivo, (2) forma en que el sistema entrega indicadores para realizar cierta tarea y (3) retroalimentación del sistema. Este método de evaluación está centrado en un tema especialmente importante desde el punto de vista de la usabilidad [49]: la carga cognitiva que se impone al usuario durante el uso del sistema. También, puede ser realizado en diferentes etapas del desarrollo de un sistema software.

El recorrido cognitivo requiere tener un conocimiento amplio acerca de aspectos cognitivos por lo que la necesidad de formación de los evaluadores es alta. La

aportación de usabilidad por parte de este método es importante, pero el coste de aplicación es alto, por lo que el grado de aportación frente al esfuerzo es medio.

Los MEU: recorrido pluralista, inspección de estándares y análisis de acciones, no son considerados en esta investigación. Teniendo en cuenta la información de la Tabla 3 del Anexo B.3, el recorrido pluralista no es considerado porque resulta poco práctico para simular (en papel u otros materiales) las posibles acciones del usuario, además de que para su ejecución resulta complicado agrupar tantos participantes para una sola sesión. La inspección de estándares no es considerada porque requiere evaluadores con un amplio conocimiento de estándares (nivel de formación alto) y además, porque no tiene en cuenta la funcionalidad de las acciones a evaluar. El análisis de acciones no fue seleccionado principalmente porque requiere de expertos del más alto nivel, lo cual para la mayoría de las organizaciones resulta costoso conseguir.

3.2.2 Selección de los métodos de prueba

Los métodos de prueba para la evaluación de usabilidad seleccionados como objeto de estudio en la presente investigación son los siguientes:

- *Experimentos formales*: este método de prueba fue seleccionado porque es bastante utilizado y reconocido como método de medición de la usabilidad, además, porque su realización produce datos objetivos, primarios y cuantitativos (de distintos niveles de complejidad) que pueden ser estadísticamente analizados [1]. El grado de esfuerzo para realizar este método es alto por el coste que supone montar un laboratorio de usabilidad o una instalación con características similares, sin embargo, el aporte de usabilidad es significativo.

La necesidad de formación para realizar este método de prueba es media, puesto que debe tenerse cierto entrenamiento (pero no excesivo) para aprender la mecánica de las pruebas. Su aplicabilidad es alta, ya que puede ajustarse para evaluar diferentes sistemas software interactivos.

- *Cuestionarios y entrevistas*: estos métodos de indagación fueron seleccionados principalmente porque son simples y económicos, además de que proporcionan información (cualitativa y cuantitativa) relacionada a la satisfacción subjetiva del usuario y pueden emplearse en cualquier etapa del ciclo de desarrollo. Debido al carácter flexible de las entrevistas, puede indagarse más profundamente a los usuarios con el fin de obtener información complementaria.

Las entrevistas y cuestionarios tienen una aplicabilidad alta, puesto que son útiles para evaluar todo tipo de sistema. La necesidad de formación es importante sin resultar excesiva (nivel medio). El esfuerzo de elaboración, distribución y análisis de los resultados es considerable, sin embargo, permiten identificar un número importante de problemas de usabilidad, por lo que el grado de aportación frente al esfuerzo es medio.

- *Interacción constructiva*: este método de prueba fue seleccionado porque permite identificar un buen número de problemas de usabilidad, encuentra las razones por las que ocurren los problemas y puede ser aplicado en diferentes etapas del ciclo de desarrollo. Además, la interacción constructiva soluciona las desventajas del método:

pensando en voz alta, puesto que dos participantes expresan sus impresiones de forma más natural, conversando entre sí.

La interacción constructiva tiene una aplicabilidad alta, puesto que sirve para evaluar diferentes sistemas interactivos, sin embargo, la interacción de los participantes no representa el uso de algunos sistemas en condiciones normales. La necesidad de formación es importante sin resultar excesiva (nivel medio). La aportación de usabilidad por parte de este método es importante, pero el coste de aplicación es alto (dado que necesita un número doble de participantes), por lo que el grado de aportación frente al esfuerzo es medio.

- *Método del conductor*: este método de prueba fue seleccionado porque está enfocado en descubrir las necesidades de información de los usuarios en el sistema, detecta las razones por las que ocurren los problemas y puede emplearse en cualquier etapa del ciclo de desarrollo. Este método de evaluación es similar a la interacción constructiva, la diferencia está en que no interactúan dos usuarios sino un evaluador y un usuario, por lo que también se da una conversación natural donde las intervenciones del evaluador son controladas.

El método del conductor tiene alta aplicabilidad, ya que puede emplearse para evaluar diferentes sistemas interactivos, sin embargo, se pierde el grado de representación de la realidad ya que los sistemas evaluados no son comúnmente usados en compañía de una persona guía. La necesidad de formación es importante sin resultar excesiva (nivel medio). La aportación de usabilidad por parte de este método es importante, y el coste de aplicación es medio, por lo que el grado de aportación frente al esfuerzo es medio.

Los MEU: pensando en voz alta, grabación del uso, medida de prestaciones y test retrospectivo, no son considerados en esta investigación. Con base en la información de la Tabla 3 del Anexo B.3, el pensamiento en voz alta no es considerado porque interfiere en la conducta normal del usuario, lo cual influye en la interacción con el sistema. La grabación del uso no es considerada porque su realización requiere un nivel de formación alto por parte de los evaluadores, el esfuerzo para establecer el equipamiento es alto, y además, es especialmente indicado sólo para analizar sitios web (nivel de aplicabilidad bajo). La medida de prestaciones no fue seleccionada a causa de que no asegura que lo medido está relacionado con la meta de usabilidad que se investiga, no obtiene información subjetiva (opiniones, actitudes, satisfacción), además, el ambiente utilizado no es natural para el usuario, por lo que puede sesgar la actuación del mismo. El test retrospectivo no fue seleccionado principalmente porque demora como mínimo el doble de tiempo necesario que con cualquier otro método.

Una vez seleccionados los MEU objeto de estudio, el paso a seguir consiste en llevar a cabo las fases que conforman el *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software*, sin embargo, antes de obtener la especificación colaborativa de dichos métodos es presentada una propuesta de notación que contribuye a la representación de los procesos. Esta notación es presentada en la siguiente sección.

3.3 Propuesta de notación para el modelado de procesos colaborativos extendiendo la notación

HAMSTERS

El *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software* [15] está basado en la *Metodología para el desarrollo de procesos colaborativos* [8], la cual provee un *Modelo de Facilitación del Proceso* (MFP) para representar el flujo del proceso y otros elementos de las actividades colaborativas.

Centrándonos en el MFP, este presenta una serie de carencias que pueden mejorarse con el propósito de enriquecer la representación gráfica del modelo. Para tal fin, han sido estudiadas una serie de notaciones formales entre las cuales HAMSTERS (*Human-centered Assessment and Modeling to Support Task Engineering for Resilient Systems*) [50] se destaca por ofrecer un conjunto de elementos apropiados para complementar la representación gráfica del MFP. Así, esta sección presenta una propuesta de notación para el modelado de procesos colaborativos extendiendo la notación HAMSTERS, la cual es descrita más adelante.

3.3.1 Análisis de notaciones existentes

Teniendo en cuenta que el MFP [8] presenta el flujo del proceso colaborativo diseñado y los elementos: identificador, *thinklet*, *patrón de colaboración* y nombre de las actividades (en un rectángulo dividido en 4 secciones); se ha identificado la necesidad de complementar dicho modelo con elementos que contribuyan principalmente a la comprensión, entendimiento y uso futuro del mismo. Para tal fin, se han estudiado una serie de notaciones formales como se muestra más adelante.

En el MFP, propuesto por investigadores en Ingeniería de Colaboración, los autores de la presente investigación (con base en la experiencia de estudios previos [47][51]) han identificado una serie de carencias que pueden mejorarse con el propósito de complementar la representación gráfica del modelo, esas son:

1. Representación secuencial del flujo del proceso. Hay actividades que pueden realizarse de forma independiente, concurrente, iterativa, opcional, con paso de información, entre otras, lo cual no es posible representar con los elementos que ofrece un MFP.
2. Dificultad para identificar jerarquías de actividades. En el caso de las actividades que han sido definidas como colaborativas y tienen asociados varios (2 o más) *patrones de colaboración*, es compleja la identificación de jerarquías entre las subactividades correspondientes a los *thinklets* asociados (a cada patrón de colaboración).
3. Ausencia de elementos que indican la entrada y/o salida de información. Los elementos que ofrece el MFP son insuficientes para representar la existencia de entradas (recursos necesarios) y/o salidas (entregables a generar) de cada actividad que conforma el proceso. Para una persona que desea ejecutar un proceso colaborativo resulta adecuado presentar elementos que señalen la entrada y/o salida de información.
4. No es posible identificar los participantes que realizan las actividades.
5. Falta de información relacionada a las actividades colaborativas. Las actividades que se han definido como colaborativas tienen asociado un patrón de colaboración y un *thinklet* (en los campos respectivos del rectángulo), sin embargo, no se presenta información suficiente relacionada a las sub-actividades (o pasos) que conforman el *thinklet*.

Un MFP consiste básicamente en la representación de un conjunto de actividades o tareas para alcanzar un objetivo, las cuales deben realizarse de acuerdo al protocolo concreto establecido. Con base en lo anterior, al observar un MFP es posible notar la relación existente con las notaciones utilizadas en el campo del Análisis de Tareas (TA, por sus siglas en inglés *Task Analysis*) [52].

En la disciplina HCI e Ingeniería de Software en general, existe un amplio conocimiento soportando que el TA es útil para los desarrolladores y/o diseñadores de sistemas interactivos. Por este motivo, se considera apropiado analizar si las notaciones de TA existentes son de utilidad para este trabajo. Con este propósito, ha sido analizado un conjunto de notaciones existentes para el modelado de tareas, tales como: CTT (*Concur Task Trees*) [53], HTA (*Hierarchical Task Analysis*) [54][55] y HAMSTERS [50]. Estas notaciones permiten modelar las tareas que un usuario puede llevar a cabo en un sistema interactivo, además cada una ofrece un conjunto particular de elementos y es especialmente útil para un tipo específico de sistema. La Tabla 2 presenta (en la columna de la izquierda) el conjunto de requisitos necesarios para poder representar el MFP con base en el TA y el soporte que cada una de las notaciones analizadas (en las columnas segunda, tercera y cuarta) puede proporcionar. Esta tabla fue elaborada con base en la experiencia que se tiene usando la *Metodología para el Desarrollo de Procesos Colaborativos*.

Tabla 2. Comparación del soporte que cada notación puede ofrecer a los requisitos necesarios para representar el MFP.

Requerimiento	CTT	HTA	HAMSTERS
Realización de actividades de forma concurrente.	X		X
Realización de actividades de forma iterativa.	X	X	X
Realización de actividades de forma opcional.	X	X	X
Representación de entrada y/o salida de información en las actividades.			X
Representación de jerarquías de actividades.	X	X	X
Representación de los participantes/roles que realizan las actividades.			
Representación de actividades físicas.	X		X
Representación de actividades cognitivas (de análisis, toma de decisiones).	X		X
Representación de entrada y/o de salida de datos del sistema.	X	X	X
Representación de actividades colaborativas/cooperativas.	X		X
Representación de compartición de información.			
Representación de actividades cognitivas colaborativas (de análisis, toma de decisiones).			
Representación de entrada colaborativa de datos al sistema.			

Con base en la información de la Tabla 2 y, considerando las falencias del MFP antes mencionadas, se ha optado por utilizar la notación HAMSTERS (basada en la notación CTT) ya que ofrece un conjunto de elementos apropiados para complementar la representación gráfica de la información en el MFP:

- Permite mostrar gráficamente las relaciones (conurrencia, iteración, entre otras) existentes entre las actividades para alcanzar un objetivo.
- Notación fácil de usar y aplicable para representar actividades en diferentes sistemas software interactivos.
- Genera una representación gráfica en forma de árbol permitiendo una descomposición jerárquica de las actividades.
- Es posible representar actividades colaborativas/cooperativas.

Adicionalmente, a diferencia de CCT, HAMSTERS incluye extensiones, tales como:

actividades interactivas con representación de entrada y/o salida de información y precondiciones para la ejecución de ciertas actividades.

Según lo anterior, se considera que HAMSTERS es la notación que resulta más apropiada para complementar la representación gráfica del MFP. Aún así, la información presentada en la Tabla 2 deja entrever que dicha notación todavía debe ser complementada, de tal forma que pueda ser posible representar información detallada sobre las actividades colaborativas y los participantes/roles que realizan dichas actividades. La siguiente sección describe en mayor detalle la notación HAMSTERS.

3.3.2 HAMSTERS

HAMSTERS [50] es una notación para el modelado de tareas propuesta por investigadores del Instituto de Investigación en Informática de Toulouse de la Universidad Paul Sabatier (Toulouse, Francia). Ha sido inspirada en notaciones existentes, especialmente en la notación CTT [56], y ha sido diseñada para mantener compatibilidad con CTT, puesto que los modelos son jerárquicos y representados gráficamente mediante operadores entre tareas [50]. HAMSTERS incluye otros elementos como [50]: actividades interactivas más detalladas, precondiciones asociadas a la ejecución de tareas, flujo de datos a través de las tareas del modelo, entre otros. Además, esta notación proporciona extensiones adicionales requeridas para la estructuración de modelos. En [50] se presentan los tipos de tareas y la representación de entradas y/o salidas de información definida en HAMSTERS.

Como en CTT [56], cada tarea particular del modelo puede ser opcional, iterativa o ambos. La propiedad *opcional* indica que la tarea no requiere ser ejecutada para la meta a alcanzar. La propiedad *iterativa* indica que una tarea puede ser ejecutada una o más veces. En la notación se utilizan además, una serie de operadores para facilitar la descripción de las relaciones temporales existentes entre las tareas. Adoptando la notación CTT, en [57] se presentan los operadores temporales que utiliza HAMSTERS. En la siguiente sección se presenta la propuesta de extensión de la notación para utilizar en el MFP.

3.3.3 Propuesta: extensión de la notación HAMSTERS para utilizar en el MFP

Con base en la información antes presentada (secciones 3.3.1 y 3.3.2), y teniendo en cuenta las carencias que presenta el MFP, todos los elementos de la notación HAMSTERS (como son: tipos y propiedades de tareas, relaciones entre tareas, entrada/salida de información) serán considerados para complementar la representación gráfica del MFP. Las extensiones realizadas a HAMSTERS se presentan en los siguientes puntos. Los ejemplos de las extensiones de la notación corresponden a la especificación colaborativa del método evaluación heurística.

- Representación de una actividad. La imagen que representa una tarea/actividad (en HAMSTERS) fue reemplazada por el rectángulo (dividido en 4 secciones) utilizado en el MFP tradicionalmente definido. De esta manera, el nombre de la actividad está acompañado del identificador, *thinklet* y *patrón de colaboración*. Adicionalmente, en la actividad se indican los participantes y si existe entrada y/o salida de información (ver Figura 5).

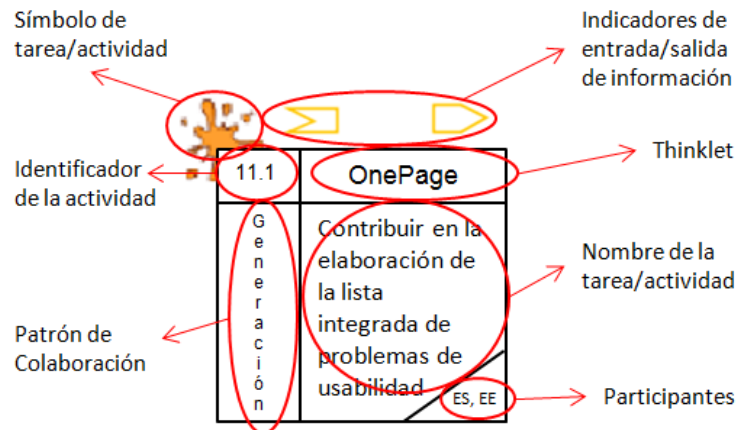


Figura 5. Representación de una actividad en el MFP, usando elementos de HAMSTERS.

- Identificación de los participantes en las actividades. Mediante la notación HAMSTERS no es posible representar los roles de los participantes de una tarea/actividad, por tal razón, en la esquina inferior derecha del rectángulo (que representa la actividad) se indica el símbolo correspondiente a los participantes de la actividad (ver Figura 5). Para el caso de la especificación colaborativa de MEU, la Tabla 3 presenta los símbolos utilizados en el MFP correspondientes a los participantes de los métodos de evaluación.

Tabla 3. Símbolos de los participantes.

Participante	Símbolo
Evaluador Supervisor	ES
Evaluadores Expertos	EE
Usuario (s)	U
Representante de la Organización	RO

- Relaciones entre tareas/actividades y entrada/salida de información. Las Figuras 6 y 7 muestran la representación de la actividad: *Creación de una lista integrada de problemas*, en el MFP tradicional y usando solamente elementos de HAMSTERS, respectivamente. En el MFP tradicional (ver Figura 6) las actividades están conectadas por flechas que indican la dirección del flujo del proceso, por lo cual la representación es secuencial. Usando sólo elementos de HAMSTERS (ver Figura 7) es posible identificar diferentes tipos de relaciones entre las actividades y si existen entradas/salidas de información, pero no se representan elementos como: *patrones de colaboración*, *thinklets* e información acerca de actividades colaborativas.

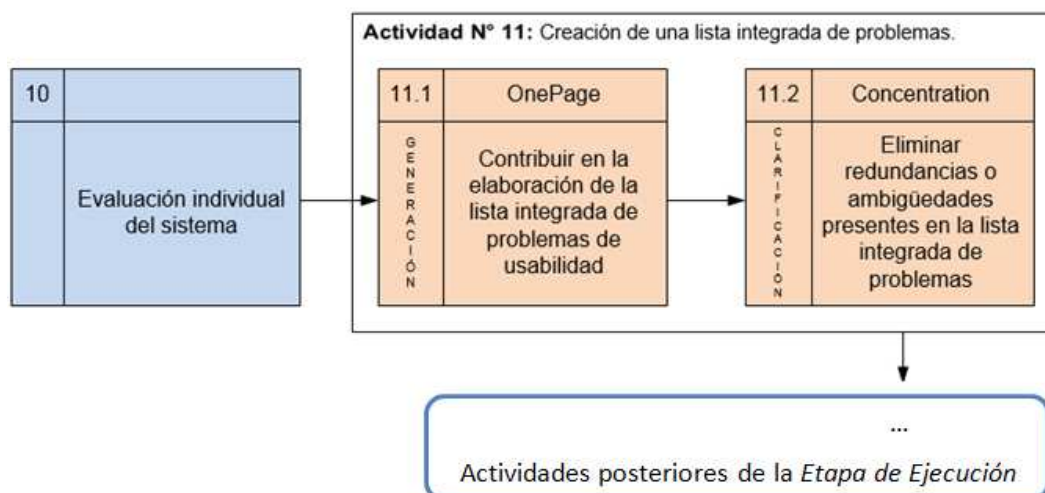


Figura 6. Representación de las actividades en el MFP tradicionalmente definido.

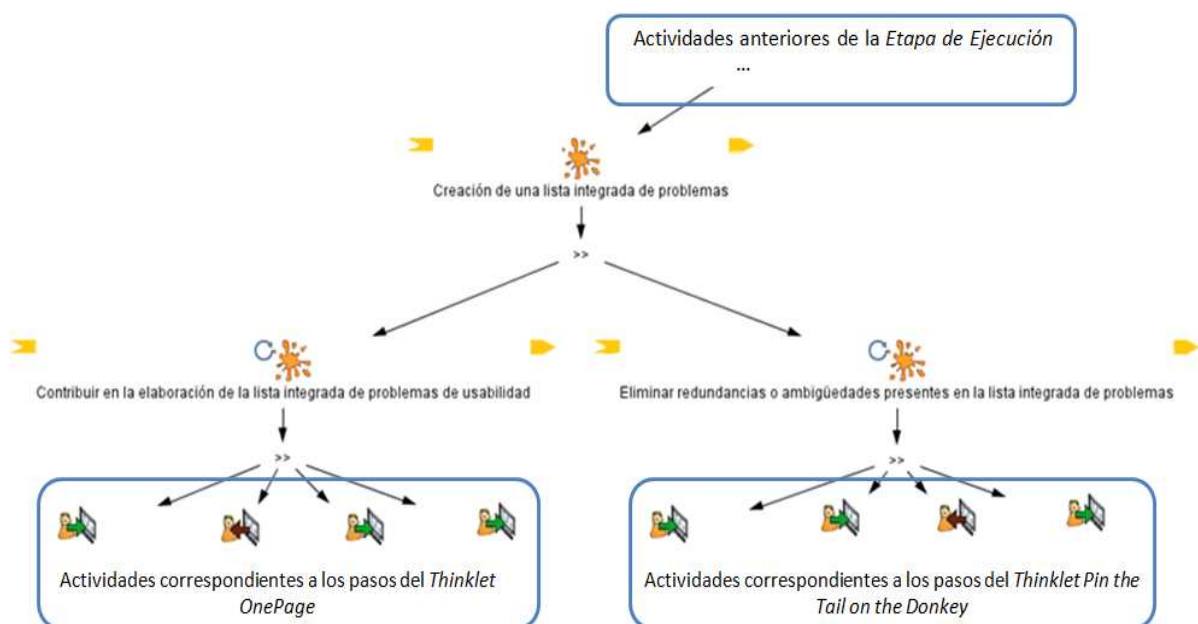






Figura 7. Representación de las actividades usando únicamente elementos de HAMSTERS.

- Actividades/tareas colaborativas detalladas. Las notaciones CTT y HAMSTERS permiten representar actividades colaborativas/cooperativas. Para este caso, dicha representación tiene un alto nivel de generalidad, ya que no es posible representar información detallada a cerca de este tipo de actividades, como por ejemplo la toma de decisiones (consenso) en grupo, análisis de la información, entre otra. Con base en lo anterior, la notación se extiende mediante una serie de actividades/tareas que complementan la información relacionada a las actividades colaborativas. La Tabla 4 presenta las imágenes a utilizar para representar las subactividades que conforman una actividad colaborativa, dichas subactividades corresponden a los pasos definidos en el *thinklet* utilizado.

Tabla 4. Representación gráfica de actividades colaborativas.

Actividad/Tarea	Imagen
Compartir información.	
Actividad cognitiva colaborativa (análisis).	
Actividad cognitiva colaborativa (toma de decisión).	
Entrada colaborativa de datos al sistema.	

Teniendo en cuenta la información antes presentada, la Figura 8 presenta una porción del MFP extendiendo la notación HAMSTERS. En esta pueden apreciarse las extensiones realizadas a HAMSTERS con el propósito de satisfacer las carencias identificadas en el MFP (ver sección 3.3.1).

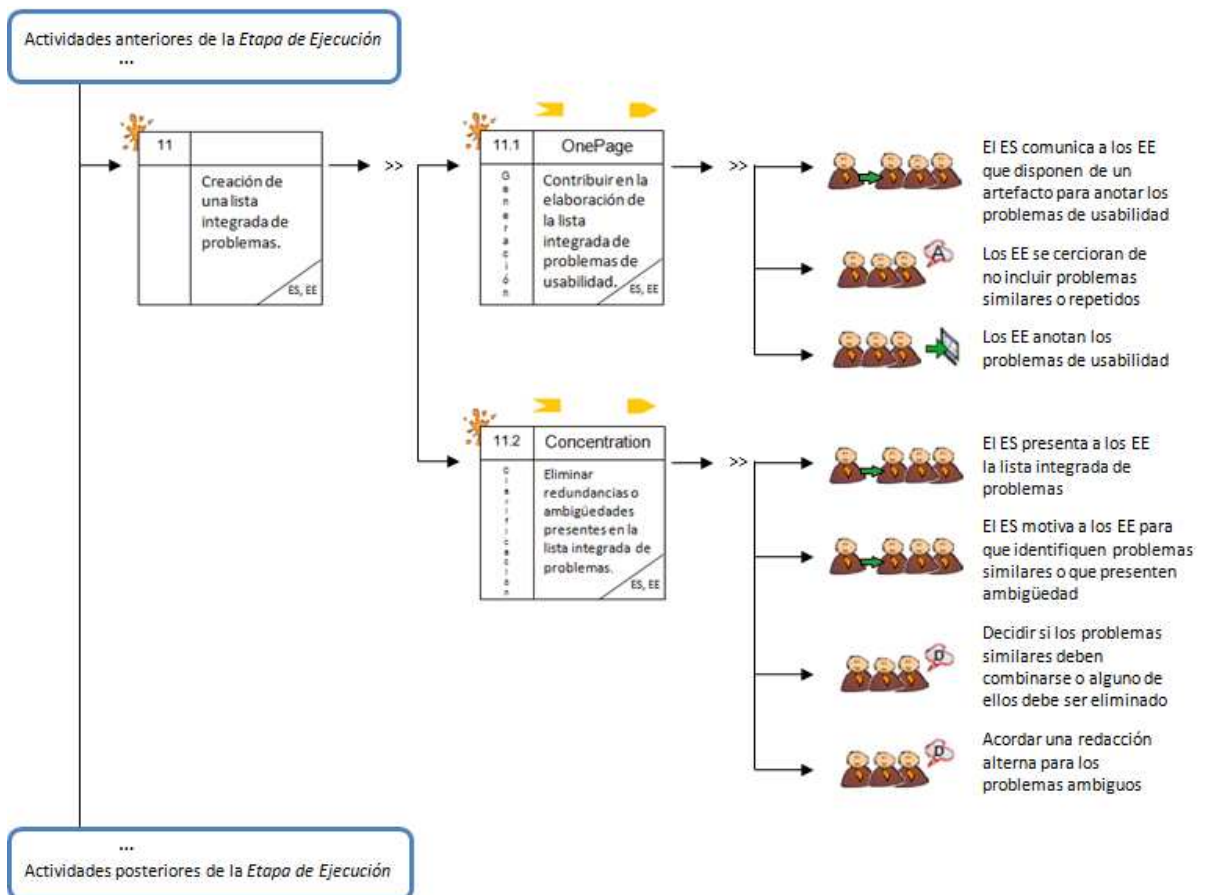


Figura 8. MFP extendiendo la notación HAMSTERS.

3.3.4 Evaluación preliminar de la notación propuesta

Con el propósito de evaluar si los elementos de HAMSTERS incluidos al MFP son apropiados y satisfacen las carencias identificadas (ver sección 3.3.1), fue realizado el caso de estudio presentado a continuación.

3.3.4.1 Diseño del caso de estudio

Considerando los elementos incluidos al MFP a partir de la notación HAMSTERS, fue realizado un caso de estudio mediante el cual se obtuvo información acerca de un MFP tradicional vs uno que incluye elementos de HAMSTERS, esto con el fin de realizar algunas comparaciones basadas en una serie de características descritas más adelante.

Para obtener información acerca de los MFP (tradicional vs uno que incluye elementos de HAMSTERS), fue realizada una encuesta a un grupo de 11 expertos con el siguiente perfil: experiencia y conocimiento en el área de Ingeniería de Colaboración y experiencia en la ejecución de por lo menos un método de evaluación de usabilidad. La encuesta fue planteada con el objetivo de determinar si los elementos de HAMSTERS incluidos en el MFP son apropiados y fáciles de comprender, teniendo en cuenta las consideraciones de personas que trabajan en áreas de investigación relacionadas.

La encuesta fue elaborada utilizando el sistema SUS (*System Usability Scale*) [58], de tal manera que cada pregunta tiene 5 opciones de respuesta. El significado de las opciones indica que la nota mínima (1) corresponde a una evaluación que reprueba o califica de mala manera la característica que se está evaluando, mientras que la nota máxima (5) corresponde a una aprobación o que la característica está siendo calificada positivamente.

Las características sobre las cuales se obtuvo información de los MFP presentados en la encuesta son las siguientes:

- *Compleitud*: se refiere a si el MFP provee la información necesaria para llevar a cabo el proceso colaborativo especificado.
- *Expresividad/Representatividad*: se refiere a si la información presentada gráficamente en el MFP es lo suficientemente explícita para su comprensión.
- *Facilidad de usar*: se refiere a la rapidez con que el MFP puede ser utilizado para llevar a cabo el proceso especificado de forma colaborativa sin mayores problemas.
- *Facilidad de entender*: refleja la rapidez con que se entiende la información del MFP y pueda ser llevado a cabo el proceso colaborativo especificado.
- *Cantidad de información por actividad*: se refiere a si la cantidad de información presentada por actividad es suficiente para representar los aspectos primordiales de cada una de ellas.

Así, a cada participante se le solicitó valorar dichas características en una escala de 1 a 5 respecto a un MFP tradicional y uno que incluye elementos de HAMSTERS. Los MFP objeto de estudio corresponden al método de evaluación de usabilidad: *evaluación heurística*. Las valoraciones de los participantes que diligenciaron la encuesta fueron el insumo para el análisis de resultados presentado más adelante.

3.3.4.2 Resultados obtenidos

Una vez los participantes diligenciaron la encuesta, se obtuvieron los siguientes resultados para cada uno de los MFP.

MFP TRADICIONAL

Descripción: En este MFP se representa cada actividad como un rectángulo que se divide en cinco campos: (i) En la parte superior izquierda se indica el número de secuencia. (ii) La zona central y de mayor tamaño contiene el nombre descriptivo de la actividad. (iii) El rectángulo de la izquierda presenta el nombre del *patrón de colaboración* asociado a la actividad. (iv) El nombre del *thinklet* se ubica en el campo superior y (v) en la esquina superior derecha se ubica el tiempo (en minutos) estimado para realizar la actividad. Las flechas indican la dirección del desarrollo del proceso. Las actividades que no tienen asociado un *patrón de colaboración* y *thinklet*, es porque no se realizan de forma colaborativa. Las actividades con una tonalidad más oscura representan actividades generales, mientras que aquellas con una tonalidad más clara representan las subactividades que conforman cada actividad general. En este caso, las actividades que conforman la evaluación heurística se han dividido en tres etapas: planeación, ejecución y análisis de resultados. Este MFP está enfocado en la evaluación heurística de aplicaciones de televisión digital interactiva. La Figura 9 presenta una porción del MFP correspondiente a la etapa de ejecución de la evaluación heurística.

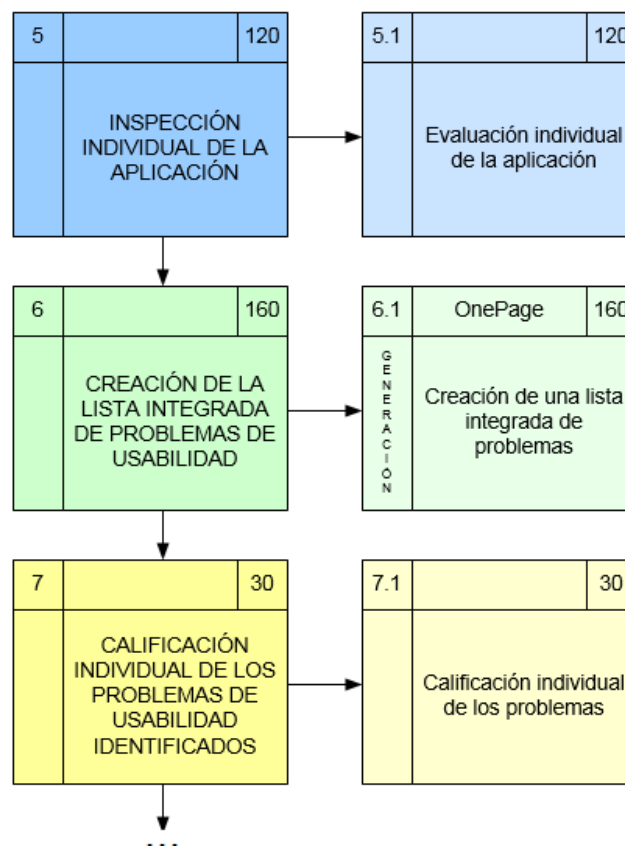


Figura 9. Porción del MFP tradicional – *Etapa de ejecución* de la evaluación heurística.

Los resultados obtenidos para este MFP, según las valoraciones de cada una de las características, son presentados en la Tabla 5.

Tabla 5. Resultados obtenidos para el MFP tradicional.

Participante	Características				
	Complejidad	Expresividad/ Representatividad	Facilidad de usar	Facilidad de entender	Cantidad de información por actividad
1	2	2	3	3	3
2	4	3	3	3	4
3	4	2	3	4	3
4	4	5	5	5	3
5	4	4	2	2	2
6	4	3	3	4	3
7	4	5	4	5	3
8	4	4	5	5	4
9	5	5	4	5	5
10	2	2	4	4	4
11	4	4	3	5	4
Promedio	3,72	3,54	3,54	4,09	3,45
σ	0,90	1,21	0,93	1,04	0,82

MFP QUE INCLUYE ELEMENTOS DE HAMSTERS

Descripción: En este MFP se representa cada actividad como un rectángulo que se divide en cinco campos: (i) En la parte superior izquierda se indica el número de secuencia. (ii) La zona central y de mayor tamaño contiene el nombre descriptivo de la actividad. (iii) El rectángulo de la izquierda presenta el nombre del *patrón de colaboración* asociado a la actividad. (iv) El nombre del *thinklet* se ubica en el campo superior y (v) en la esquina inferior derecha se indican los participantes de la actividad (evaluador supervisor – ES, evaluadores expertos – EE, usuarios – U, representante de la organización – RO). Las flechas indican la dirección del desarrollo del proceso. Las actividades que no tienen asociado un *patrón de colaboración* y *thinklet*, es porque no se realizan de forma colaborativa. Este MFP presenta información sobre entradas/salidas de las actividades, tipos de relaciones entre actividades (independientes, concurrentes, entre otras) mediante operadores de la notación CTT y actividades colaborativas detalladas. En este caso, las actividades que conforman la evaluación heurística se han dividido en tres etapas: planeación, ejecución y análisis de resultados. Este MFP está enfocado en la evaluación heurística de sistemas software interactivos. La Figura 10 presenta una porción del MFP correspondiente a la etapa de ejecución de la evaluación heurística.

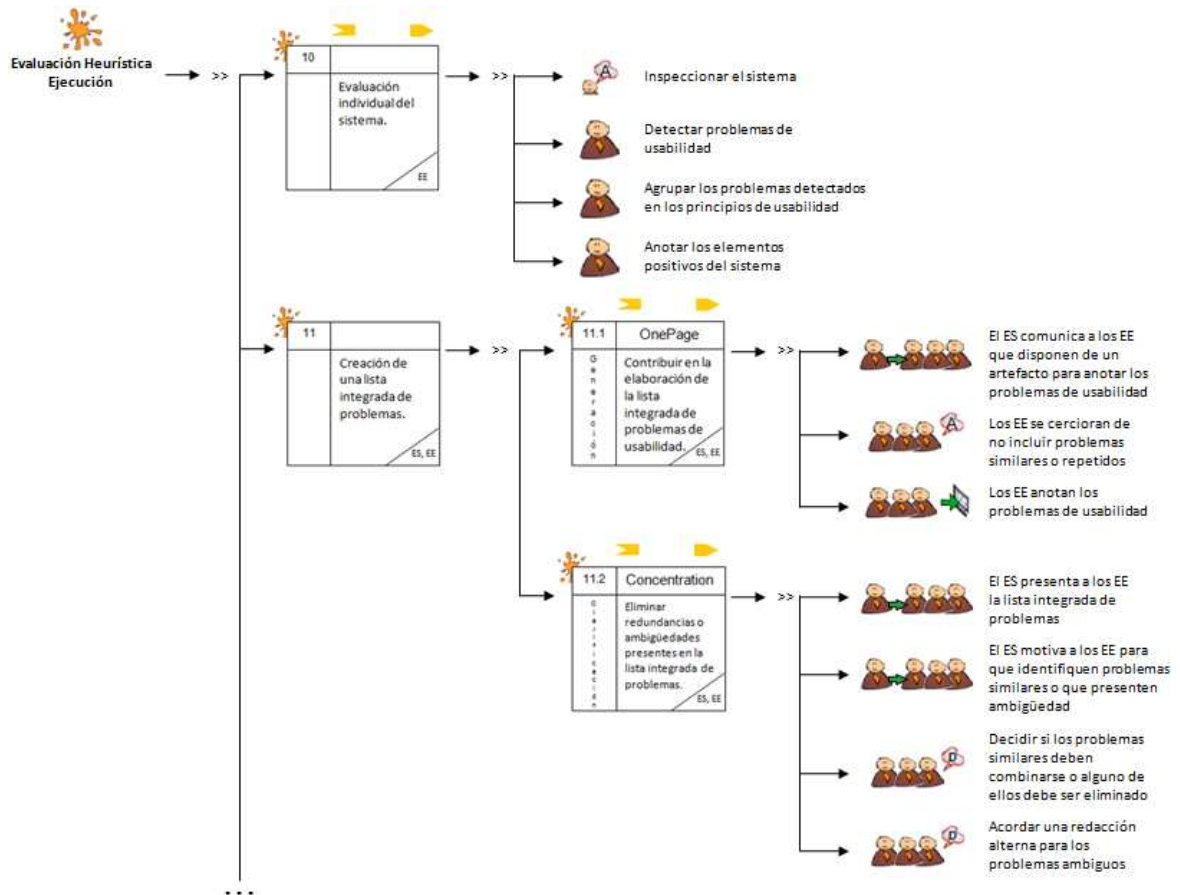


Figura 10. Porción del MFP que incluye elementos de HAMSTERS – Etapa de ejecución de la evaluación heurística.

Los resultados obtenidos para este MFP, según las valoraciones de cada una de las características, son presentados en la Tabla 6.

Tabla 6. Resultados obtenidos para el MFP que incluye elementos de HAMSTERS.

Participante	Características				
	Complejidad	Expresividad/ Representatividad	Facilidad de usar	Facilidad de entender	Cantidad de información por actividad
1	4	4	4	4	5
2	4	3	3	3	4
3	5	5	4	4	5
4	3	4	5	5	4
5	4	4	4	3	4
6	4	3	2	2	2
7	5	5	4	5	5
8	4	5	5	5	5
9	5	4	5	5	5
10	5	5	5	5	5
11	4	4	4	5	5
Promedio	4,27	4,18	4,09	4,18	4,45
σ	0,64	0,75	0,94	1,07	0,93

3.3.4.3 Análisis de resultados

Una vez recopilados y procesados los resultados de las encuestas (incluyendo promedios y desviación estándar), fue posible realizar algunas comparaciones entre el MFP tradicional y el que incluye elementos de HAMSTERS, como se presenta a continuación.

En cuanto a la interpretación de los resultados presentados en la Tabla 5, la calificación promedio más alta corresponde a la característica *Facilidad de entender*, mientras que la más baja corresponde a la *Cantidad de información por actividad*. Esto sugiere que la simplicidad del MFP tradicional contribuye a que las personas entiendan rápidamente la información presentada en el modelo. Sin embargo, la misma simplicidad del modelo afecta su representatividad, debido a que la cantidad de información presentada no es suficiente para visualizar aspectos primordiales de las actividades.

De igual forma, con base en los resultados de la Tabla 6, la calificación promedio más alta corresponde a la característica *Cantidad de información por actividad*, mientras que la más baja corresponde a la *Facilidad de usar*. Lo anterior indica que los elementos de HAMSTERS adicionados a las actividades (como son: participantes, indicadores de entrada y/o salida de información y actividades colaborativas detalladas) contribuyen a la representación de información primordial que requiere una persona encargada de llevar a cabo un proceso colaborativo. Por otro lado, el bajo promedio en la calificación de la *Facilidad de uso* se pudo haber presentado porque los tipos de relaciones entre actividades no se indican explícitamente (solamente es indicado el símbolo según la notación CTT). Con lo cual, si la persona encargada de ejecutar el proceso colaborativo no conoce los tipos de relación entre actividades (definidos en CTT), demorará más tiempo en aprender a utilizar el MFP. Con base en lo anterior, si en el modelo se representara el símbolo y significado del mismo, lo esperado sería que aumente la rapidez con que el MFP puede ser utilizado para llevar a cabo el proceso especificado de forma colaborativa sin mayores inconvenientes.

En las Tablas 5 y 6 se destacan los valores de desviación estándar más altos (mayores que 1), indicando que las opiniones de los participantes han sido dispersas, teniendo distintos puntos de vista sobre la misma característica. Respecto al MFP tradicional las características en las cuales se presentó una mayor desviación estándar son: *Expresividad/Representatividad* y *Facilidad de entender*, mientras que en el MFP que incluye elementos de HAMSTERS sólo una característica presentó un valor mayor a 1, la *Facilidad de entender*. La *Expresividad/Representatividad*, en el caso del MFP que incluye elementos de HAMSTERS, mejoró la calificación promedio obtenida, además hubo mayor consenso entre los expertos pues la desviación estándar es menor a 1. Adicionalmente, los resultados de las tablas mencionadas dejan entrever que en los dos MFP evaluados la *Facilidad de entender* es una característica en la cual no se alcanzó un consenso entre los participantes que diligenciaron la encuesta, por lo cual la mejora de este aspecto se considera como trabajo futuro a realizar.

En general, el promedio de calificaciones para las 5 características evaluadas en el MFP que incluye elementos de HAMSTERS es más alto respecto a las calificaciones obtenidas en el MFP tradicional. Esto indica que la inclusión de elementos de la notación HAMSTERS a un MFP permite obtener una serie de beneficios que contribuyen al momento de diseñar y/o ejecutar un proceso colaborativo determinado. En la siguiente sección se muestran los principales beneficios de la notación extendida.

3.3.5 Beneficios de la notación extendida: MFP+HAMSTERS

La inclusión de elementos de la notación HAMSTERS a un MFP permite obtener una serie de beneficios que contribuyen al momento de diseñar y/o ejecutar un proceso colaborativo determinado. Estos serían algunos beneficios de la notación extendida [59]:

- Mayor expresividad/representatividad: con el uso de los elementos que ofrece la notación HAMSTERS se tiene la capacidad de proveer mucha más información en el MFP. Esto es, en el MFP tradicionalmente definido era posible observar el nombre, *patrón de colaboración*, *thinklet* e identificador de las actividades, así como también la relación entre ellas (mediante flechas). Ahora, con el uso de HAMSTERS será posible observar toda la información antes mencionada, con los siguientes elementos adicionales: actividades detalladas de los participantes, diferentes tipos de relaciones entre actividades, indicadores de entrada y/o salida de información y participantes. Así, para una persona interesada en llevar a cabo un proceso colaborativo determinado, la información presentada en el MFP resulta más fácil de comprender.
- Flexibilidad en la representación del MFP: es posible representar actividades concurrentes e iterativas que conforman un proceso colaborativo determinado. También, mediante los conectores que ofrece la notación, el flujo de las actividades no es representado sólo de forma secuencial, como se hacía en el MFP tradicional.
- Sincronización de actividades: dado que la notación HAMSTERS está basada en CTT, es posible sincronizar dos actividades. Esto es un hecho positivo, por ejemplo cuando se necesita intercambiar información, debido a que la información de salida de una actividad es la información de entrada de la actividad siguiente.
- Descomposición jerárquica de las actividades: permite comprender mucho mejor la secuencia de subactividades que conlleva la realización de una actividad general que hace parte de un proceso especificado de forma colaborativa.
- La notación gráfica de las actividades (en forma de árbol) permite a una persona visualizar de mejor manera la descomposición jerárquica de las actividades.

3.4 Especificación colaborativa de los MEU objeto de estudio

Esta sección presenta el procedimiento y resultados obtenidos en el desarrollo de las fases que componen el *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software*, esto con el fin de obtener la especificación colaborativa de los MEU seleccionados como objeto de estudio. Inicialmente se lista una serie de actividades recurrentes destacadas, luego, es presentado el proceso y resultados obtenidos en las fases que componen el modelo únicamente para el método: *evaluación heurística*. Por restricción de extensión del documento, la especificación colaborativa de los demás MEU objeto de estudio es presentada en los Anexos F a K.

3.4.1 Actividades recurrentes destacadas

Como fue mencionado en la sección 2.1.2.1.2, los MEU siguen un mismo proceso en donde son definidas una serie de actividades para la planeación, ejecución y análisis de resultados de los mismos. Es así como en los métodos de evaluación se repiten algunas actividades definidas en la etapa de análisis de resultados. Estas actividades recurrentes destacadas están relacionadas principalmente con la elaboración de la lista de problemas de usabilidad identificados en el sistema, el análisis e interpretación de la información recolectada y la sugerencia de recomendaciones de diseño para solucionar los problemas detectados.

Dichas actividades son realizadas repetidamente en cada método de evaluación y pueden completarse usando un diseño de proceso similar cada vez que son ejecutadas. Así, en este punto resulta de utilidad la Ingeniería de Colaboración la cual está enfocada en el diseño de prácticas de trabajo colaborativo para tareas recurrentes, como las antes mencionadas.

Con base en lo anterior, el Anexo D presenta el proceso colaborativo definido para completar las actividades: (1) identificar problemas de usabilidad, (2) análisis e interpretación de la información recolectada y (3) realizar recomendaciones para solucionar los problemas de usabilidad. El proceso colaborativo definido para llevar a cabo dichas actividades será reutilizado en la especificación colaborativa de los MEU objeto de estudio.

3.4.2 Especificación colaborativa del método: *evaluación heurística*

3.4.2.1 Fase 1 – Diagnóstico de la técnica

En esta fase es descrito el MEU, incluyendo información como: entregables, participantes y demás características relevantes del método.

Descripción

Ver descripción de la evaluación heurística en el Anexo B.1.

Entregables

- Lista de problemas de usabilidad identificados en el prototipo o sistema evaluado.
- Priorización de los problemas de usabilidad según su criticidad y severidad.
- Contribuciones relacionadas al análisis e interpretación de los resultados obtenidos.
- Lista de recomendaciones de diseño que dan solución a los problemas de usabilidad.
- Lista de elementos positivos del sistema evaluado.

Requerimientos

- Cualquier tipo de prototipo (funcional o no funcional) o sistema final.

Etapas del proceso de desarrollo

Puede ser realizado en todas las etapas del desarrollo de un sistema, incluso muy tempranamente, por lo que no es necesaria la disponibilidad de una versión final de este. Sin embargo, se recomienda que el sistema a evaluar tenga cierto grado de avance o funcionalidad, para que los expertos puedan evaluarlo de una mejor manera y se puedan obtener resultados más completos.

Participantes

- *Evaluador supervisor*: persona encargada de dirigir la evaluación de usabilidad; asume el rol de moderador en las sesiones grupales de la evaluación.
- *Evaluadores expertos*: son los participantes más importantes de la evaluación, ya que influyen directamente en los resultados que presenta la realización de este método de evaluación. El grupo debe estar conformado por 3-5 evaluadores. Se recomienda que los evaluadores tengan entre 3 y 5 años de experiencia en el tema de evaluación de usabilidad y conocimiento acerca del dominio del sistema a evaluar.
- *Representante de la organización*: persona de la organización que tiene conocimiento acerca del sistema a evaluar. Es un participante importante, pero no necesariamente obligatorio.

Ventajas

- Es un método de evaluación económico.
- No consume tiempo de los usuarios.
- Permite identificar gran cantidad de problemas de usabilidad (mayores y menores), priorizarlos e interpretar su criticidad.
- Aplicable en etapas tempranas de desarrollo, pueden evaluarse prototipos verbales, en papel, de interfaces, ejecutables.

Desventajas

- Un número elevado de expertos aumentan los costos de la evaluación.
- Podrían ignorarse problemas específicos del dominio.

Descripción del proceso

Ver descripción del proceso de la evaluación heurística en el Anexo B.1.

3.4.2.2 Fase 2 – Descomposición de la técnica

En esta fase son identificadas las actividades generales que conforman el MEU. La Tabla 7 presenta las actividades generales que conforman la evaluación heurística, destacando que estas se agrupan en dos etapas de trabajo individual y dos etapas de trabajo grupal.

Tabla 7. Actividades generales de la *evaluación heurística*.

Actividades
<i>Trabajo individual</i>
Cada evaluador trabaja independientemente 1-2 horas, evaluando primero el sistema en general, después en detalles.
Cada evaluador prepara un listado individual de potenciales problemas de usabilidad y un listado de aspectos positivos.
<i>Trabajo grupal, coordinado por un evaluador supervisor</i>
Los evaluadores reúnen los problemas identificados en un solo listado.
<i>Trabajo individual</i>
Cada evaluador califica independientemente la severidad y la frecuencia de cada uno de los problemas del listado común.
Se calculan promedios de severidad, frecuencia y criticidad (severidad + frecuencia) para cada problema.
Se establecen rankings de problemas.
<i>Trabajo grupal, coordinado por un evaluador supervisor</i>
Los resultados obtenidos son analizados e interpretados.
Se sugieren recomendaciones de diseño para solucionar los problemas identificados.

3.4.2.3 Fase 3 – Establecer tareas colaborativas

En esta fase son detalladas las actividades que componen cada una de las actividades generales del método de evaluación (identificadas en la fase anterior) y son identificadas las que se realizarían de forma colaborativa [8]

La evaluación heurística es un método, que por naturaleza, presenta características que lo distinguen como un método colaborativo, ya que en su ejecución participan varios evaluadores. Considerando lo anterior, al momento de aplicar el modelo, el trabajo realizado consistió en refinar y especificar el proceso que conforma dicho método de evaluación.

Para identificar cuáles actividades se realizarían de forma colaborativa, los autores del presente trabajo consideraron los criterios definidos en [60] para consensuar el carácter colaborativo de las actividades que conforman el método de evaluación. Dichos criterios son:

1. La ejecución de la actividad y el plan de trabajo requieren incluir varias personas, las cuales pueden tener diferentes roles.
2. Se requiere contar con personas que tienen un alto grado de experticia en un área específica de conocimiento, para la ejecución de las actividades.
3. Es necesario compartir conocimiento, recursos e información con otras personas.
4. Es necesario tener en cuenta los diferentes aportes, opiniones y puntos de vista de otros integrantes del grupo que está ejecutando la actividad.

Ahora bien, las actividades que componen el método de evaluación han sido agrupadas en 3 etapas: planeación, ejecución y análisis de resultados. La *etapa de planeación* incluye las actividades de preparación realizadas antes de ejecutar el método. La *etapa de ejecución* se compone de las actividades que realizan los *evaluadores expertos* durante la inspección del sistema. La *etapa de análisis de resultados* incluye las actividades de análisis e interpretación de la información recolectada en la *etapa de ejecución*. Las Tablas 8 a 10 listan las actividades que componen cada etapa, dichas actividades corresponden a una recopilación de distintas fuentes bibliográficas relacionadas, tales como [1][5][44][61].

Tabla 8. Actividades de la *etapa de planeación*.

Nº	Actividad/Descripción	Colaborativa
1	Definir el sistema a evaluar: el <i>representante de la organización</i> define el sistema que será evaluado por el grupo de expertos en usabilidad.	NO
2	Elaborar presentación general del sistema: el <i>representante de la organización</i> elabora una presentación del sistema para enviar al <i>evaluador supervisor</i> .	NO
3	Revisar la presentación general del sistema a evaluar: el <i>evaluador supervisor</i> revisa la presentación elaborada por el <i>representante de la organización</i> , con el objetivo de familiarizarse con el sistema a evaluar.	NO
4	Identificar los posibles expertos a participar en la evaluación: el <i>evaluador supervisor</i> identifica un conjunto de posibles evaluadores expertos para que participen en la evaluación de usabilidad.	NO
5	Seleccionar los expertos que van a participar en la evaluación de usabilidad: el <i>evaluador supervisor</i> selecciona un conjunto de evaluadores (de 3 a 5) que tengan un nivel de experiencia bueno, cierta familiaridad con el dominio del sistema y estén disponibles.	NO
6	Identificar el conjunto de heurísticas a utilizar: el <i>evaluador supervisor</i> define el conjunto de heurísticas más adecuado y específico a las características del sistema que se quiere evaluar.	NO
7	Elaborar el documento guía para la evaluación: el <i>evaluador supervisor</i> prepara el documento guía que tendrán en cuenta los <i>evaluadores expertos</i> durante la evaluación de usabilidad.	NO
8	Proveer a los expertos la información general del sistema y el documento guía de la evaluación.	NO
9	Solucionar preguntas de los expertos: el <i>evaluador supervisor</i> responde a preguntas de los expertos relacionadas a la información proporcionada.	NO

Las actividades que conforman la *etapa de planeación* se estima que no requieren de trabajo colaborativo para su ejecución debido a que no cumplen todos los criterios definidos en [60]. Hay actividades que podrían ser colaborativas dependiendo la forma cómo sea tomada una decisión dentro de una organización o por el *evaluador supervisor*, pero para efectos de esta investigación esas actividades no son consideradas colaborativas. De igual forma, hay actividades en las que participan varias personas (por ejemplo la actividad N° 9: *Solucionar preguntas de los evaluadores*), pero dichas actividades no requieren trabajo colaborativo para su ejecución debido a la forma en cómo se llevarían a cabo, es decir, son actividades en las que no necesariamente participan todos los miembros del grupo.

Tabla 9. Actividades de la etapa de ejecución.

Nº	Actividad/Descripción	Colaborativa
10	Evaluación individual del sistema: cada evaluador experto inspecciona el sistema, para ello debe establecer su propio proceso o método de inspección, proporcionar escenarios de uso cuando sea necesario, agrupar los problemas detectados en los principios de usabilidad y anotar los elementos positivos del sistema.	NO
11	Creación de una lista integrada de problemas: los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> presentan los problemas detectados en el sistema y los principios heurísticos que incumple cada uno, con el fin de elaborar una lista integrada con los problemas de usabilidad identificados por cada evaluador.	SI
12	Calificación individual de los problemas de la lista integrada: cada evaluador asigna calificaciones de severidad y frecuencia a los problemas de la lista integrada (obtenida en la actividad 11).	NO

Tabla 10. Actividades de la etapa de análisis de resultados.

Nº	Actividad/Descripción	Colaborativa
13	Promediar las calificaciones de los evaluadores: el <i>evaluador supervisor</i> calcula promedios de severidad, frecuencia y criticidad para cada problema, teniendo en cuenta las calificaciones asignadas por los evaluadores. Adicionalmente, calcula la desviación estándar de los resultados correspondientes a las evaluaciones de los evaluadores.	NO
14	Generar un ranking de problemas en orden de importancia: el <i>evaluador supervisor</i> , según las calificaciones promediadas, realiza una clasificación o priorización de los problemas encontrados, según la criticidad y severidad.	NO
15	Análisis e interpretación de los resultados: los <i>evaluadores expertos</i> interpretan los resultados correspondientes a la cantidad de problemas identificados, los cuales están agrupados según el principio de usabilidad que incumplen. Los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> , interpretan los resultados de las calificaciones, la desviación estándar y los rankings (de criticidad y severidad) de los problemas detectados.	SI
16	Realizar recomendaciones para solucionar los problemas de usabilidad: los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> , proponen soluciones o recomendaciones de diseño para corregir los problemas de usabilidad identificados.	SI
17	Identificar elementos positivos del sistema: los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> identifican los aspectos positivos del sistema para darlos a conocer a los desarrolladores del mismo, con el fin de que ellos conserven esas buenas prácticas de diseño y programación.	SI
18	Elaborar el informe final de la evaluación: el <i>evaluador supervisor</i> elabora el informe de la evaluación en el cual incluye toda la información del proceso.	NO

Según las Tablas 9 y 10, una serie de actividades requieren de trabajo colaborativo debido a que cumplen con los criterios mencionados. Así, la ejecución de las actividades requiere: (1) incluir mínimo 3 evaluadores y un evaluador encargado, (2) que los evaluadores tengan un nivel de formación/experiencia medio/alto en la realización de evaluaciones de usabilidad o ser profesionales en el área de HCI, (3) compartir conocimiento, recursos e información entre los evaluadores y (4) tener en cuenta los diferentes aportes, opiniones y puntos de vista de los evaluadores que participan en el proceso de inspección.

Asociación de patrones de colaboración a las actividades colaborativas

Una vez especificadas las actividades y definidas cuáles de estas serán realizadas de forma colaborativa, a cada actividad definida como colaborativa se le asocia uno o más patrones de colaboración [8]. Para la selección de los patrones fueron considerados los subpatrones de cada uno de estos, los cuales pueden relacionarse a la actividad colaborativa en cuestión. Es importante resaltar que a una actividad colaborativa compleja se le pueden asociar varios patrones de colaboración. Adicionalmente, para cada una de las actividades colaborativas debe identificarse la siguiente información para su ejecución [8]:

- Entradas: recursos necesarios para ejecutar la actividad.
- Resultados esperados: lo que se pretende obtener una vez finalizada la actividad.
- Participantes: grupo de personas que realizan la actividad.
- Listado de patrones de colaboración: cada subactividad colaborativa se relaciona con uno o más patrones de colaboración. Para cada selección es necesario incluir la respectiva justificación.

La Tabla 11 presenta los resultados de asociar los patrones de colaboración a la actividad colaborativa N° 11: *Creación de una lista integrada de problemas*.

Tabla 11. Asociación de patrones a la actividad N° 11: *Creación de una lista integrada de problemas*.

Actividad: Creación de una lista integrada de problemas.	
Actividades relacionadas: 10 (Evaluación individual del sistema).	
Descripción: los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> presentan los problemas detectados en el sistema y los principios heurísticos que incumple cada uno, con el fin de elaborar una lista integrada con los problemas de usabilidad identificados por cada evaluador. En la lista no deben aparecer problemas repetidos o muy similares.	
Entradas:	
<ul style="list-style-type: none"> • Lista de problemas de usabilidad identificados por cada evaluador. 	
Resultados esperados:	
<ul style="list-style-type: none"> • Lista integrada de problemas de usabilidad. 	
Participantes: <i>Evaluador supervisor y evaluadores expertos</i> .	
Patrones	Justificación
Generación	Este patrón permite a los evaluadores realizar contribuciones para generar una lista integrada de los problemas de usabilidad identificados.
Clarificación	Permite lograr el entendimiento común de conceptos manejados por los evaluadores, además, permite dar explicaciones y descripciones alternativas de un concepto.

La Tabla 12 presenta los resultados de asociar los patrones de colaboración a la actividad colaborativa N° 15: *Análisis e interpretación de los resultados*.

Tabla 12. Asociación de patrones a la actividad N° 15: *Análisis e interpretación de los resultados*.

Actividad: Análisis e interpretación de los resultados.	
Actividades relacionadas: 10 (Evaluación individual del sistema), 13 (Promediar las calificaciones de los evaluadores), 14 (Generar un ranking de problemas en orden de importancia).	
Descripción: los <i>evaluadores expertos</i> interpretan los resultados correspondientes a la cantidad de problemas identificados, los cuales están agrupados según el principio de usabilidad que incumplen. Los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> , interpretan los resultados de las calificaciones, la desviación estándar y los rankings (de criticidad y severidad) de los problemas detectados.	
Entradas:	
<ul style="list-style-type: none"> • Lista integrada de problemas de usabilidad. • Cantidad de problemas de usabilidad, los cuales están organizados por principios. • Promedios de las calificaciones de los evaluadores. • Rankings de problemas de usabilidad ordenados según su criticidad y severidad. 	
Resultados esperados:	
<ul style="list-style-type: none"> • Contribuciones acerca del análisis e interpretación de los resultados. 	
Participantes: <i>Evaluador supervisor y evaluadores expertos</i> .	
Patrones	Justificación
Generación	Este patrón de colaboración permite a los evaluadores compartir aportes relacionados a la

Actividad: Análisis e interpretación de los resultados.	
	interpretación de la información recolectada, a partir de su experiencia, conocimiento y comentarios de los demás miembros del grupo.
Construcción de consenso	Permite a los evaluadores comprender los aportes realizados a partir del análisis de la información recolectada.

La Tabla 13 presenta los resultados de asociar los patrones de colaboración a la actividad colaborativa N° 16: *Realizar recomendaciones para solucionar los problemas de usabilidad*.

Tabla 13. Asociación de patrones a la actividad N° 16: *Realizar recomendaciones para solucionar los problemas de usabilidad*.

Actividad: Realizar recomendaciones para solucionar los problemas de usabilidad.	
Actividades relacionadas: 11 (Creación de una lista integrada de problemas), 15 (Análisis e interpretación de los resultados).	
Descripción: los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> , proponen soluciones o recomendaciones de diseño para corregir los problemas de usabilidad identificados.	
Entradas:	
<ul style="list-style-type: none"> • Lista integrada de problemas de usabilidad. • Rankings de problemas de usabilidad ordenados según su criticidad y severidad. 	
Resultados esperados:	
<ul style="list-style-type: none"> • Lista de recomendaciones de diseño que dan solución a los problemas de usabilidad. 	
Participantes: <i>Evaluador supervisor y evaluadores expertos</i> .	
Patrones	Justificación
Generación	Este patrón permite a los evaluadores generar recomendaciones de diseño para solucionar los problemas de usabilidad identificados, teniendo en cuenta su experiencia y conocimiento.
Construcción de consenso	Permite a los evaluadores discutir sobre las recomendaciones de solución propuestas, con el fin de identificar la (s) más adecuada (s).

La Tabla 14 presenta los resultados de asociar los patrones de colaboración a la actividad colaborativa N° 17: *Identificar elementos positivos del sistema*.

Tabla 14. Asociación de patrones a la actividad N° 17: *Identificar elementos positivos del sistema*.

Actividad: Identificar elementos positivos del sistema.	
Actividades relacionadas: 10 (Evaluación individual del sistema).	
Descripción: los <i>evaluadores expertos</i> en compañía del <i>evaluador supervisor</i> identifican los aspectos positivos del sistema evaluado para darlos a conocer a los desarrolladores del mismo, con el fin de que ellos conserven esas buenas prácticas de diseño y programación.	
Entradas:	
<ul style="list-style-type: none"> • Lista de elementos positivos del sistema, anotados por los evaluadores durante la inspección individual. 	
Resultados esperados:	
<ul style="list-style-type: none"> • Lista de elementos positivos del sistema evaluado. 	
Participantes: <i>Evaluador supervisor y evaluadores expertos</i> .	
Patrones	Justificación
Construcción de consenso	Permite a los evaluadores discutir sobre los elementos positivos identificados en el sistema.

3.4.2.4 Fase 4 – Relación con proceso colaborativo

Luego de asignar los patrones de colaboración a las actividades colaborativas, se requiere asociar a cada patrón seleccionado un *thinklet* adecuado a los procesos que conforman cada

actividad. Las tablas de relación de *thinklets* presentan el nombre de la actividad colaborativa, patrones de colaboración asignados y *thinklets* relacionados a cada patrón con su respectiva justificación, considerando los *aspectos para seleccionar los thinklets* [47].

La Tabla 15 presenta los resultados obtenidos de la relación de *thinklets* a la actividad colaborativa N° 11: *Creación de una lista integrada de problemas*.

Tabla 15. Relación de *thinklets* a la actividad N° 11: *Creación de una lista integrada de problemas*.

Actividad: Creación de una lista integrada de problemas.		
Patrones	Thinklet	Justificación de selección del thinklet
Generación	OnePage	<ul style="list-style-type: none"> Este <i>thinklet</i> permite a los evaluadores contribuir fácil y simultáneamente en la elaboración de la lista integrada de problemas de usabilidad. Los pasos especificados en este <i>thinklet</i> se ajustan de forma adecuada al proceso que conforma la actividad.
Clarificación	Concentration	<ul style="list-style-type: none"> El uso de este <i>thinklet</i> permite el entendimiento común de conceptos manejados por los evaluadores y limpiar la lista integrada de problemas la cual puede contener redundancias o ambigüedades. Los pasos de este <i>thinklet</i> se ajustan de forma adecuada al proceso que conforma la actividad. La combinación de este <i>thinklet</i> con el anterior (<i>OnePage</i>) es pertinente ya que en el <i>mapa de selección</i> el tipo de relación entre estos dos <i>thinklets</i> es excelente.

La Tabla 16 presenta los resultados obtenidos de la relación de *thinklets* a la actividad colaborativa N° 15: *Análisis e interpretación de los resultados*.

Tabla 16. Relación de *thinklets* a la actividad N° 15: *Análisis e interpretación de los resultados*.

Actividad: Análisis e interpretación de los resultados.		
Patrones	Thinklet	Justificación de selección del thinklet
Generación	OnePage	<ul style="list-style-type: none"> Este <i>thinklet</i> permite a los evaluadores contribuir fácil y simultáneamente en la elaboración de la lista de aportes. Los pasos de este <i>thinklet</i> se ajustan de forma adecuada al proceso que conforma la actividad.
Construcción de consenso	MoodRing	<ul style="list-style-type: none"> El uso de este <i>thinklet</i> estimula la discusión entre los evaluadores para comprender los aportes realizados por cada uno de ellos. Los pasos de este <i>thinklet</i> se ajustan de forma adecuada al proceso que conforma la actividad. La combinación de este <i>thinklet</i> con el anterior (<i>OnePage</i>) es posible según el <i>mapa de selección</i>.

La Tabla 17 presenta los resultados obtenidos de la relación de *thinklets* a la actividad colaborativa N° 16: *Realizar recomendaciones para solucionar los problemas de usabilidad*.

Tabla 17. Relación de *thinklets* a la actividad N° 16: *Realizar recomendaciones para solucionar los problemas de usabilidad*.

Actividad: Realizar recomendaciones para solucionar los problemas de usabilidad.		
Patrones	Thinklet	Justificación de selección del thinklet
Generación	OnePage	<ul style="list-style-type: none"> Este <i>thinklet</i> permite a los evaluadores contribuir fácil y simultáneamente en la realización de recomendaciones para solucionar los problemas de usabilidad identificados. Los pasos de este <i>thinklet</i> se ajustan de forma adecuada al proceso que conforma la actividad.
Construcción de consenso	MoodRing	<ul style="list-style-type: none"> El uso de este <i>thinklet</i> estimula la discusión sobre las recomendaciones de solución propuestas por los evaluadores, con el fin de identificar la (s) más adecuada (s). Los pasos de este <i>thinklet</i> se ajustan de forma adecuada al proceso que

Actividad: Realizar recomendaciones para solucionar los problemas de usabilidad.		
		conforma la actividad. <ul style="list-style-type: none"> • La combinación de este <i>thinklet</i> con el anterior (<i>OnePage</i>) es posible según el <i>mapa de selección</i>.

La Tabla 18 presenta los resultados obtenidos de la relación de *thinklets* a la actividad colaborativa N° 17: *Identificar elementos positivos del sistema*.

Tabla 18. Relación de *thinklets* a la actividad N° 17: *Identificar elementos positivos del sistema*.

Actividad: Identificar elementos positivos del sistema.		
Patrones	<i>Thinklet</i>	Justificación de selección del <i>thinklet</i>
Construcción de consenso	MoodRing	<ul style="list-style-type: none"> • El uso de este <i>thinklet</i> estimula la discusión entre los evaluadores para identificar los elementos positivos del sistema evaluado. • Los pasos de este <i>thinklet</i> se ajustan de forma adecuada al proceso que conforma la actividad.

3.4.2.5 Fase 5 – Documentación de la técnica

A partir de la información obtenida anteriormente, es generada la documentación definitiva del diseño colaborativo, para esto la Ingeniería de Colaboración ha definido los siguientes elementos [8]: *Modelo de Facilitación del Proceso* y *Agenda Detallada*. Estos documentos presentan la información relacionada al proceso que debe realizarse para ejecutar las actividades colaborativas que conforman los MEU.

3.4.1.5.1 Documentación del diseño - etapa de planeación

Las Figuras 11 y 12 presentan el *Modelo de Facilitación del Proceso* (MFP) de las actividades que conforman la *etapa de planeación*. Las actividades que no tienen asociado un patrón de colaboración y un *thinklet*, es porque que no se realizan de forma colaborativa.

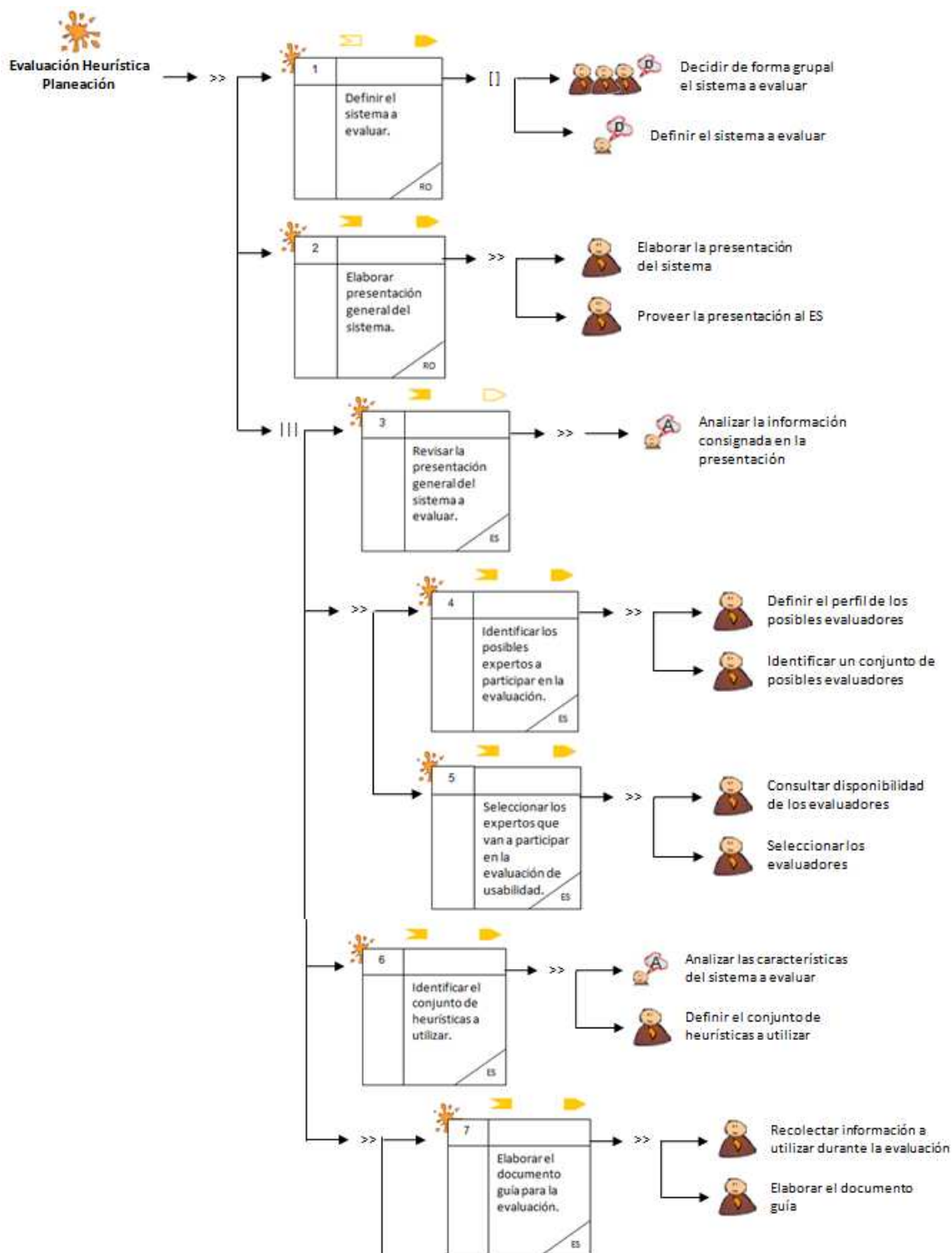


Figura 11. MFP de la etapa de planeación de la evaluación heurística. Parte 1.

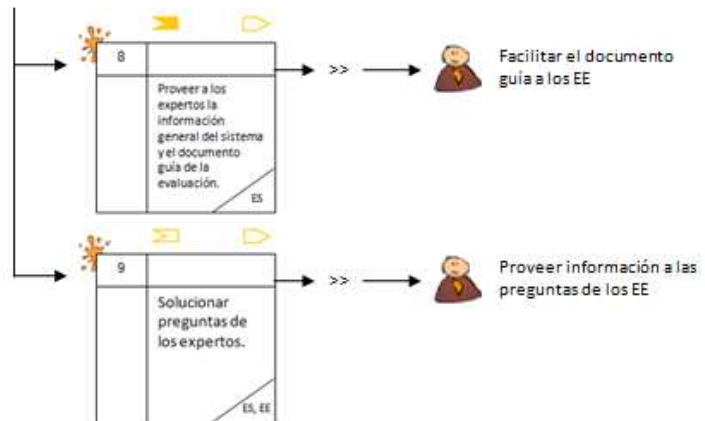


Figura 12. MFP de la etapa de planeación de la evaluación heurística. Parte 2.

La Tabla 19 presenta la *agenda detallada* con las actividades que conforman la *etapa de planeación*. Algunas actividades descritas en las *agendas detalladas* no tienen información en las columnas: *thinklet* y *patrón* y *proceso colaborativo*, debido a que no requieren de un proceso colaborativo para su ejecución.

Nota: En la columna *participantes* de las *agendas detalladas* se utilizan las siguientes abreviaturas para identificar los participantes de cada actividad: RO (Representante de la Organización), ES (Evaluador supervisor) y EE (Evaluadores Expertos).

Tabla 19. Agenda detallada con las actividades que conforman la etapa de planeación.

Nº	Actividad	Entregable	Pregunta/ Asignación	Patrón y Thinklet	Proceso Colaborativo	Participantes
1	Definir el sistema a evaluar.	Nombre del sistema a evaluar.	¿Cuál es el sistema a evaluar?			RO
2	Elaborar presentación general del sistema.	Presentación general del sistema a evaluar.	Realizar una presentación general del sistema a evaluar.			RO
3	Revisar la presentación general del sistema a evaluar.		Revisar en detalle la presentación general del sistema a evaluar.			ES
4	Identificar los posibles expertos a participar en la evaluación.	Lista de posibles evaluadores a participar en la evaluación de usabilidad.	¿Quiénes son los posibles evaluadores a participar en la evaluación de usabilidad?			ES
5	Seleccionar los expertos que van a participar en la evaluación de usabilidad.	Lista de evaluadores a participar en la evaluación de usabilidad.	Seleccionar los evaluadores a participar en la evaluación de usabilidad.			ES
6	Identificar el conjunto de heurísticas a utilizar.	Lista de heurísticas a utilizar en la evaluación.	Definir el conjunto de heurísticas más adecuado y específico a las características del sistema a evaluar.			ES
7	Elaborar el documento guía para la evaluación.	Documento guía para la evaluación, el cual será entregado a los expertos.	Construir el documento guía para la evaluación.			ES
8	Proveer a los expertos la información general del sistema y el documento guía de la evaluación.		Enviar a expertos la información general del sistema y el documento guía de la evaluación.			ES
9	Solucionar preguntas de los expertos.		Solucionar las preguntas que tengan los expertos relacionadas a la información suministrada.			ES, EE
<p>Observaciones</p> <ul style="list-style-type: none"> En la actividad N° 1, el <i>representante de la organización</i> define el sistema o partes del sistema a evaluar, esto último ya que el sistema puede estar conformado de múltiples funcionalidades y solo interesa evaluar algunas de ellas. En la actividad N° 2, el <i>representante de la organización</i> debe utilizar una plantilla para hacer la presentación del sistema, la cual incluya la siguiente información: objetivo del sistema, alcance de la evaluación, forma de acceder al sistema e información de contacto. Adicionalmente, es recomendable que el <i>representante de la organización</i> provea el manual de usuario del sistema junto con la presentación general. En la actividad N° 4, para facilidad del <i>evaluador supervisor</i>, conviene disponer de un registro con información de evaluadores expertos (como por ejemplo: una base de datos, documento de Excel u otro archivo que contenga información detallada de evaluadores), de tal forma, que al momento de identificar posibles participantes en la evaluación, se pueda encontrar fácilmente evaluadores que tengan experiencia en realizar evaluaciones heurísticas, conozcan las características del dominio del sistema y estén disponibles. 						

Nº	Actividad	Entregable	Pregunta/ Asignación	Patrón y Thinklet	Proceso Colaborativo	Participantes
						<ul style="list-style-type: none"> • En la actividad N°5, el grupo debe estar conformado por 3-5 evaluadores. • En la actividad N° 6, es recomendable utilizar un conjunto de heurísticas correspondientes al dominio del sistema a evaluar. Las heurísticas o principios heurísticos en que se basan los evaluadores pueden ser de distintas fuentes o autores. Generalmente, las heurísticas de Nielsen [5] son las más utilizadas para guiar este tipo de evaluación, sin embargo, estas son genéricas, por lo que es necesaria una posible adaptación o personalización de las heurísticas al sistema que se quiere evaluar. Para la realización de esta actividad se recomienda utilizar la herramienta software: Open-HEREDEUX (Open HEuristic REsource for Designing and Evaluating User eXperience) [62], propuesta por el Grupo de Investigación en Interacción Persona Ordenador e Integración de Base de Datos (GRIHO) de la Universidad de Lleida (España), la cual identifica un conjunto de heurísticas apropiadas a considerar durante la evaluación de usabilidad de diferentes sistemas interactivos. • En la actividad N° 7, el documento guía debe incluir información sobre el sistema a evaluar, las heurísticas a utilizar (definidas en la actividad N° 6) y procedimiento de evaluación. • En la actividad N°9, el <i>representante de la organización</i> puede participar de la reunión para proporcionar información detallada del sistema a evaluar.

3.4.1.5.2 Documentación del diseño - etapa de ejecución

La Figura 13 presenta el MFP de las actividades que conforman la etapa de ejecución.

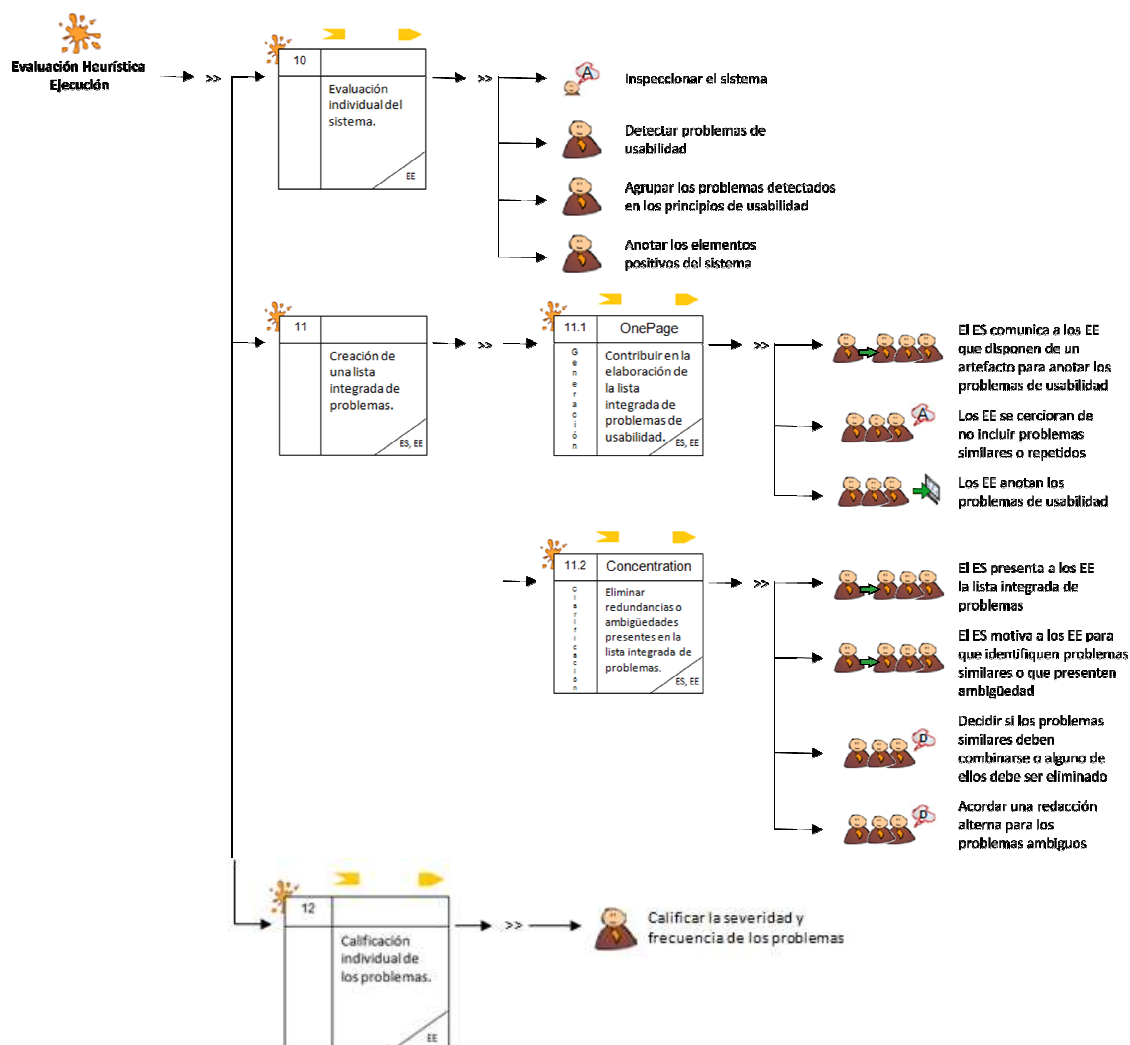


Figura 13. MFP de la etapa de ejecución de la evaluación heurística.

A continuación, la Tabla 20 presenta la agenda detallada con las actividades que conforman la etapa de ejecución.

Tabla 20. Agenda detallada con las actividades que conforman la etapa de ejecución.

Nº	Actividad	Entregable	Pregunta/Asignación	Patrón y Thinklet	Proceso Colaborativo	Participantes
10	Evaluación individual del sistema.	Plantilla diligenciada por cada evaluador con el listado de problemas identificados.	Cada evaluador experto debe inspeccionar el sistema.			EE
Actividad N°1 1: Creación de una lista integrada de problemas.						
11.1	Contribuir en la elaboración de la	Lista integrada preliminar de	Elaborar la lista integrada	Generación (OnePage)	Seguir el proceso colaborativo	ES, EE

Nº	Actividad	Entregable	Pregunta/Asignación	Patrón y Thinklet	Proceso Colaborativo	Participantes
	lista integrada de problemas de usabilidad.	problemas de usabilidad.	problemas de usabilidad.		definido en el Anexo D.1 para el <i>thinklet OnePage</i> .	
11.2	Eliminar redundancias o ambigüedades presentes en la lista integrada de problemas.	Lista integrada limpia (sin redundancias o ambigüedades) de problemas de usabilidad.	Limpia la lista integrada de problemas la cual puede contener redundancias o ambigüedades.	Clarificación (Concentration)	Seguir el proceso colaborativo definido en el Anexo D.1 para el <i>thinklet Concentration</i> .	ES, EE
12	Calificación individual de los problemas.	Calificación individual de los problemas de usabilidad de la lista integrada.	Cada evaluador experto debe asignar calificaciones a los problemas de la lista integrada (obtenida en la actividad 11).			EE

Observaciones

- En la actividad N° 10, deben tenerse en cuenta los siguientes aspectos:
 - Es recomendable que cada evaluador realice al menos dos iteraciones inspeccionando el sistema.
 - El evaluador debe recordar que puede darse el caso que un problema de usabilidad incumpla más de un principio heurístico.
 - Los problemas identificados en esta actividad deben registrarse en una plantilla con la siguiente información: identificador y definición del problema, comentarios o explicaciones, ejemplos de ocurrencia, principio (s) incumplidos (s), pantallas explicativas (opcional) y elementos positivos del sistema.
- En la actividad N° 11, deben tenerse en cuenta los siguientes aspectos:
 - En la lista integrada no deben repetirse problemas.
 - La lista integrada de problemas debe presentarse en una plantilla con la siguiente información: identificador y definición del problema, ejemplos de ocurrencia y principio (s) que incumple el problema.
- En la actividad N° 12, para la calificación de los problemas, los evaluadores asignan una nota de severidad (escala de 0 a 4) y una nota de frecuencia (escala de 0 a 4) a cada problema, permitiendo calcular la criticidad del mismo (Criticidad = Severidad + Frecuencia). Para la calificación de los problemas, se utilizan las siguientes escalas de severidad y frecuencia:

Nota	Severidad	Frecuencia
4	Catastrófico	(4) > 90%
3	Mayor	(3) 51-90%
2	Menor	(2) 11-50%
1	Cosmético	(1) 1-10%
0	No es un problema	(0) < 1%

- Las escalas muestran que mientras mayor es la calificación, ya sea en cuanto a severidad o frecuencia, el problema de usabilidad es mayor, lo que se verá reflejado en la criticidad del problema que corresponde a la suma de las dos calificaciones anteriores.
- De acuerdo a la severidad de los problemas de usabilidad, se tiene que: 1) Los problemas catastróficos impiden al usuario realizar una tarea o dan lugar a pérdidas catastróficas de datos o de tiempo. Estos problemas deben ser solucionados antes de que el sistema salga al público. 2) Los problemas mayores entorpecen de manera significativa la realización de tareas, pero el usuario puede encontrar una solución temporal. Es importante arreglar estos problemas, los evaluadores deben darles alta prioridad. 3) Los problemas menores son irritantes para el usuario pero no entorpecen de manera significativa la ejecución de tareas. Los evaluadores deben darles baja prioridad. 4) Los problemas cosméticos no resultan irritantes para los usuarios. No necesitan ser resueltos a menos que se disponga de tiempo extra en el proyecto.

3.4.1.5.3 Documentación del diseño - etapa de análisis de resultados

Las Figuras 14 y 15 presentan el MFP de las actividades que conforman la *etapa de análisis de resultados*.

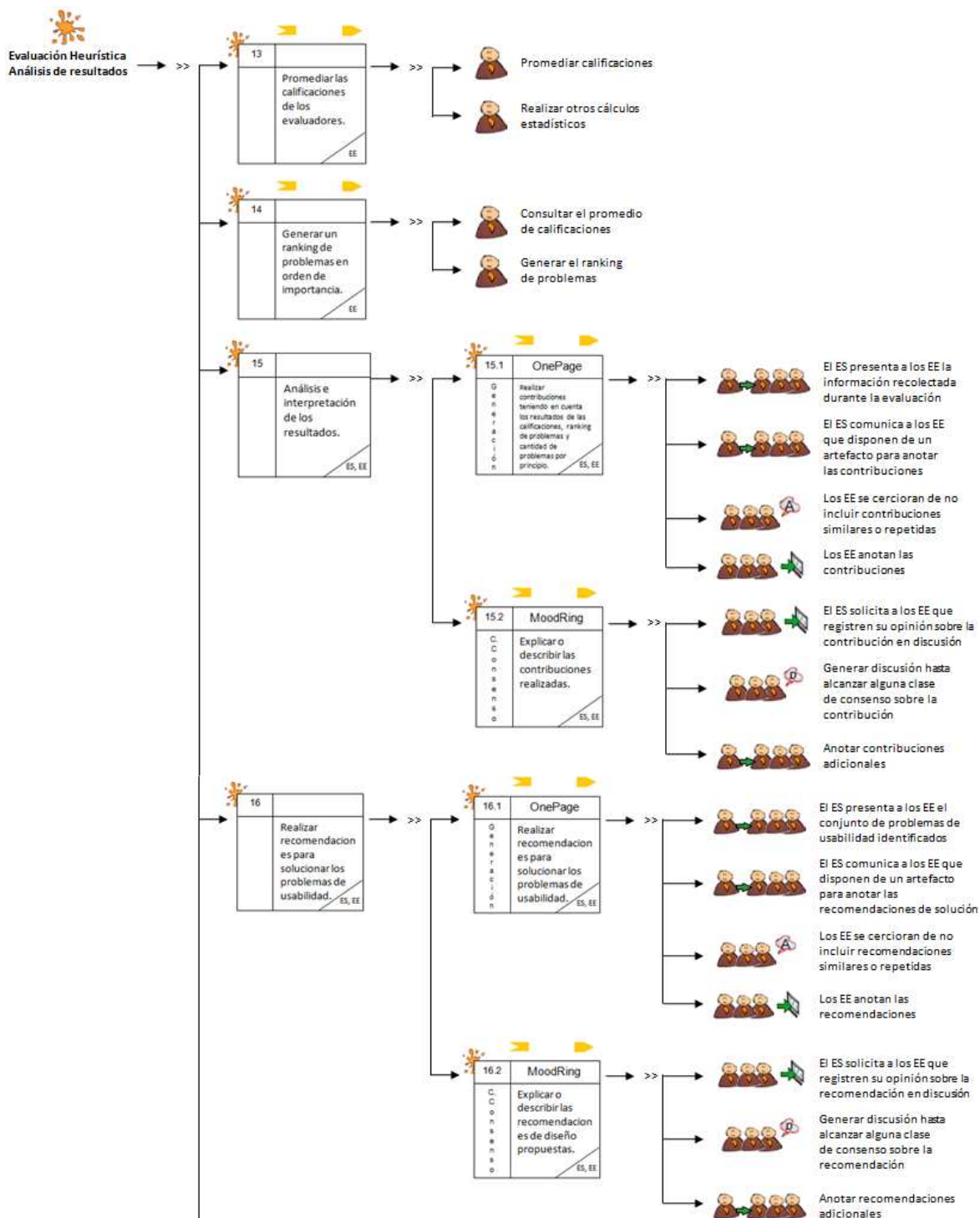


Figura 14. MFP de la etapa de análisis de resultados de la evaluación heurística. Parte 1.

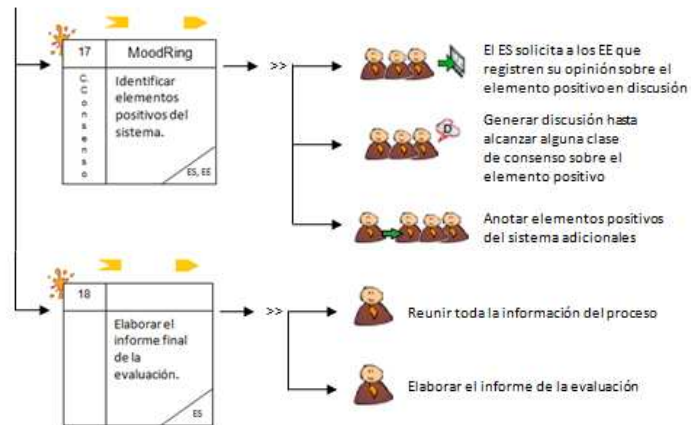


Figura 15. MFP de la etapa de análisis de resultados de la evaluación heurística. Parte 2.

En la Tabla 21 se presenta la *agenda detallada* con las actividades que conforman la *etapa de análisis de resultados*.

Tabla 21. Agenda detallada con las actividades que conforman la etapa de análisis de resultados.

Nº	Actividad	Entregable	Pregunta/ Asignación	Patrón y Thinklet	Proceso Colaborativo	Participantes
13	Promediar las calificaciones de los evaluadores.	Promedios de las calificaciones asignadas por los evaluadores expertos.	Calcular el promedio de las calificaciones de los evaluadores expertos.			ES
14	Generar un ranking de problemas en orden de importancia.	Ranking de problemas de usabilidad según su criticidad y severidad.	Generar un ranking de problemas de usabilidad según su criticidad y severidad.			ES
Actividad N°15: Análisis e interpretación de los resultados.						
15.1	Realizar contribuciones teniendo en cuenta los resultados de las calificaciones, ranking de problemas y cantidad de problemas por principio.	Contribuciones preliminares a partir del análisis de la información recolectada en la evaluación.	Analizar la información recolectada en la evaluación (resultados de las calificaciones, ranking de problemas, cantidad de problemas por principio) para realizar contribuciones.	Generación (OnePage)	Seguir el proceso colaborativo definido en el Anexo D.2 para el <i>thinklet OnePage</i> .	ES, EE
15.2	Explicar o describir las contribuciones realizadas.	Contribuciones finales a partir del análisis de la información recolectada en la evaluación.	Cada evaluador experto debe explicar o describir las contribuciones que ha realizado.	Construcción de consenso (MoodRing)	Seguir el proceso colaborativo definido en el Anexo D.2 para el <i>thinklet MoodRing</i> .	ES, EE
Actividad N°16: Realizar recomendaciones para solucionar los problemas de usabilidad.						
16.1	Realizar recomendaciones para solucionar los problemas de usabilidad.	Conjunto preliminar de recomendaciones para solucionar los problemas de usabilidad.	Proponer recomendaciones de diseño para solucionar los problemas de usabilidad.	Generación (OnePage)	Seguir el proceso colaborativo definido en el Anexo D.3 para el <i>thinklet OnePage</i> .	ES, EE
16.2	Explicar o describir las recomendaciones de diseño propuestas.	Recomendaciones de diseño finales para solucionar los problemas de usabilidad.	Cada evaluador experto debe explicar o describir las recomendaciones de diseño propuestas, con el fin de identificar la (s) más adecuada (s).	Construcción de consenso (MoodRing)	Seguir el proceso colaborativo definido en el Anexo D.3 para el <i>thinklet MoodRing</i> .	ES, EE
17	Identificar elementos positivos del sistema.	Lista de elementos positivos del sistema.	Cada evaluador debe explicar o describir los elementos positivos que ha identificado en el sistema, a partir de las anotaciones hechas en la actividad N°9.	Construcción de consenso (MoodRing)	<ol style="list-style-type: none"> 1. El <i>evaluador supervisor</i> solicita a los <i>evaluadores expertos</i> que registren su opinión sobre el elemento positivo en discusión. 2. El <i>evaluador supervisor</i> solicita a los evaluadores que hablen respecto al elemento positivo identificado. 3. El <i>evaluador supervisor</i> motiva a los evaluadores a que expresen su nueva opinión sobre el elemento positivo objeto de estudio, si escuchan algún comentario que cambie la idea que tenían inicialmente. 4. Generar discusión hasta alcanzar alguna clase de 	ES, EE

Nº	Actividad	Entregable	Pregunta/ Asignación	Patrón y Thinklet	Proceso Colaborativo	Participantes
					consenso sobre el elemento positivo identificado. 5. Continuar moderando la actividad hasta un tiempo previamente determinado o hasta que no se estén dando contribuciones en la discusión por parte de algún evaluador.	
18	Elaborar el informe final de la evaluación.	Informe final de la evaluación heurística.	El <i>evaluador supervisor</i> debe elaborar el informe de la evaluación incluyendo toda la información del proceso			ES
<p>Observaciones:</p> <ul style="list-style-type: none"> • En la actividad N° 14, los problemas con mayor calificación deben ser ordenados al principio del ranking, desde los que necesitan corrección urgente, hasta los problemas que en realidad no tienen importancia o no corresponden (un problema con calificación cero, no corresponde a un problema de usabilidad). En esta actividad para calcular los promedios, desviación estándar, etc., conviene el uso de herramientas software que agilicen el cálculo de dichos valores. • En la actividad N° 18, el <i>evaluador supervisor</i> debe incluir en el informe final la siguiente información: lista de problemas de usabilidad identificados en el prototipo o sistema evaluado, priorización de los problemas de usabilidad según su criticidad y severidad, contribuciones relacionadas al análisis e interpretación de los resultados obtenidos, lista de recomendaciones de diseño que dan solución a los problemas de usabilidad y lista de elementos positivos del sistema evaluado. 						

3.4.2.6 Fase 6 – Validación

En esta fase es realizada la validación del diseño colaborativo de la *evaluación heurística*. La validación del diseño colaborativo de este método de evaluación fue llevada a cabo mediante dos de las formas de validación propuestas en el modelo. La primera de ellas es la *simulación*, la cual permitió al grupo de trabajo realizar una serie de preguntas para probar la lógica del diseño, corroborar que a cada actividad se le crearán los entregables requeridos y de esta manera mejorar el diseño del método. La segunda es la *prueba piloto*, la cual permitió al grupo de trabajo aplicar el método de evaluación diseñado de forma colaborativa e identificar problemas en el diseño y verificar si la ejecución del método puede llevarse a cabo con los recursos definidos.

3.4.1.6.1 Simulación

Objetivo y proceso

El objetivo de la *simulación* es verificar el conjunto de actividades que conforman el diseño colaborativo del método de evaluación, así como también verificar el conjunto de entregables especificados en cada actividad. Esta validación es realizada con el fin de probar la lógica de diseño y a partir de ella realizar las mejoras pertinentes.

Para llevar a cabo la *simulación* son revisadas cada una de las actividades que conforman el método de evaluación, además, son realizadas una serie de preguntas para cada actividad colaborativa, las respuestas obtenidas sirven de referente para realizar las mejoras respectivas.

Participantes de la simulación

Las personas que participaron en el proceso de *simulación* para validar el diseño colaborativo de la evaluación heurística son: Cesar Collazos de la Universidad del Cauca (Colombia), Cristian Rusu de la Pontificia Universidad Católica de Valparaíso (Chile), Toni Granollers de la Universidad de Lleida (España) y Andrés Fernando Solano autor del presente trabajo.

El documento guía utilizado para la validación del diseño de la evaluación heurística mediante la *simulación* y los resultados obtenidos, son presentados en el Anexo E.

3.4.1.6.2 Prueba piloto

Objetivo

El objetivo de la *prueba piloto* es aplicar los procesos diseñados de forma colaborativa, con el fin de identificar problemas en el diseño y verificar si la ejecución de los procesos pueden llevarse a cabo con los recursos definidos.

La *prueba piloto* ha sido realizada sobre sistemas software interactivos correspondientes a diferentes áreas de aplicación. Las secciones 4.3.1.2.2, 4.3.2.2.2 y 4.3.3.2.2, presentan la información relacionada a la realización de la *prueba piloto* de la evaluación heurística.

Capítulo 4

Ningún jugador es tan bueno como todos juntos. Alfredo Di Stéfano

Ejecución de los MEU colaborativos

Este capítulo presenta información relacionada a la ejecución y análisis de resultados de los MEU objeto de estudio sobre sistemas que pertenecen a las áreas de aplicación: televisión digital interactiva, web transaccional y aplicaciones móviles. Aquí, son presentadas las áreas de aplicación estudiadas y los sistemas software interactivos objeto de estudio, correspondientes a dichas áreas. El análisis de los resultados obtenidos en la ejecución de los MEU es realizado con base en un conjunto de métricas previamente definidas. La Figura 16 presenta la estructura del presente capítulo.

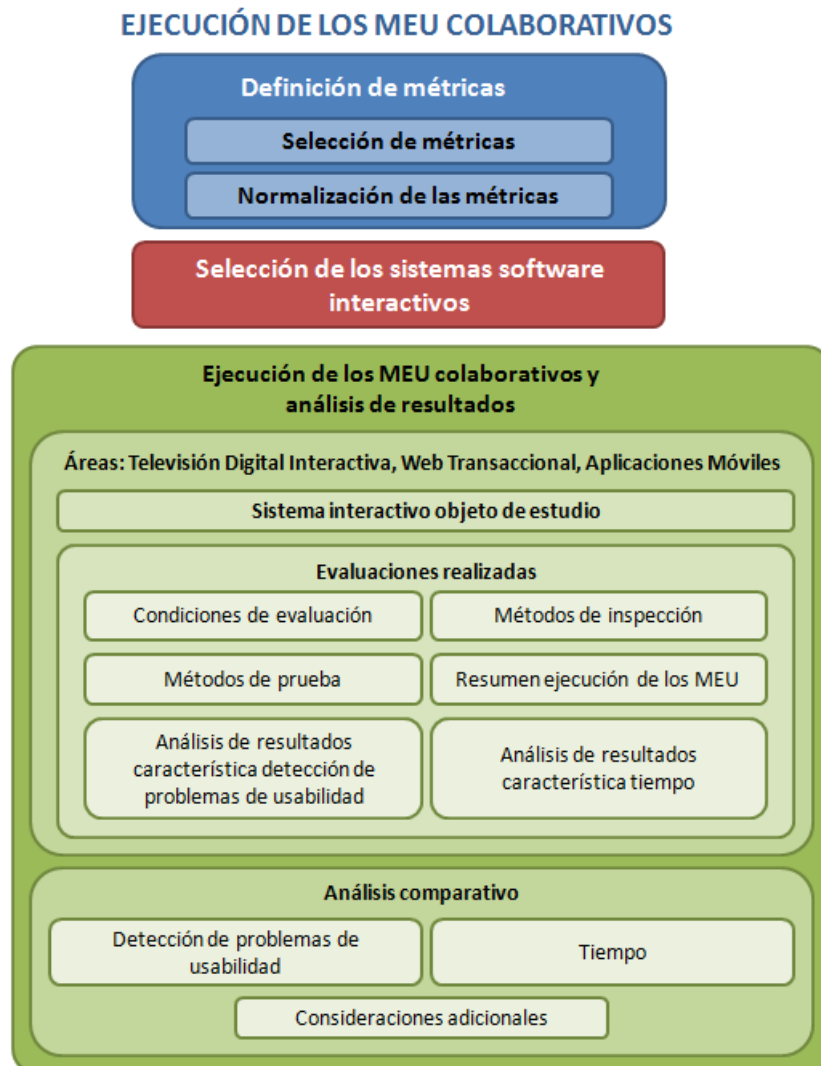


Figura 16. Estructura del capítulo 4.

4.1 Definición de métricas

Para el análisis de resultados es necesario definir un conjunto de métricas que permitan medir de forma objetiva los resultados obtenidos a partir de la ejecución de los MEU objeto de estudio. Para ello, luego de un proceso de observación y revisión de la literatura, se obtuvo una serie de métricas a partir de la ejecución de distintos métodos de evaluación, las cuales fueron agrupadas en las siguientes características:

Característica N°1: *Detección de problemas de usabilidad*

- Cantidad total de problemas identificados
- Cantidad de problemas críticos/severos
- Cantidad de problemas frecuentes
- Cantidad de problemas NO críticos
- Cantidad de problemas por funcionalidad

Característica N°2: *Recurso humano*

- Cantidad de expertos/evaluadores
- Cantidad de usuarios
- Cantidad de implicados
- Experiencia (en años) de expertos/evaluadores

Característica N°3: *Equipamiento*

- Cantidad de herramientas/tecnologías software requeridas
- Cantidad de dispositivos hardware requeridos
- Cantidad de materiales requeridos

Característica N°4: *Tiempo*

- Tiempo empleado para completar una tarea
- Tiempo invertido para recuperarse de errores
- Tiempo empleado para completar el método
- Tiempo empleado para completar la etapa de planeación
- Tiempo empleado para completar la etapa de ejecución
- Tiempo empleado para completar el análisis de resultados

Característica N°5: *Tareas*

- Cantidad de tareas propuestas
- Cantidad de tareas completadas
- Cantidad de tareas completadas por perfil de usuario
- Porcentaje de tareas completadas

Luego de definir el conjunto preliminar de métricas, fue elaborada una encuesta (ver Anexo L) con el fin de identificar, según la experiencia y conocimiento de un grupo de expertos, las métricas más relevantes para llevar a cabo el análisis de resultados. La encuesta fue elaborada utilizando el sistema SUS (*System Usability Scale*) [58], de tal manera que cada

pregunta tiene 5 opciones de respuesta. Fue realizado un consenso entre 11 expertos con experiencia en evaluaciones de usabilidad de sistemas interactivos (que realizan por lo menos 3 evaluaciones al año).

4.1.1 Selección de métricas

Una vez recopilados y procesados los resultados de las encuestas (incluyendo promedios y desviación estándar), fueron identificadas las métricas más relevantes según sus altos promedios. Dichas métricas son aquellas que fueron calificadas como “*importantes*” y “*muy importantes*” con base en la experiencia de los participantes que diligenciaron las encuestas.

Las métricas identificadas corresponden a las características: *detección de problemas de usabilidad*, *recurso humano* y *tiempo*. Sin embargo, al momento de hacer el análisis entre los MEU deben tenerse en cuenta las métricas generadas por los propios métodos, por lo cual las métricas de la característica *recurso humano* no serán consideradas como criterios para discriminar entre los MEU objeto de estudio. La razón es que dichas métricas no están relacionadas con el método de evaluación en sí, sino con una sesión de prueba en la que este es usado, lo cual es diferente. Por ejemplo, el número de personas que intervienen en la ejecución de un método (métrica *Cantidad de Implicados*) no debería ser un criterio para comparar entre varios MEU pues se estaría atribuyendo a una métrica un valor que no es generado por el propio método (los requisitos del método o las exigencias de trabajo del mismo). Lo mismo sucede con la métrica *Experiencia (en años) de expertos/evaluadores*, puesto que la experiencia de los evaluadores que participan en un método no revela nada sobre este. La Tabla 22 presenta la descripción de las métricas seleccionadas.

Las métricas seleccionadas corresponden a medidas base (o directas) según la teoría de la medición, esto indica que no dependen de ninguna otra medida y cuya forma de medir es un método de medición [58][63]. Por otro lado, las métricas que pertenecen a la característica *detección de problemas de usabilidad* están asociadas a un tipo de *escala absoluta* [63] ya que sólo hay una forma posible de medir: contando; mientras que las métricas de la característica *tiempo* están asociadas a un tipo de *escala de ratio* [63], la cual tiene un punto fijo de referencia: el cero (ningún valor puede ser menor a cero).

Ahora bien, una vez realizado el proceso de medición, los valores de las métricas no están entre 0 y 1 (exceden a 1), por lo cual debe ser utilizada una *tabla de normalización* para llevarlos a escala de valores entre 0 y 1 (ver en la sección 4.1.2 la información relacionada a la normalización de las métricas). Luego de normalizar los valores, las métricas generan un número real que está en un rango de 0 a 1. Así, las métricas proporcionan evidencias positivas si los valores son cercanos a 1.

Nota: en el caso de las métricas relacionadas al *tiempo*, cuyos valores “buenos” son los que se acercan a cero, habría necesidad de efectuar un cálculo como este: $V_c = 1 - V$. De modo que cuando el valor (V) de la métrica se acerque más a cero, el valor complementario (V_c) será más cercano a 1, tal que las métricas puedan llevarse a valores con sentido positivo (o creciente).

Nota: respecto a las métricas correspondientes a la característica *tiempo*, no ha sido establecido un tiempo base para realizar las etapas de planeación, ejecución y análisis de resultados. La razón es que el tiempo puede variar según el número de evaluadores y usuarios que participen en el proceso de evaluación. Por otro lado, la velocidad de ejecución de las etapas de un MEU (planeación, ejecución y análisis de resultados) no es un valor por

sí misma si no se asegura una mínima calidad en los resultados que se deriven de ellas. Por tal razón, el *evaluador supervisor* (rol definido en la especificación colaborativa de los MEU) es el encargado de asegurar cierto grado de calidad en los resultados/entregables obtenidos.

Tabla 22. Descripción de las métricas a considerar en el análisis de resultados de los MEU.

Métrica	Descripción	Interpretación
Característica: Detección de problemas de usabilidad		
Cantidad total de problemas identificados (CTP)	Esta métrica corresponde a la cantidad total de problemas de usabilidad identificados en el sistema evaluado.	Cuanto más problemas de usabilidad sean detectados, el valor de la métrica se acerca más a 1.
Cantidad de problemas críticos (CPC)	Esta métrica corresponde a la cantidad de problemas críticos identificados en el sistema evaluado.	Cuanto más problemas críticos sean detectados, el valor de la métrica se acerca más a 1.
Cantidad de problemas frecuentes (CPF)	Esta métrica corresponde a la cantidad de problemas frecuentes identificados en el sistema evaluado.	Cuanto más problemas frecuentes sean detectados, el valor de la métrica se acerca más a 1.
Característica: Tiempo		
Tiempo empleado para completar la etapa de planeación (TEP)	Esta métrica corresponde al tiempo empleado para realizar las actividades que conforman la <i>etapa de planeación</i> .	Cuanto menos tiempo sea empleado para realizar las actividades de la <i>etapa de planeación</i> , mejor. Habría que definir: $TEP_c = 1 - TEP$, para definir relación: menor (tiempo) – mayor (valor).
Tiempo empleado para completar la etapa de ejecución (TEE)	Esta métrica corresponde al tiempo empleado para realizar las actividades que conforman la <i>etapa de ejecución</i> .	Cuanto menos tiempo sea empleado para realizar las actividades de la <i>etapa de ejecución</i> , mejor. Habría que definir: $TEE_c = 1 - TEE$, para definir relación: menor (tiempo) – mayor (valor).
Tiempo empleado para completar el análisis de resultados (TEA)	Esta métrica corresponde al tiempo empleado para completar el análisis de resultados.	Cuanto menos tiempo sea empleado para analizar los resultados, mejor. Habría que definir: $TEA_c = 1 - TEA$, para definir relación: menor (tiempo) – mayor (valor).

Como fue mencionado anteriormente, las métricas que pertenecen a la característica *recurso humano* no serán consideradas para discriminar entre los MEU ejecutados. Sin embargo, cabe destacar que la métrica *Cantidad de Usuarios* proporciona evidencias positivas si el número de usuarios que participa en el método de prueba es mayor o igual al establecido en la Tabla 5 (Resumen comparativo entre los métodos de prueba) del Anexo B.3. Por otro lado, respecto a la métrica *Cantidad de Implicados*, esta proporciona evidencias positivas cuando en el método participa un número de usuarios mayor o igual al establecido, varios (mínimo 3) evaluadores y por lo menos un *representante de la organización*. Por tal razón, en la especificación colaborativa de los MEU objeto de estudio se indica explícitamente los participantes de cada método. Por último, respecto a la métrica *Experiencia (en años) de expertos/evaluadores*, esta proporciona evidencias positivas entre más alta sea, pues influye directamente en la cantidad y calidad de los resultados obtenidos en la ejecución de los métodos de evaluación (de inspección y prueba).

4.1.2 Normalización de las métricas

Una vez realizado el proceso de medición (por cada área de aplicación), los valores de las métricas no están entre 0 y 1 (exceden a 1). Por tal razón, conviene utilizar una *tabla de*

normalización para llevar las medidas a una escala de valores entre 0 y 1, y luego hacer el respectivo análisis e interpretación de los resultados. Así, para la elaboración de las *tablas de normalización* de las métricas correspondientes a las características: *detección de problemas de usabilidad* y *tiempo*, la Tabla 23 presenta un ejemplo de esquema de *tabla de normalización*.

Tabla 23. Ejemplo de esquema de *tabla de normalización*.

Clasificación del valor obtenido de la métrica	Calificación a asignar para "normalizar"
Valor_Métrica > A	0,99
B < Valor_Métrica <= A	0,90
C < Valor_Métrica <= B	0,80
D < Valor_Métrica <= C	0,70
E < Valor_Métrica <= D	0,60
F < Valor_Métrica <= E	0,50
G < Valor_Métrica <= F	0,40
H < Valor_Métrica <= G	0,30
I < Valor_Métrica <= H	0,20
J < Valor_Métrica <= I	0,10
Valor_Métrica <= J	0,01

En la Tabla 23 los valores A y J son el máximo y mínimo de la distribución, respectivamente. Para las métricas que pertenecen a la característica *detección de problemas de usabilidad*, el número de intervalos es establecido siguiendo algunas pautas definidas en [64]: en primer lugar, son identificados los valores máximo y mínimo de la distribución, y en segundo lugar, estos valores se restan y es buscado un número entero mayor que la diferencia el cual sea divisible por el número de intervalos que se desean establecer (en este caso 7, correspondiente al número de MEU objeto de estudio). Además, las *tablas de normalización* toman como valores de referencia (valores máximos) la cantidad de problemas detectados (críticos y frecuentes) en la evaluación heurística. Este método de evaluación es la base de comparación para el análisis de los métodos de prueba debido a que es uno de los métodos más eficientes y utilizados [65]. De esta manera, en la Tabla 23 el valor A puede ser la cantidad total de problemas identificados en la evaluación heurística, cantidad de problemas críticos o cantidad de problemas frecuentes, según la métrica a normalizar.

Para el caso de las métricas correspondientes a la característica *tiempo* también fueron consideradas las pautas definidas en [64] para establecer el número de intervalos (entre 11 y 15). Es importante resaltar que la medida de las métricas correspondientes a la característica *tiempo* es una aproximación, puesto que el tiempo para completar las actividades que conforman los MEU puede variar según el número de evaluadores y usuarios que participen en el proceso de evaluación, número de funcionalidades a evaluar, número de tareas diseñadas para los usuarios, entre otros factores. El tiempo en las actividades que son realizadas varias veces (como por ejemplo: la presentación de una prueba, la realización de tareas en el sistema por parte de los usuarios, la realización de preguntas para obtener información adicional, entre otras) corresponde al tiempo promedio que demora el *evaluador supervisor* o *usuario(s)* realizando la actividad en cuestión. De igual forma, el tiempo de las actividades colaborativas corresponde al tiempo promedio que cada evaluador invierte en la ejecución de los pasos propuestos en el(los) *thinklet(s)* respectivo(s).

4.2 Selección de los sistemas software interactivos

Con el objetivo de recolectar la mayor cantidad de información relacionada a las áreas de

aplicación (TDi, web transaccional y aplicaciones móviles), fueron seleccionados sistemas interactivos pertenecientes a cada una de estas con base en los siguientes criterios:

- **Representatividad del sistema:** el sistema interactivo seleccionado como objeto de estudio debe ser representativo del área de aplicación a la que pertenece. De esta manera, los resultados obtenidos en las evaluaciones podrían generalizarse a otros sistemas que pertenecen a dicha área.
- **Disponibilidad:** el sistema interactivo debe estar disponible de forma gratuita. De igual forma, el sistema debe ser de fácil acceso.
- **Tareas representativas en el sistema:** el sistema interactivo debe contar con una cantidad apropiada de funcionalidades y un buen nivel de navegabilidad, con lo cual será posible realizar tareas representativas.

Las áreas de aplicación antes mencionadas fueron seleccionadas porque se tiene fácil acceso a sistemas interactivos que pertenecen a cada una de ellas. Además, cabe destacar que se cuenta con una serie de heurísticas específicas para poder llevar a cabo una evaluación satisfactoria de los sistemas. Por otro lado, las áreas de aplicación seleccionadas corresponden a tecnologías emergentes, lo cual podría contribuir a la generalización de la metodología de evaluación de la usabilidad para diferentes sistemas software interactivos.

4.3 Ejecución de los MEU colaborativos y análisis de resultados

Esta sección presenta la ejecución y análisis de resultados de los MEU realizados sobre diferentes sistemas interactivos, siendo la televisión digital interactiva, web transaccional y aplicaciones móviles, las áreas de aplicación consideradas como base experimental. Al final, es realizado un análisis comparativo entre los MEU ejecutados en las diferentes áreas considerando las métricas previamente definidas (ver sección 4.1.1), las cuales van a permitir discriminar entre los métodos de evaluación ejecutados. Este análisis permitirá identificar los MEU más apropiados para conformar la *metodología de evaluación colaborativa de la usabilidad de sistemas software interactivos*.

La ejecución de los MEU corresponde a la realización de la *prueba piloto* [8] como una de las formas para validar la especificación colaborativa de dichos métodos. Una vez ejecutados los MEU fueron realizados una serie de ajustes a la especificación colaborativa, en especial a la agenda detallada y MFP de dichos métodos. *La ejecución de los MEU colaborativos es realizada bajo la premisa de que permiten obtener resultados más apropiados respecto a los definidos tradicionalmente, esto considerando los resultados obtenidos en [42].*

Respecto a los métodos de medición, cabe mencionar que no fueron utilizadas herramientas software que los automatizaran, las métricas están asociadas a un tipo de *escala absoluta* por lo cual la forma de medir fue: contando. Para conocer la medida de las métricas fue realizado un conteo manual, principalmente a partir del análisis de resultados y un seguimiento observacional durante la ejecución de los MEU (de las grabaciones de video, anotaciones en documentos guía de los evaluadores, entre otros), además de consultas a evaluadores que participaron en los procesos de evaluación (para el caso de los tiempos empleados en realizar una determinada actividad).

La siguiente sección presenta la descripción del área de aplicación y sistema interactivo objeto de estudio, así como también la ejecución y análisis de resultados de los MEU para el

caso de la TDi. Por restricción de extensión del documento, la ejecución y análisis de resultados de los métodos de evaluación en las áreas web transaccional y aplicaciones móviles, son presentados en los Anexos O y P, respectivamente.

4.3.1 Área: Televisión Digital Interactiva

La Televisión Digital Interactiva (TDi) es considerada como la convergencia de la televisión y las tecnologías de computación, que reúne tres características típicas [66]: interactividad, personalización y digitalización. En televisión un programa se refiere a “cualquier tipo de contenido, puede ser un anuncio, una película, un concurso o sonido” [67]. En la TDi el contenido de un programa, la forma o incluso el orden de presentación, puede ser afectado por el usuario. Además, el usuario tiene acceso a aplicaciones interactivas que pueden ser independientes o estrechamente relacionadas con el tema o el progreso del contenido.

La TDi sobrepasa la televisión analógica en varios aspectos: capacidad, mejor aprovechamiento del espectro, mayor inmunidad al ruido o interferencias, mejor calidad de imagen y sonido, posibilidad de transmisión de datos simultáneamente, ahorro de potencia en la transmisión. Sin embargo, la principal ventaja está en la capacidad de interactuar con el usuario. La interactividad permite al usuario ser parte activa de la programación, brindando la posibilidad de consultar o extender la información presentada, participar en foros, encuestas, chat y además controlar de cierta manera la secuencia de la información presentada [68].

En la actualidad el concepto de televisión ya no se refiere a un dispositivo específico, sino más bien a un tipo específico de contenidos presentes en casi todas partes, desde el televisor tradicional, al computador, pasando por los teléfonos móviles o las pantallas en los taxis y en toda la ciudad, llevando la televisión fuera de casa. Adicionalmente, las tendencias actuales pretenden combinar la TDi con la web, de tal manera que los usuarios puedan navegar por internet desde sus televisores, descargar y utilizar aplicaciones, descargar contenidos o personalizar la programación de una forma mejorada [68].

Las aplicaciones de TDi están definidas por un conjunto de características básicas, incluyendo aquellas que son comunes a otros sistemas interactivos. Dichas características deben considerarse para evaluar la usabilidad de este tipo de aplicaciones, esas son [69]:

- **Interactividad:** las aplicaciones de TDi deben ofrecer una comunicación bidireccional, un requerimiento fundamental de cualquier sistema interactivo. Una aplicación de TDi debe invitar al usuario a participar, con el objetivo de que tenga una experiencia más activa mientras observa un programa televisivo. La interactividad se refiere a la capacidad de ofrecer contenidos adicionales a los programas de televisión, permitiendo al usuario ver información asociada al contenido audiovisual, ver la programación de los canales, participar en concursos, votaciones, comprar productos o servicios, e incluso participar en la creación/personalización de los propios programas de televisión [70].
- **Personalización:** esta característica hace referencia al uso de la tecnología y a la información vista, para modificar la aplicación interactiva a cada perfil individual. Así, las aplicaciones de TDi deben permitir al usuario modificarlas en cuanto al contenido, apariencia u otros, teniendo en cuenta sus necesidades, características, preferencias personales, etc.
- **Características físicas de interacción:** los usuarios tienen una visión óptima a cierta distancia de la pantalla, por ello, las aplicaciones deben tener en cuenta aspectos de

contraste y resolución de la pantalla. Ésta es una característica especialmente diferenciadora, ya que se debe considerar que los usuarios ven televisión en un entorno que está orientado hacia la relajación y comodidad. Sin embargo, en la actualidad los usuarios pueden acceder a este medio en diferentes ambientes, desde múltiples dispositivos (televisores, móviles, etc.) y usando diferentes tecnologías (alta definición, 3D, etc.).

- Consistencia de aplicaciones y contenido: las aplicaciones de TDi deben estar relacionadas con el contenido y ser relevantes para usuarios específicos; en la TDi se ofrecen diversas aplicaciones que pueden ser fáciles de usar, pero si el usuario no encuentra relación alguna con el contenido, puede que resulten irrelevantes.
- Adaptabilidad: las aplicaciones de TDi deben tener la capacidad de adaptarse a diferentes tipos de público y ambientes. Incluso, estas deben sugerir programas/contenidos a los usuarios teniendo en cuenta sus preferencias, historial de opciones seleccionadas, entre otros. Esta característica se refiere a las aplicaciones reaccionando de forma proactiva para modificar su oferta de funciones/servicios en relación al usuario, contexto, etc.

4.3.1.1 Aplicaciones de TDi objeto de estudio

Teniendo en cuenta que en la TDi los usuarios pueden acceder a una amplia gama de aplicaciones, las cuales pueden ser implementadas para diferentes entornos como la Televisión Digital Terrestre (TDT), IPTV⁴, Televisión Móvil⁵, entre otras [68], en el presente proyecto de investigación las aplicaciones objeto de estudio son aquellas que se transmiten mediante el estándar tecnológico DVB (Digital Video Broadcasting)⁶, el cual fue adoptado en Colombia por la Comisión Nacional de Televisión⁷ en el año 2008, y que además, sigue la especificación MHP (Multimedia Home Platform)⁸. Esto significa que las aplicaciones pueden visualizarse en un televisor (y no en otros dispositivos, como: móviles, tabletas, entre otros) mediante el uso de una STB (Set-Top-Box), dispositivo que permite adaptar la señal digital. El estándar de transmisión DVB tiene varios modos de distribución: DVB-T, si la distribución se hace por el aire, DVB-C, si la distribución se hace por Cable, sin embargo, el contenido y las aplicaciones no se ven afectadas en el lado del cliente por estos modos de distribución.

Las aplicaciones presentan una serie de características que permiten al usuario controlar e intercambiar información de manera activa, a través de un canal de comunicación bidireccional o canal de retorno, que permite interactuar con el objeto de la transmisión, ya sea vídeo, imágenes, audio o datos. A través del receptor de TDi (STB) se realiza la integración y ejecución de las aplicaciones sobre el contenido televisivo, pero también se brinda la conexión por canal de retorno a un conjunto de aplicaciones y servicios remotos que pueden ser accedidos con base en la información que se desee en un momento dado, o a la dinámica del sistema de interacción [69].

Las aplicaciones de TDi objeto de estudio para esta investigación, han sido desarrolladas en el Laboratorio de Televisión Digital de la Universidad del Cauca de Colombia (el Anexo M presenta la descripción de la infraestructura del laboratorio). Las aplicaciones seleccionadas

⁴ IPTV (Internet Protocol Television): Sistema donde un servicio de televisión digital es entregado a sus clientes usando el protocolo IP sobre una infraestructura de red.

⁵ TELEVISIÓN MÓVIL: Sistema donde un servicio de televisión digital es entregado a sus clientes a través de un dispositivo móvil.

⁶ Disponible en: <http://www.dvb.org/>. Consultado en Junio de 2013.

⁷ Disponible en: http://www.cntv.org.co/cntv_bop/. Consultado en Junio de 2013

⁸ Disponible en: <http://www.mhp.org/>. Consultado en Junio de 2013.

como objeto de estudio son las siguientes:

- Guía de Programación Electrónica (o EPG por sus siglas en inglés Electronic Program Guide): ofrece información sobre los programas que se están emitiendo en un canal, los programas que se van a emitir en el día, los programas que se van a emitir en los días de la semana, la descripción de los programas, hora de inicio y fin, la duración, entre otros. En la EPG se pueden consultar los programas favoritos y los que han obtenido mayores votaciones en encuestas. Los usuarios pueden agregar o quitar programas de la lista de favoritos. La EPG puede ser consultada por un usuario pulsando el botón de interactividad azul del control remoto. Las Figuras 17 y 18, a continuación, corresponden a imágenes del menú principal de la EPG y de la opción "Semana", respectivamente.



Figura 17. Menú principal de la EPG.



Figura 18. Opción "Semana" de la EPG.

- Tablón o mini blog: permite a los usuarios publicar mensajes o noticias para que sean consultadas o comentadas por miembros de una comunidad. La Figura 19 corresponde a un conjunto de mensajes que se han publicado en el Tablón.



Figura 19. Mensajes publicados en el Tablón

- Chat: permite la comunicación entre usuarios de una comunidad. Esta aplicación se ha diseñado para utilizarla en paralelo a un programa de televisión que actúa como foco de la conversación. La Figura 20 corresponde a un conjunto de mensajes que se han enviado a través del Chat.



Figura 20. Mensajes enviados a través del Chat.

Las imágenes anteriores evidencian que las aplicaciones de TDi son prototipos que no han sido terminados completamente. Estas aplicaciones han sido seleccionadas porque disponen de varias funcionalidades y un mayor nivel de navegabilidad que otras aplicaciones disponibles, así es posible realizar más tareas representativas. Las aplicaciones de TDi objeto de estudio han sido desarrolladas considerando lineamientos básicos para TV como [71][72][73][74][75][76].

4.3.1.2 Evaluaciones realizadas

4.3.1.2.1 Condiciones de evaluación

La ejecución de los métodos de evaluación (de inspección y de prueba) estuvo sujeta a las siguientes condiciones:

- Las aplicaciones de TDi evaluadas corresponden a prototipos funcionales en una etapa

avanzada de desarrollo.

- El *representante de la organización* brindó toda la información necesaria y mantuvo constante atención a las actividades realizadas por el *evaluador supervisor*.
- En la ejecución de los MEU el grupo de evaluadores estuvo conformado básicamente por investigadores de distintos temas relacionados con el área de usabilidad y/o TDi.
- Los usuarios que participaron en los métodos de prueba tienen entre 22 y 29 años de edad, experiencia en el uso de tecnologías de la información y poca experiencia en el uso de aplicaciones de TDi.
- El lugar donde fueron realizadas las evaluaciones (Laboratorio de TDi) es de fácil acceso para los participantes del proceso. El laboratorio ofrece condiciones óptimas (de infraestructura hardware y software, iluminación, muebles) para la realización de las actividades individuales y grupales.
- El encargado de realizar las pruebas con usuarios fue el *evaluador supervisor*, por lo cual otros evaluadores no participaron en el proceso de observación de las acciones de los usuarios en las aplicaciones evaluadas.
- Las pruebas con usuarios fueron grabadas y luego distribuidas a los evaluadores, lo cual ayudó en su posterior análisis puesto que no se pierde información.
- Los dispositivos hardware utilizados por evaluadores y usuarios durante las evaluaciones son (ver Figura 21): televisor de 32 pulgadas, control remoto y STB.



Figura 21. Dispositivos hardware utilizados en las evaluaciones de usabilidad.

4.3.1.2.2 Métodos de inspección

La evaluación heurística fue realizada por un conjunto de 5 evaluadores que inspeccionaron el diseño de la interfaz de las aplicaciones con base en una serie de principios específicos para aplicaciones de TDi, definidos en [69]. Este método ha permitido identificar una serie de problemas mayores y menores en las aplicaciones de TDi objeto de estudio. Si bien, no se trata de problemas que ponen en riesgo el funcionamiento de las aplicaciones, sí atentan contra la facilidad de uso, y por lo tanto, con el buen aprovechamiento que estas pudiesen tener. En general, el nivel de criticidad de los problemas es alto, un buen número de problemas (16 de 24) fueron calificados, en promedio, con notas mayores a 6 (en una escala de 0 a 8), y 8 de los 24 problemas detectados fueron calificados con notas inferiores a 6. Lo anterior tiene su causa en que las aplicaciones evaluadas no están terminadas por completo, con lo cual los prototipos evaluados presentan varios aspectos posibles de mejorar.

El recorrido cognitivo fue realizado por un conjunto de 5 evaluadores que realizaron una serie de tareas con el fin de comprobar si las interfaces son adecuadas para los usuarios. Este método de inspección ha puesto al descubierto 20 problemas de usabilidad directamente relacionados con la facilidad de aprendizaje y de uso en las aplicaciones objeto de estudio. Los problemas detectados revelan que los usuarios reales tomarán más tiempo del estimado para realizar las tareas sugeridas en la evaluación. De igual forma, los usuarios necesitarían ayuda y/o asesoría de otras personas para completar las tareas satisfactoriamente. En ese sentido, los evaluadores estimaron que la experiencia de los usuarios respecto al uso e interacción con las aplicaciones evaluadas será medianamente satisfactoria.

El Anexo N.1 presenta los entregables de las actividades que conforman la especificación colaborativa de los métodos de inspección objeto de estudio.

4.3.1.2.3 Métodos de prueba

Los experimentos formales fueron realizados por un conjunto de 4 evaluadores y 8 usuarios representativos que se adecuan al perfil de usuario definido. Luego de la firma del acuerdo de confidencialidad (ver Anexo N.2), se llevaron a cabo las evaluaciones teniendo en cuenta la lista de tareas diseñada con base en los problemas críticos identificados en la evaluación heurística antes realizada (ver ranking de criticidad en Anexo N.1.1). Los sucesos ocurridos durante los experimentos fueron grabados (previa autorización de los usuarios) y distribuidos entre los evaluadores para su análisis (cada evaluador analizó 2 registros). La ejecución de este método permitió identificar 16 problemas de usabilidad en las funcionalidades objeto de estudio, entre los cuales fueron confirmados 8 de los 16 problemas más críticos identificados en la evaluación heurística.

Los cuestionarios fueron realizados luego de llevar a cabo los experimentos formales, por lo cual el número de usuarios que diligenciaron los cuestionarios antes (pre-test) y después (post-test) de los experimentos es 8 (mismo número de usuarios que participaron en los experimentos formales), una cifra aceptable para obtener conclusiones sobre la percepción subjetiva de los usuarios. Los cuestionarios post-test permitieron obtener resultados alentadores respecto a la satisfacción subjetiva de los usuarios con las aplicaciones de TDi evaluadas. Hay promedios que superan la nota 4 (en una escala de 1 a 5), por lo que se puede decir que, en general, los usuarios están conformes con la consistencia y control sobre las aplicaciones, sin embargo, consideran que el diseño de estas puede mejorarse para aumentar significativamente la productividad y satisfacción al usarlas.

El método interacción constructiva fue realizado por un conjunto de 4 evaluadores y 12 usuarios que se adecuan al perfil de usuario definido, por lo que, debido a la naturaleza de la prueba, fueron realizadas 6 interacciones constructivas. Previa firma del acuerdo de confidencialidad (ver Anexo N.2), las parejas de usuarios exploraban libremente las funcionalidades de las aplicaciones de TDi, al mismo tiempo que intercambiaban sus impresiones a viva voz. Los sucesos ocurridos durante las pruebas fueron grabados (previa autorización de los usuarios) y distribuidos entre los evaluadores para su posterior análisis. El desarrollo de este método de prueba permitió identificar 25 problemas de usabilidad, entre los cuales fueron confirmados los 16 problemas más críticos identificados en la evaluación heurística (ver ranking de criticidad en Anexo N.1.1). La interacción constructiva permitió identificar problemas en distintas partes de las aplicaciones de TDi, por lo que este método resulta apropiado para evaluar un sistema de forma global.

Las entrevistas fueron realizadas luego de que los usuarios interactuaron con las aplicaciones mediante el método interacción constructiva. Así, fueron 12 los usuarios entrevistados (mismo número de usuarios que participaron en la interacción constructiva), con el fin de obtener información de su percepción acerca de distintos aspectos de las aplicaciones de TDi. A partir de los resultados de las entrevistas, podría decirse que los usuarios consideraron poco amigable la interacción con las aplicaciones. La insuficiente información disponible y la deficiente navegación a través de las aplicaciones, fueron problemas identificados en la evaluación heurística y confirmados mediante las entrevistas.

El método del conductor fue realizado por un conjunto de 4 evaluadores y 8 usuarios representativos que se adecuan al perfil de usuario definido. Previa firma del acuerdo de confidencialidad (ver Anexo N.2), los usuarios realizaron las tareas indicadas por el conductor sobre las aplicaciones de TDi (con escenario de uso preestablecido). Los sucesos ocurridos durante la prueba fueron grabados (previa autorización de los usuarios) y distribuidos entre los evaluadores para su análisis (cada evaluador analizó 2 registros). La ejecución de este método permitió identificar 15 problemas de usabilidad en las funcionalidades estudiadas, entre los cuales fueron confirmados 11 de los 16 problemas más críticos identificados en la evaluación heurística (ver ranking de criticidad en Anexo N.1.1).

El Anexo N.3 presenta los entregables de las actividades que conforman la especificación colaborativa de los métodos de prueba objeto de estudio.

4.3.1.2.4 Resumen ejecución de los MEU

La Tabla 24 presenta, en resumen, información acerca de la ejecución de los MEU (de inspección y de prueba) en el área de TDi. El análisis de resultados es presentado más adelante en la sección 4.3.1.3.

Tabla 24. Resumen de la ejecución de los MEU en el área de TDi.

Método de evaluación	# Evaluadores	# Usuarios	# Problemas	# Problemas críticos confirmados	# Problemas no detectados en la evaluación heurística
Evaluación heurística	5	No aplica	24	16	No aplica
Recorrido cognitivo	5	No aplica	20	7	9
Experimentos formales	4	8	16	8	6
Cuestionarios	4	8	0	0	0
Interacción constructiva	4	12	25	16	5
Entrevistas	4	12	12	5	4
Método del conductor	4	8	15	11	2

Las siguientes secciones presentan el análisis de resultados considerando las características: *detección de problemas de usabilidad y tiempo*.

4.3.1.3 Análisis de resultados

4.3.1.3.1 Característica: detección de problemas de usabilidad

Métrica: Cantidad total de problemas identificados (CTP)

La Tabla 25 presenta la *tabla de normalización* propuesta para la métrica CTP. En la tabla

fue tomado como valor de referencia (valor máximo) la cantidad de problemas detectados en la evaluación heurística (ver Anexo N.1.1).

Tabla 25. Tabla de normalización para la métrica CTP.

Clasificación del valor obtenido de la métrica	Calificación a asignar para "normalizar"
Valor_Métrica > 24	0,95
20 < Valor_Métrica <= 24	0,80
16 < Valor_Métrica <= 20	0,65
12 < Valor_Métrica <= 16	0,50
8 < Valor_Métrica <= 12	0,35
4 < Valor_Métrica <= 8	0,20
Valor_Métrica <= 4	0,05

Con base en la información de la Tabla 25, la Tabla 26 presenta las medidas obtenidas y normalizadas de la métrica CTP.

Tabla 26. Medidas de la métrica CTP.

Método de evaluación	CTP	CTP normalizada
Evaluación heurística	24	0,80
Recorrido cognitivo	20	0,65
Experimentos formales	16	0,50
Cuestionarios	0	0,05
Interacción constructiva	25	0,95
Entrevistas	12	0,35
Método del conductor	15	0,50

Los MEU con los valores más altos respecto a la métrica CTP son la interacción constructiva y evaluación heurística. Mediante la interacción constructiva fue identificado el mayor número de problemas de usabilidad, esto considerando que las aplicaciones evaluadas son prototipos en un estado avanzado de desarrollo, razón por la cual fue identificada una cantidad significativa de problemas. Por otro lado, los métodos de evaluación con los valores más bajos son los cuestionarios y entrevistas. Esto obedece a que en dichos métodos las preguntas planteadas están orientadas a obtener información acerca de la satisfacción subjetiva de los usuarios, con lo cual la cantidad de problemas de usabilidad identificados se reduce significativamente. Mediante la ejecución de los cuestionarios no fueron identificados problemas de usabilidad, sin embargo, se obtuvo una serie de cálculos estadísticos con base en las respuestas de los usuarios los cuales están relacionados a la satisfacción subjetiva de los mismos, resultados que no dejan de ser relevantes.

Métrica: Cantidad de problemas críticos (CPC)

La Tabla 27 presenta la *tabla de normalización* propuesta para la métrica CPC. En la tabla fue tomado como valor de referencia (valor máximo) la cantidad de problemas críticos detectados en la evaluación heurística (ver Anexo N.1.1).

Tabla 27. Tabla de normalización para la métrica CPC.

Clasificación del valor obtenido de la métrica	Calificación a asignar para "normalizar"
Valor_Métrica = 16	1
13 < Valor_Métrica < 16	0,84
10 < Valor_Métrica <= 13	0,67
7 < Valor_Métrica <= 10	0,50
4 < Valor_Métrica <= 7	0,33
1 <= Valor_Métrica <= 4	0,16

Clasificación del valor obtenido de la métrica	Calificación a asignar para "normalizar"
Valor_Métrica = 0	0

Con base en la información de la Tabla 27, la Tabla 28 presenta las medidas obtenidas y normalizadas de la métrica CPC.

Tabla 28. Medidas de la métrica CPC.

Método de evaluación	CPC	CPC normalizada
Evaluación heurística	16	No aplica
Recorrido cognitivo	7	0,33
Experimentos formales	8	0,50
Cuestionarios	0	0
Interacción constructiva	16	1
Entrevistas	5	0,33
Método del conductor	11	0,67

Métrica: Cantidad de problemas frecuentes (CPF)

La Tabla 29 presenta la *tabla de normalización* propuesta para la métrica CPF. En la tabla fue tomado como valor de referencia (valor máximo) la cantidad de problemas frecuentes detectados en la evaluación heurística (ver Anexo N.1.1).

Tabla 29. Tabla de normalización para la métrica CPF.

Clasificación del valor obtenido de la métrica	Calificación a asignar para "normalizar"
Valor_Métrica = 15	1
12 < Valor_Métrica < 15	0,84
9 < Valor_Métrica <= 12	0,67
6 < Valor_Métrica <= 9	0,50
3 < Valor_Métrica <= 6	0,33
0 < Valor_Métrica <= 3	0,16
Valor_Métrica = 0	0

Con base en la información de la Tabla 29, la Tabla 30 presenta las medidas obtenidas y normalizadas de la métrica CPF.

Tabla 30. Medidas de la métrica CPF.

Método de evaluación	CPF	CPF normalizada
Evaluación heurística	15	No aplica
Recorrido cognitivo	4	0,33
Experimentos formales	9	0,50
Cuestionarios	0	0
Interacción constructiva	13	0,84
Entrevistas	5	0,33
Método del conductor	10	0,67

Los MEU que identificaron una mayor *Cantidad de problemas críticos (CPC)* y *Cantidad de problemas frecuentes (CPF)* son la interacción constructiva y el método del conductor. Aquí, se destaca el hecho de que la interacción constructiva confirmó los problemas críticos y la mayoría de problemas frecuentes detectados en la evaluación heurística. Por el contrario, los métodos con los valores más bajos en las métricas CPC y CPF, son: recorrido cognitivo, entrevistas y cuestionarios. Cabe mencionar que mediante el recorrido cognitivo fue identificada una cantidad significativa de problemas de usabilidad (CTP) y un número considerable de problemas que no fueron detectados en la evaluación heurística, sin

embargo, estos corresponden a problemas menores. Por otro lado, mediante los cuestionarios y entrevistas no fue identificado un número significativo de problemas de usabilidad, pero estos funcionaron como buen complemento para capturar las impresiones subjetivas de los usuarios.

4.3.1.3.2 Característica: tiempo

Para las métricas correspondientes a la característica *tiempo*, la *tabla de normalización* propuesta (en minutos) está dada de la siguiente manera (ver Tabla 31).

Tabla 31. Tabla de *normalización* para las métricas correspondientes a la característica *tiempo*.

Clasificación del valor obtenido de la métrica (minutos)	Calificación a asignar para "normalizar"
Tiempo_Actividades > 370	0,99
333 < Tiempo_Actividades <= 370	0,90
296 < Tiempo_Actividades <= 333	0,80
259 < Tiempo_Actividades <= 296	0,70
222 < Tiempo_Actividades <= 259	0,60
185 < Tiempo_Actividades <= 222	0,50
148 < Tiempo_Actividades <= 185	0,40
111 < Tiempo_Actividades <= 148	0,30
74 < Tiempo_Actividades <= 111	0,20
37 < Tiempo_Actividades <= 74	0,10
0 < Tiempo_Actividades <= 37	0,01

Respecto a la característica *tiempo* en el área de TDi, es importante mencionar que la mayoría de actividades colaborativas relacionadas a la etapa de planeación, ejecución y análisis de resultados, fueron realizadas de forma síncrona en un mismo lugar (laboratorio de TDi de la Universidad del Cauca). Por lo cual, las contribuciones por parte de los evaluadores fueron recolectadas en un periodo de tiempo corto.

Métrica: Tiempo empleado para completar la etapa de planeación (TEP)

La Tabla 32 presenta las medidas de la métrica TEP.

Tabla 32. Medidas de la métrica TEP.

Método de evaluación	TEP (minutos)	Valor normalizado de TEP	Complemento de TEP
Evaluación heurística	340	0,90	0,10
Recorrido cognitivo	205	0,50	0,50
Experimentos formales	378	0,99	0,01
Cuestionarios	108	0,20	0,80
Interacción constructiva	125	0,30	0,70
Entrevistas	55	0,10	0,90
Método del conductor	180	0,40	0,60

Los MEU con los valores más altos respecto a la métrica TEP son los cuestionarios y entrevistas, los cuales son métodos de interrogación que no necesitaron una extensa preparación. Los métodos con los valores más bajos son la evaluación heurística y experimentos formales, puesto que la *etapa de planeación* incluye un mayor número de actividades en comparación a la de los demás MEU objeto de estudio. La preparación de la evaluación heurística empleó un menor tiempo en comparación a los experimentos formales; en la evaluación heurística fue necesario un acercamiento a las aplicaciones de TDi y sus principales características, y los principios de usabilidad aplicados al caso, mientras que en los experimentos formales fue necesario un diseño de tareas detallado que incluye criterios de éxito y fracaso, tiempos máximos, entre otra información.

Métrica: Tiempo empleado para completar la etapa de ejecución (TEE)

La Tabla 33 presenta las medidas de la métrica TEE.

Tabla 33. Medidas de la métrica TEE.

Método de evaluación	TEE (minutos)	Valor normalizado de TEE	Complemento de TEE
Evaluación heurística	185	0,40	0,60
Recorrido cognitivo	135	0,30	0,70
Experimentos formales	32	0,01	0,99
Cuestionarios	15	0,01	0,99
Interacción constructiva	58	0,10	0,90
Entrevistas	20	0,01	0,99
Método del conductor	32	0,01	0,99

Los MEU con los valores más altos respecto a la métrica TEE son: cuestionarios, entrevistas, método del conductor y experimentos formales, por lo cual el tiempo promedio que invierte el *evaluador supervisor*, *evaluadores expertos* y *usuarios* realizando las actividades de la *etapa de ejecución* de dichos métodos, es reducido. Los cuestionarios y entrevistas tienen los valores más altos en esta métrica pues disponen de la información inmediatamente, por lo que rápidamente pueden hacerse los análisis estadísticos necesarios de las respuestas de los usuarios y sacar conclusiones. Por otro lado, el método más demorado en la etapa de ejecución fue la evaluación heurística. Este método, a pesar de ser calificado como simple para desarrollar, necesita de un período de tiempo considerable para que los evaluadores inspeccionen el sistema tranquilamente. A esto se suma que la inspección del sistema incluye fases de trabajo individual y grupal (según la especificación colaborativa propuesta para dicho método).

Métrica: Tiempo empleado para completar el análisis de resultados (TEA)

La Tabla 34 presenta las medidas de la métrica TEA.

Tabla 34. Medidas de la métrica TEA.

Método de evaluación	TEA (minutos)	Valor normalizado de TEA	Complemento de TEA
Evaluación heurística	215	0,50	0,50
Recorrido cognitivo	79	0,20	0,80
Experimentos formales	240	0,60	0,40
Cuestionarios	68	0,10	0,90
Interacción constructiva	122	0,30	0,70
Entrevistas	62	0,10	0,90
Método del conductor	165	0,40	0,60

Respecto a la métrica TEA, los MEU con los valores más altos son: entrevistas, cuestionarios y recorrido cognitivo. Esto indica que en dichos métodos el tiempo promedio invertido por los evaluadores analizando e interpretando la información (grabaciones en el caso de las entrevistas, cuestionarios diligenciados y consolidado de resultados, y anotaciones en los documentos guía en el caso del recorrido cognitivo) fue relativamente bajo. Lo anterior también obedece a que la información recolectada fue distribuida, en la medida de lo posible, en partes iguales entre los evaluadores que participaron en el proceso. Por otro lado, en los experimentos formales (método más demorado respecto al análisis de resultados) los evaluadores invirtieron una cantidad de tiempo significativa analizando e interpretando la información recolectada (como: grabaciones, cálculos estadísticos con base en las acciones de los usuarios, documentos guía diligenciados, entre otra).

4.3.2 Área: Web Transaccional

Con el rápido crecimiento de la web, el número de aplicaciones E-Commerce está continuamente en aumento. De esta manera, la usabilidad juega un rol importante pues las aplicaciones web transaccionales deben permitir a los usuarios lograr sus objetivos con efectividad, eficiencia y satisfacción, siempre teniendo en cuenta que las aplicaciones deben ser comprensibles y fáciles de usar [77]. Las aplicaciones web transaccionales son sistemas software particularmente complejos, en donde existe una interacción constante con el usuario y pueden ser repositorios de gran cantidad de información, además de ofrecer servicios que pueden ser accedidos desde cualquier lugar y desde variados dispositivos [9]. Ejemplos de sistemas relacionados a la categoría web transaccional son: sistemas bancarios, compras en línea, sistemas de reservas, entre otras, los cuales permiten al usuario realizar transacciones (no necesariamente financieras).

Una aplicación web puede ser definida como [78]: *“un sistema software basado en las tecnologías y estándares del World Wide Web Consortium (W3C) que ofrece recursos específicos de la web, como contenidos y servicios, mediante una interfaz de usuario denominada navegador web”*. Con esta definición no sólo se confirma la importancia de las aplicaciones web, sino también de las aplicaciones web transaccionales, ya que pueden ofrecer una gran variedad de recursos a un gran número de usuarios. Las aplicaciones web transaccionales han sido creadas para proveer mayor interactividad, permitiendo al usuario no solo interactuar con la aplicación mediante la lectura, sino que también generando actualizaciones sobre el contenido asociado [78].

Considerando un sistema de información turística, este podría permitir, por ejemplo, actualizar el contenido de una manera descentralizada o hacer una reserva de habitaciones. El requisito previo para esto es que los sistemas de bases de datos permitan un manejo eficiente y consistente de la creciente oferta de contenidos en aplicaciones web y ofrezcan la posibilidad de consultas estructuradas [78].

Las aplicaciones web transaccionales están definidas por un conjunto de características básicas, incluyendo aquellas que son comunes a otras aplicaciones web. Dichas características son [1][78]:

- Control en las transacciones: para mantener la seguridad y consistencia de los datos involucrados. Por ejemplo, un usuario transfiere dinero de una cuenta a otra de un mismo banco; la cantidad de dinero que se descuenta de la cuenta emisora debe ser igual a la que se suma en la cuenta receptora. De no ser así, la acción (transacción) no se realiza.
- Capacidad de corregir errores ocurridos durante una transacción: el usuario debe poder deshacer las operaciones realizadas, manteniendo los datos tal cual estaban antes del error.
- Control y administración de múltiples transacciones, determinando prioridades entre éstas. Por ejemplo, un usuario está haciendo la reserva de un asiento en un vuelo, dicho asiento debe ser bloqueado temporalmente hasta que se concrete la transacción, porque otro usuario podría estar queriendo reservar el mismo asiento en el mismo momento.
- Rapidez: deben ser capaces de responder rápidamente, en general la respuesta no debe ser mayor a un par de segundos.
- Fiabilidad: deben ser altamente fiables, de lo contrario podría afectar a usuarios, un

negocio, la reputación de una organización, etc. En caso de fallas, se deben proveer mecanismos de recuperación y respaldo de datos.

- Inflexibilidad: no pueden aceptar información distinta a la establecida. Por ejemplo, el sistema transaccional de una aerolínea debe aceptar reservas de múltiples agencias de viajes. Cada reserva debe contener los mismos datos obligatorios, con determinadas características.
- Amigables al usuario: ya que se trata de sistemas de información interactivos dirigidos al público, estos deben ser amigables y contar con altos estándares de diseño gráfico y visual.

4.3.2.1 Sitio web transaccional objeto de estudio

Considerando los criterios definidos en la sección 4.2, el sitio web transaccional seleccionado es Booking.com⁹, el cual es un sistema (final) que ofrece el servicio de reservas hoteleras y funciona a nivel internacional. Este sistema está disponible de forma gratuita, es de fácil acceso y presenta una cantidad apropiada de funcionalidades con buen nivel de navegabilidad. Booking.com fue seleccionado porque una *reserva hotelera* es un ejemplo representativo de transacción, por lo cual los resultados obtenidos en las evaluaciones podrían generalizarse a otros sistemas que pertenecen al área de web transaccional.

La empresa Booking.com ofrece el servicio de reservas de alojamiento online. Según datos de la empresa, cada día se reservan más de 550.000 noches a través de este sistema. El objetivo de Booking.com es ayudar a todo tipo de viajeros a descubrir, reservar y disfrutar de los mejores alojamientos de todo el mundo sea cual sea su presupuesto. Además, el propósito de este sistema es garantizar los mejores precios para todo tipo de alojamientos (pequeños cuartos, apartamentos, lujosas suites, entre otros). Booking.com está disponible en 41 idiomas, con una oferta que supera los 418.397 hoteles en 193 países. Esta información se encuentra disponible en la URL de la empresa.

4.3.2.2 Evaluaciones realizadas

Por restricción de extensión del documento, la información relacionada a la ejecución y análisis de resultados de los MEU objeto de estudio, es presentada en el Anexo O.

4.3.3 Área: Aplicaciones móviles

Las aplicaciones móviles son aquellas que fueron desarrolladas para ejecutarse en dispositivos móviles. El término "móvil" hace referencia a que puede accederse a los datos, las aplicaciones y los dispositivos desde cualquier lugar [79]. Ahora bien, la alta penetración de la telefonía móvil en el mercado, la fabricación de teléfonos inteligentes cada vez más potentes y el desarrollo de nuevos protocolos de telecomunicaciones, han permitido que los dispositivos móviles sean quizás el dispositivo más utilizado en la actualidad [80]. Los teléfonos móviles han pasado de ser simples herramientas para comunicarse a convertirse en auténticos centros de gestión para tareas cotidianas [81]. Con base en lo anterior, la usabilidad toma un lugar de importancia en el éxito de una aplicación móvil.

Para el desarrollo de aplicaciones móviles deben considerarse ciertas restricciones que tiene el hardware de los dispositivos, tales como [82]: dimensiones reducidas, bajo poder de cómputo, escasa capacidad de almacenamiento, variación impredecible de la calidad de red, ancho de banda limitado, entre otras. Los dispositivos móviles son suficientemente livianos como para ser transportados por personas y disponen de la capacidad de batería adecuada

⁹ Disponible en la URL: www.booking.com. Consultado en Marzo de 2014.

para funcionar de forma autónoma [82]. Algunos ejemplos de aplicaciones móviles son: mapas y navegación, búsqueda, juegos, mensajería, entre otras.

Las aplicaciones móviles deben manejar requerimientos, como [82]: potencial interacción con otras aplicaciones, manejo de sensores (por ejemplo el acelerómetro para responder a movimientos del dispositivo o pantalla táctil para responder a gestos), seguridad, bajo consumo de energía, entre otros. Adicionalmente, al momento de ejecutar una aplicación en un dispositivo móvil conviene tener en cuenta las siguientes consideraciones [83]:

- Eficiencia y efectividad: la aplicación debe ofrecer al usuario lo que busca de una manera rápida y sencilla. La aplicación debe satisfacer necesidades en el menor tiempo posible y sin necesidad de que el usuario tenga conocimiento acerca de su uso o requiera un largo proceso de aprendizaje.
- Distintas velocidades y características de red: al ser dispositivos móviles atravesarán distintas redes.
- Errores de red: la falta de una red de datos disponible es muy probable, en estos casos la aplicación debe tener un modo fuera de línea.
- Variación del rendimiento de la plataforma de hardware: lo ideal es que la aplicación esté disponible para tantos dispositivos como sea posible. Esto significa soportar diferentes dispositivos y diferentes plataformas.
- Distintos tamaños y resoluciones de pantallas: los diferentes dispositivos cuentan con diferentes pantallas con distintas funcionalidades.
- Difícil de probar las aplicaciones por completo: dada la variedad de dispositivos, se torna difícil probar todos los dispositivos actuales y los que ingresan al mercado.

4.3.3.1 Aplicación móvil objeto de estudio

Considerando los criterios definidos en la sección 4.2, la aplicación a evaluar es Dropbox¹⁰ en su versión gratuita para dispositivos móviles, la cual corresponde a una aplicación *Cloud Mobile*. Dropbox está disponible en versiones gratuitas y de pago, cada una de las cuales tiene opciones variadas, es de fácil acceso y proporciona una serie de funcionalidades con buen nivel de navegabilidad.

Mobile Cloud Computing (MCC) [84] se refiere a una infraestructura en la que el almacenamiento y procesamiento de datos tienen lugar fuera del dispositivo móvil, con lo cual las aplicaciones *Cloud Mobile* pueden utilizarse en una variedad de dispositivos móviles. MCC integra la tecnología *Cloud Computing* [81] en el entorno móvil, superando obstáculos relacionados con el rendimiento (por ejemplo, duración de la batería, almacenamiento, ancho de banda), el ambiente (por ejemplo, la heterogeneidad, escalabilidad y disponibilidad) y la seguridad (por ejemplo, la fiabilidad y la privacidad) [85].

Dropbox es un sistema de almacenamiento de archivos en Internet con el propósito de realizar copia de seguridad de los mismos, además de poder tenerlos sincronizados entre varios dispositivos. Al colocar los archivos en Dropbox, estos pueden obtenerse desde cualquier computador (con acceso a Internet), teléfonos móviles o tablets e incluso en el sitio web de Dropbox (www.dropbox.com).

En Dropbox es posible compartir archivos con otras personas, editar documentos, agregar fotos automáticamente, observar videos desde cualquier lugar, compartir fotos con amigos,

¹⁰ Disponible en: <https://www.dropbox.com/>, o en la tienda de aplicaciones de Google.

trabajar en equipo como si fuera utilizado un solo computador. Todo es privado de manera automática, además es posible controlar quién tiene acceso a qué archivos.

4.3.3.2 Evaluaciones realizadas

Por restricción de extensión del documento, la información relacionada a la ejecución y análisis de resultados de los MEU objeto de estudio, es presentada en el Anexo P.

4.4.4 Análisis comparativo

Con base en la ejecución de los MEU sobre las áreas de aplicación objeto de estudio, y considerando las condiciones de evaluación bajo las cuales estos fueron ejecutados, se tiene el siguiente análisis comparativo de resultados. En primer lugar, son analizadas las métricas asociadas a la característica *detección de problemas de usabilidad*, y en segundo lugar, son analizadas aquellas relacionadas a la característica *tiempo*.

4.4.4.1 Detección de problemas de usabilidad

Métrica: Cantidad total de problemas identificados (CTP)

La Tabla 35 presenta las medidas normalizadas de la métrica CTP según las áreas de aplicación objeto de estudio.

Tabla 35. Medidas de la métrica CTP por área de aplicación.

Método de evaluación	CTP normalizada por área de aplicación		
	TDi	Web transaccional	Aplicaciones móviles
Evaluación heurística	0,80	0,80	0,80
Recorrido cognitivo	0,65	0,50	0,35
Experimentos formales	0,50	0,50	0,50
Cuestionarios	0,05	0,05	0,05
Interacción constructiva	0,95	0,65	0,80
Entrevistas	0,35	0,35	0,35
Método del conductor	0,50	0,65	0,65

Según la Tabla 35, la evaluación heurística e interacción constructiva permitieron identificar la mayor cantidad de problemas de usabilidad en los sistemas interactivos objeto de estudio, puesto que en estos métodos los sistemas son evaluados de forma global. Esto es, los evaluadores (en la evaluación heurística) y las parejas de usuarios (en la interacción constructiva) exploran libremente todas las funcionalidades que ofrece el sistema, con lo cual el número de problemas detectados incrementa en buena medida.

Los resultados de la evaluación heurística fueron la base de comparación para el análisis de los métodos de prueba debido a que es uno de los métodos de evaluación más eficientes y más utilizados. Los resultados de la Tabla 35 confirman las buenas referencias de la evaluación heurística, por lo cual este método resulta apropiado para evaluar diferentes sistemas interactivos. Sin embargo, cabe resaltar que es altamente necesario disponer de un conjunto de heurísticas específicas para evaluar sistemas interactivos que pertenecen a áreas de aplicación diferentes de las estudiadas en la presente investigación.

Respecto a la interacción constructiva, es destacable que en cada área de aplicación detectó (en promedio) un porcentaje mayor o igual al 80% de los problemas identificados en la evaluación heurística. Estos resultados positivos podrían tener su causa en que dicho método de prueba es llevado a cabo de forma más natural por la pareja de usuarios ya que verbalizan sus impresiones conjuntamente. Con base en lo anterior, la interacción

constructiva resulta apropiada para la evaluación global de sistemas interactivos.

Por otro lado, se tienen los métodos: conductor, experimentos formales y recorrido cognitivo, los cuales utilizan una lista de tareas específicas para realizar en el sistema, por tal razón, el número de problemas detectados se minimiza (en comparación a los métodos que hacen una evaluación global) pues estos corresponden únicamente a las funcionalidades que han sido evaluadas. Entre los métodos antes mencionados, el conductor permitió identificar la mayor cantidad de problemas en dos de las áreas de aplicación objeto de estudio. Los resultados positivos obtenidos en el método del conductor se deben a que el evaluador (o conductor) puede controlar el curso de la prueba, esto con el fin de descubrir las necesidades de información de los usuarios en el sistema evaluado, considerando los problemas de usabilidad críticos identificados en métodos de inspección previamente ejecutados (en este caso la evaluación heurística). En general, el método del conductor permitió identificar la mayor cantidad de problemas entre aquellos que utilizan una lista de tareas específicas, razón por la cual este es un método candidato a conformar la *metodología para la evaluación colaborativa de la usabilidad de sistemas interactivos*.

Con base en lo anterior, los métodos: interacción constructiva y conductor, posiblemente son llevados a cabo en un contexto que no es el más real, ya que por ejemplo, en un ambiente de TDi o aplicaciones móviles, no interactúan dos usuarios al mismo tiempo con una aplicación (a través del mando a distancia o dispositivo móvil). Sin embargo, son métodos que funcionan muy bien y permiten recolectar buena cantidad de información acerca de las percepciones subjetivas de los usuarios, puesto que ellos se sienten con más confianza para expresar sus impresiones. Con lo cual, en áreas de aplicación diferentes a las estudiadas en este trabajo, dichos métodos podrían obtener resultados más apropiados respecto a otros métodos “clásicos” en el tema de evaluación de usabilidad.

Respecto a los métodos de indagación: entrevistas y cuestionarios, no fue identificada una cantidad significativa de problemas, es más, mediante los cuestionarios no fueron detectados problemas de usabilidad. Esto obedece a que en dichos métodos las preguntas planteadas están orientadas a obtener información acerca de la satisfacción subjetiva de los usuarios, con lo cual la cantidad de problemas de usabilidad identificados se reduce significativamente. Sin embargo, estos métodos de interrogación son buen complemento de otros métodos de prueba para obtener información adicional. Por ejemplo, en los cuestionarios, a partir del consolidado de resultados estadísticos (calculado con base en las respuestas de los usuarios) es posible obtener información acerca de la satisfacción subjetiva de los usuarios.

Métrica: Cantidad de problemas críticos (CPC)

La Tabla 36 presenta las medidas normalizadas de la métrica CPC según las áreas de aplicación objeto de estudio.

Tabla 36. Medidas de la métrica CPC por área de aplicación.

Método de evaluación	CPC normalizada por área de aplicación		
	TDi	Web transaccional	Aplicaciones móviles
Evaluación heurística	No aplica	No aplica	No aplica
Recorrido cognitivo	0,33	1	0,60
Experimentos formales	0,50	1	0,60
Cuestionarios	0	0	0
Interacción constructiva	1	1	1
Entrevistas	0,33	0,75	0,60
Método del conductor	0,67	1	0,80

Con base en la Tabla 36, la interacción constructiva y el método del conductor confirmaron la mayor cantidad de problemas críticos identificados en la evaluación heurística. Es destacable que, en cada una de las áreas de aplicación, la interacción constructiva confirmó todos los problemas críticos identificados en la evaluación heurística. Adicionalmente, la interacción constructiva fue uno de los métodos que permitió identificar la mayor cantidad de problemas de usabilidad (métrica CTP) en los sistemas interactivos estudiados, con lo cual se pone en evidencia el buen nivel de efectividad de este método.

El conductor es el segundo método con resultados positivos acerca de la confirmación de problemas críticos. Además, este es el tercer método que permitió identificar una mayor cantidad de problemas de usabilidad en los sistemas evaluados, detectando (en promedio) un porcentaje mayor o igual al 60% de los problemas identificados en la evaluación heurística. Por lo anterior, la realimentación (acciones e impresiones de los usuarios) que se obtiene con el método del conductor es de significativa relevancia.

Métrica: Cantidad de problemas frecuentes (CPF)

La Tabla 37 presenta las medidas normalizadas de la métrica CPF según las áreas de aplicación objeto de estudio.

Tabla 37. Medidas de la métrica CPF por área de aplicación.

Método de evaluación	CPF normalizada por área de aplicación		
	TDi	Web transaccional	Aplicaciones móviles
Evaluación heurística	No aplica	No aplica	No aplica
Recorrido cognitivo	0,33	0,84	0,50
Experimentos formales	0,50	0,67	0,50
Cuestionarios	0	0	0
Interacción constructiva	0,84	0,84	1
Entrevistas	0,33	0,67	0,75
Método del conductor	0,67	0,67	0,75

Con base en la Tabla 37, en cada una de las áreas de aplicación la interacción constructiva confirmó la mayor cantidad de problemas frecuentes identificados en la evaluación heurística. El método interacción constructiva no sólo obtuvo resultados positivos en esta métrica (CPF), sino también en la CPC y CTP. Esto lo convierte en un candidato que seguramente hará parte de la metodología de evaluación a proponer. Por otro lado, el método del conductor se encuentra en segundo lugar respecto a resultados positivos en la confirmación de problemas frecuentes y en tercer lugar se tiene las entrevistas.

4.4.4.2 Tiempo

Los tiempos involucrados en cada método pueden ser determinantes para escoger uno de estos sobre otro, principalmente cuando el tiempo es escaso, lo cual se da en la mayoría de los casos. Por lo anterior, los resultados obtenidos en las métricas asociadas a la característica *tiempo* darán soporte a la identificación de los MEU que van a conformar la *metodología para la evaluación colaborativa de la usabilidad de sistemas interactivos*.

Métrica: Tiempo empleado para completar la etapa de planeación (TEP)

La Tabla 38 presenta las medidas normalizadas de la métrica TEP según las áreas de aplicación objeto de estudio.

Tabla 38. Medidas de la métrica TEP por área de aplicación.

Método de evaluación	TEP normalizada por área de aplicación		
	TDi	Web transaccional	Aplicaciones móviles
Evaluación heurística	0,10	0,29	0,22
Recorrido cognitivo	0,50	0,43	0,43
Experimentos formales	0,01	0,01	0,08
Cuestionarios	0,80	0,78	0,91
Interacción constructiva	0,70	0,78	0,83
Entrevistas	0,90	0,92	0,91
Método del conductor	0,60	0,64	0,50

De acuerdo con la Tabla 38, las entrevistas y cuestionarios emplearon la menor cantidad de tiempo para su preparación. En general, la preparación de estos métodos consta de pocas actividades, entre las cuales la más destacada consiste en elaborar un conjunto de preguntas para realizar a los usuarios. Otros métodos con resultados positivos en cuanto a la métrica TEP son la interacción constructiva y conductor. El método con los resultados menos favorecidos consiste en los experimentos formales, puesto que su preparación incluye una mayor cantidad de actividades en comparación a los demás MEU objeto de estudio, además que requiere un diseño de tareas detallado (que incluye criterios de éxito y fracaso, tiempos máximos, entre otra información).

Métrica: Tiempo empleado para completar la etapa de ejecución (TEE)

La Tabla 39 presenta las medidas normalizadas de la métrica TEE según las áreas de aplicación objeto de estudio.

Tabla 39. Medidas de la métrica TEE por área de aplicación.

Método de evaluación	TEE normalizada por área de aplicación		
	TDi	Web transaccional	Aplicaciones móviles
Evaluación heurística	0,60	0,71	0,75
Recorrido cognitivo	0,70	0,85	0,83
Experimentos formales	0,99	0,92	0,91
Cuestionarios	0,99	0,99	0,99
Interacción constructiva	0,90	0,92	0,91
Entrevistas	0,99	0,99	0,99
Método del conductor	0,99	0,92	0,99

Según la Tabla 39, al igual que en la métrica TEP, las entrevistas y cuestionarios emplearon la menor cantidad de tiempo para su ejecución. Estos métodos disponen de la información inmediatamente, por lo que rápidamente pueden hacerse los análisis necesarios de las respuestas de los usuarios. Otros métodos con resultados positivos son: el conductor, experimentos formales e interacción constructiva. Claramente esta métrica depende del número de tareas sugeridas a los usuarios o funcionalidades a evaluar, por tal razón, en cada área de aplicación los métodos de evaluación que requieren un diseño de tareas fueron realizados considerando un mismo conjunto de tareas, esto con el fin de evitar resultados dispersos acerca de diferentes tareas.

Por otro lado, el método que empleó la mayor cantidad de tiempo en la *etapa de ejecución* en las diferentes áreas de aplicación es la evaluación heurística. Este método, a pesar de ser calificado como eficiente y simple para desarrollar, necesita de un período de tiempo considerable para que los evaluadores inspeccionen el sistema tranquilamente, además que la inspección del sistema incluye fases de trabajo individual y grupal (según la especificación colaborativa propuesta para el método).

Métrica: Tiempo empleado para completar el análisis de resultados (TEA)

La Tabla 40 presenta las medidas normalizadas de la métrica TEA según las áreas de aplicación objeto de estudio.

Tabla 40. Medidas de la métrica TEA por área de aplicación.

Método de evaluación	TEA normalizada por área de aplicación		
	TDi	Web transaccional	Aplicaciones móviles
Evaluación heurística	0,50	0,78	0,83
Recorrido cognitivo	0,80	0,92	0,91
Experimentos formales	0,40	0,36	0,29
Cuestionarios	0,90	0,85	0,91
Interacción constructiva	0,70	0,85	0,67
Entrevistas	0,90	0,92	0,99
Método del conductor	0,60	0,78	0,91

Teniendo en cuenta la información de la Tabla 40, las entrevistas, cuestionarios y recorrido cognitivo emplearon la menor cantidad de tiempo en el análisis de resultados. Por lo cual, el tiempo promedio invertido por los evaluadores analizando e interpretando la información recolectada (como grabaciones, consolidado de resultados de los cuestionarios, anotaciones en documentos guía) es relativamente bajo. Respecto a las entrevistas realizadas en las áreas web transaccional y aplicaciones móviles, cabe mencionar que el poco tiempo demorado en el análisis de resultados no necesariamente estuvo relacionado con las características del método, sino a la disponibilidad limitada de los evaluadores para participar en las actividades colaborativas sugeridas. Por tal razón, en el análisis e interpretación de los registros de las entrevistas se obtuvieron pocas contribuciones por parte del grupo de evaluadores. En ese sentido, el tiempo de las entrevistas pudiese haber sido mayor si los evaluadores hubieran participado activamente.

Las entrevistas y cuestionarios permiten obtener información complementaria (cualitativa y cuantitativa) a la ejecución de otros métodos de prueba, sin embargo, dichos métodos de interrogación permiten identificar una cantidad limitada de problemas pues están enfocados en obtener información acerca de la satisfacción subjetiva de los usuarios (experiencia con el sistema, perspectivas, impresiones, preferencias, etc.). Ahora bien, considerando los resultados positivos en las métricas asociadas a la característica *tiempo*, los métodos de interrogación (entrevistas y cuestionarios) resultan apropiados para complementar la ejecución de otros métodos de prueba, debido a que su preparación, ejecución y análisis, no demandan una cantidad significativa de tiempo.

El tiempo empleado en las actividades de análisis está sujeto a la cantidad de información recolectada y cantidad de problemas de usabilidad identificados. En ese sentido, si es identificado un número significativo de problemas, los evaluadores demoran más tiempo analizando la causalidad de estos. Teniendo en cuenta lo anterior, otros métodos con resultados positivos respecto a la métrica TEA son el conductor y la interacción constructiva, pues el tiempo de análisis de la información (registros de las pruebas) está relacionado directamente con el tiempo de duración de los registros (los evaluadores necesitan visualizar la interacción de los usuarios con el sistema para luego realizar el respectivo análisis e interpretación de las acciones). Esto corresponde a que los métodos mencionados obtuvieron también resultados positivos en la métrica TEE.

Por otro lado, los métodos que emplearon la mayor cantidad de tiempo (por sus bajos valores) en el análisis de resultados son la evaluación heurística y experimentos formales.

Por un lado, las evaluaciones heurísticas emplearon un período de tiempo considerable en el que los evaluadores realizaron sus contribuciones con base en el análisis de la información (rankings de criticidad, severidad y frecuencia), identificaron elementos positivos de los sistemas evaluados, entre otras actividades. Por otro lado, en los experimentos formales (método más demorado en la *etapa de análisis de resultados*) los evaluadores analizaron las acciones realizadas por los usuarios en cada una de las tareas propuestas, el tiempo demorado en alcanzar cada tarea, casos de fracaso, entre otra información, lo cual requirió que cada evaluador dedicara una cantidad de tiempo importante.

Respecto a los métodos de prueba, el tiempo dedicado al análisis de resultados depende de la cantidad de usuarios que participan en las evaluaciones. Además, el tiempo promedio que invierten los evaluadores depende en buena medida de la forma en cómo es distribuida la información (registros de las pruebas) entre ellos. Por otro lado, los resultados obtenidos en la métrica TEA, relacionados al tiempo promedio que invierte un evaluador en la ejecución de los pasos propuestos en el(los) *thinklet(s)* asociado(s) a las actividades de análisis, podrían indicar que el trabajo colaborativo contribuye de manera positiva en el tiempo requerido para analizar la información recolectada en la ejecución de los MEU. Esto debido a que la información es distribuida entre varias personas (grupo de evaluadores) lo cual reduce el tiempo y esfuerzo dedicado por el responsable de la evaluación, además de que el análisis de la información no estaría limitado a la percepción de una sola persona. Adicionalmente, se podría estimar que las combinaciones de *thinklets* propuestas permiten obtener experiencias positivas en cuanto al tiempo requerido para obtener una serie de contribuciones por parte de los evaluadores.

Adicional a los puntos antes mencionados, la métrica *Experiencia (en años) de expertos/evaluadores* (EAE) influye directamente en la cantidad y calidad de las contribuciones obtenidas durante el análisis de la información recolectada en la ejecución de los MEU. La Tabla 41 presenta la experiencia promedio de los evaluadores que participaron en la ejecución de los MEU sobre las diferentes áreas de aplicación.

Tabla 41. Experiencia (en años) de expertos/evaluadores por área de aplicación.

Método de evaluación	Experiencia (en años) de expertos/evaluadores (EAE) – TDi					Experiencia promedio
	Evaluador 1	Evaluador 2	Evaluador 3	Evaluador 4	Evaluador 5	
Evaluación heurística	Más de 10	6	4	5	2	5,4
Recorrido cognitivo	5	2	3	2	3	3
Experimentos formales	3	2	3	3	No aplica	2,75
Cuestionarios	3	2	3	3	No aplica	2,75
Interacción constructiva	3	2	3	3	No aplica	2,75
Entrevistas	3	2	3	3	No aplica	2,75
Método del conductor	6	4	5	3	No aplica	4,5
Método de evaluación	Experiencia (en años) de expertos/evaluadores (EAE) – Web transaccional					Experiencia promedio
	Evaluador 1	Evaluador 2	Evaluador 3	Evaluador 4	Evaluador 5	
Evaluación heurística	Más de 10	5	6	4	5	6
Recorrido cognitivo	5	4	6	5	No aplica	5
Experimentos formales	5	3	3	5	No aplica	4
Cuestionarios	5	3	3	5	No aplica	4

Interacción constructiva	4	4	3	3	No aplica	3,5
Entrevistas	4	4	3	3	No aplica	3,5
Método del conductor	3	5	3	5	No aplica	4
Método de evaluación	Experiencia (en años) de expertos/evaluadores (EAE) – Móviles					Experiencia promedio
	Evaluador 1	Evaluador 2	Evaluador 3	Evaluador 4	Evaluador 5	
Evaluación heurística	Más de 10	5	5	5	5	6
Recorrido cognitivo	7	5	5	5	No aplica	5,5
Experimentos formales	5	3	3	5	No aplica	4
Cuestionarios	5	3	3	5	No aplica	4
Interacción constructiva	4	3	3	5	No aplica	3,75
Entrevistas	4	3	3	5	No aplica	3,75
Método del conductor	3	3	3	5	No aplica	3,5

Con base en la Tabla 41, en las evaluaciones heurísticas participaron evaluadores con experiencia en evaluaciones de usabilidad de sistemas interactivos (5 años en promedio), lo cual supone que incrementa la confiabilidad de los resultados obtenidos. En los demás MEU participaron evaluadores con menor cantidad de años de experiencia (entre 3 y 5 años en promedio). En ese sentido, el hecho de contar con evaluadores más experimentados permitiría obtener resultados más apropiados en el análisis e interpretación de los datos.

En el caso de la interacción constructiva llevada a cabo en el área de TDi, participaron evaluadores que, en general, tenían poca experiencia en evaluaciones de usabilidad (entre 2 y 3 años de experiencia), sin embargo, estos identificaron un buen número de problemas de usabilidad (mayor al de la evaluación heurística). Dado que la realimentación de los usuarios es el principal insumo para detectar problemas de usabilidad, la clave estaría en la interpretación adecuada de las acciones, impresiones y demás comentarios de los usuarios. Ahora bien, la participación de evaluadores experimentados en actividades como: analizar la criticidad de los problemas, proponer posibles soluciones de diseño, entre otras, incrementa el grado de confiabilidad de la información obtenida, a pesar de que contar con la colaboración de expertos es complicado (por cuestiones de disponibilidad) y costoso. En ese sentido, los evaluadores que participen en los MEU deben ser personas que conozcan en cierto grado las características que debe poseer un sistema usable y que tengan cierta experiencia en el método a ejecutar y dominio del sistema, algunos serán más expertos que otros, pero lo óptimo es que todos lo sean en niveles aceptables.

Considerando lo anterior, resulta complicado asegurar que la experiencia de los evaluadores es causante de un mayor rendimiento al momento de detectar problemas de usabilidad, porque en algunos métodos han participado más o menos evaluadores con más años de experiencia que otros. Además, en la vida real no es tarea fácil conseguir evaluadores con alto nivel de experiencia en la evaluación de usabilidad de sistemas interactivos que conozcan las características del dominio del sistema y estén disponibles. Ahora bien, según Nielsen [5], diferentes evaluadores encuentran distintos problemas de usabilidad, por lo cual la participación de varios de ellos (con diferentes niveles de experiencia) permite que se encuentre una mayor cantidad de problemas de usabilidad, lo cual contribuye a la eficiencia del método en ejecución.

4.4.4.3 Consideraciones adicionales

4.4.4.3.1 Ejecución de los MEU

Es importante resaltar que en la ejecución de los MEU se incluye información acerca de la preparación, desarrollo y análisis de resultados, incluyendo los problemas de usabilidad identificados en cada uno de ellos, por lo que la investigación no pretende ir más allá de esto, es decir, no se presentan recomendaciones para solucionar los problemas detectados en los sistemas interactivos objeto de estudio, ya que esto sobrepasa los límites de los objetivos de la investigación. La presente investigación se enfoca en los resultados obtenidos con cada método respecto a los problemas de usabilidad, y las características asociadas a cada uno, que lo harán más adecuado que otro para ciertos casos.

Considerando el estado actual de la literatura, son escasos los trabajos relacionados que incluyen aspectos de trabajo colaborativo a los procesos de evaluación de usabilidad. En [9] se han propuesto una serie de combinaciones de métodos para evaluar la usabilidad de aplicaciones web transaccionales, por lo que el presente trabajo contribuye a dicha propuesta pues contempla MEU adicionales y ha generado la especificación colaborativa de los métodos objeto de estudio. Adicionalmente, en [11] se destaca el escaso asesoramiento que existe sobre qué métodos de evaluación son apropiados para ejecutar según ciertas circunstancias, por lo que la selección de los métodos se basa en la experiencia y la pericia individual. En ese sentido, la presente investigación contribuye a los trabajos mencionados pues intenta proponer un conjunto de combinaciones de MEU (para diferentes escenarios) que serían un insumo inicial para los practicantes de la usabilidad.

En las diferentes áreas de aplicación el perfil de usuario utilizado en la ejecución de los métodos de prueba consiste en usuarios novatos (con poca/nula experiencia en el uso del sistema interactivo objeto de estudio), puesto que la experiencia previa con el sistema en evaluación puede influenciar en cómo los usuarios realizan las tareas. Los usuarios con más experiencia tienden a tener mayores porcentajes de finalización y completar las tareas con mayor rapidez que los usuarios con menos experiencia, lo cual influye de manera importante en los resultados obtenidos. Por tales razones, en las pruebas fue considerado únicamente el perfil de usuario novato con el fin de evitar la diversidad en los resultados obtenidos y poder contrastarlos con los métodos de inspección. Con base en lo anterior, como trabajo futuro se plantea realizar pruebas adicionales considerando otros perfiles de usuario.

Según los resultados de las métricas asociadas a la característica *detección de problemas de usabilidad*, los métodos interacción constructiva y conductor, a pesar de que presentan un nivel considerable de subjetividad, presentan resultados positivos por encima de los experimentos formales, lo cual podría indicar que se obtienen resultados (y realimentación) apropiados con métodos de evaluación que promueven una interacción directa con los usuarios (participantes de las pruebas).

En las métricas relacionadas a la característica *tiempo* es discutible si un valor “pequeño” constituye un hecho positivo pues podría dudarse de la calidad de los entregables generados en cada actividad. De igual forma, la velocidad de ejecución de las etapas de un método de evaluación (planeación, ejecución y análisis de resultados) no es un valor por sí misma si no se asegura una mínima calidad en los resultados que se deriven de ellas. Por tal razón, el *evaluador supervisor* (rol definido en la especificación colaborativa de los MEU) es el encargado de asegurar cierto grado de calidad en los resultados/entregables obtenidos.

En términos generales, los MEU obtuvieron resultados positivos respecto a la métrica TEA, puesto que los participantes (*evaluador supervisor* y *evaluadores expertos*) emplearon relativamente poco tiempo para completar las actividades que conforman la *etapa de análisis de resultados*. Con base en lo anterior, se podría decir que, para la *etapa de análisis de resultados*, las combinaciones de *thinklets* propuestas en la especificación colaborativa de los MEU funcionan de forma apropiada, pues estarían permitiendo obtener una serie de contribuciones por parte de los evaluadores en un tiempo razonable.

4.4.4.3.2 Herramientas software de soporte

En cuanto al uso de herramientas software, en la actualidad existe una gran cantidad de herramientas que integran diferentes funcionalidades para promover el trabajo colaborativo; estas permiten la colaboración y la comunicación de varias personas de forma concurrente, y pueden ser adaptadas a las necesidades de un conjunto de actividades que se van a realizar. Debido a que hay una gran cantidad de herramientas *groupware* [86] gratuitas existentes, la selección de una de estas para realizar un proceso resulta una tarea compleja, lo cual puede dar como resultado una elección inadecuada. Teniendo en cuenta lo anterior, en la ejecución de los MEU sobre los sistemas interactivos no fue utilizada una herramienta *groupware* específica para la realización de las actividades colaborativas que conforman los métodos. La decisión obedece a que una de las mayores dificultades suele ser convencer a los participantes del proceso de evaluación de que hagan uso de una herramienta específica, por lo que esta debe tener un alto grado de usabilidad, de lo contrario, no se utilizará. Además, los evaluadores están acostumbrados a utilizar herramientas software existentes (de comunicación, cooperación y colaboración) que funcionan de forma apropiada, son flexibles y sencillas de usar, tal es el caso de las herramientas de Google. Un aspecto a considerar es que si fuera utilizada una herramienta *groupware*, posiblemente deban realizarse capacitaciones a las personas para que aprendan a utilizar la herramienta, ya que este tipo de herramientas ofrece una gran variedad de funcionalidades.

Cabe resaltar que, como aporte adicional de este trabajo, han sido especificados una serie de requerimientos para la construcción de la herramienta software *ECUSI – Evaluación Colaborativa de la Usabilidad de Sistemas Interactivos*. ECUSI es una aplicación web que guía la ejecución de las actividades que conforman un conjunto de métodos de evaluación de usabilidad. Esto haciendo uso de documentos compartidos en Google Docs que permiten la participación de diferentes personas involucradas en el proceso de evaluación. ECUSI fue concebida con el objetivo de ofrecer a los practicantes de la usabilidad una herramienta que los oriente durante el proceso de evaluación de usabilidad de diferentes sistemas interactivos. ECUSI ofrece las siguientes funcionalidades generales: (i) permite consultar la especificación colaborativa de un conjunto de MEU (de inspección y prueba), (ii) crear proyectos de evaluación los cuales están conformados por la combinación de varios MEU, (iii) ofrecer los recursos necesarios para el desarrollo de las actividades (colaborativas y no colaborativas) que conforman los MEU seleccionados en un proyecto de evaluación. ECUSI se encuentra temporalmente disponible en la siguiente URL: <http://ecussi.juancamiloceron.com/>.

Capítulo 5

El valor del producto se halla en la producción. Albert Einstein

Metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos

Este capítulo presenta la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos*. Las características específicas de los MEU, la ejecución y análisis de resultados de los métodos realizados en las áreas de aplicación objeto de estudio (TDi, web transaccional y aplicaciones móviles), así como también los análisis comparativos entre ellos, han permitido establecer una serie de combinaciones de métodos de evaluación (de inspección y prueba) las cuales conforman la metodología propuesta. La Figura 22 presenta la estructura del presente capítulo.



Figura 22. Estructura del capítulo 5.

5.1 Características sobresalientes de los MEU

Teniendo en cuenta la experiencia obtenida mediante la ejecución de los MEU en las áreas de aplicación objeto de estudio, la Tabla 42 (que toma como referencia la Tabla 3 – Ventajas y desventajas de los métodos de evaluación de usabilidad, del Anexo B.3) presenta un conjunto de características sobresalientes de cada uno de los métodos. Esto con el fin de generar un pequeño acercamiento a los factores más influyentes que van a permitir definir los métodos de evaluación que conformarán la *metodología para la evaluación colaborativa*

de la usabilidad de sistemas software interactivos.

Tabla 42. Características sobresalientes de los MEU.

Métodos de evaluación	Áreas de aplicación: TDi, Web transaccional, Aplicaciones móviles Características
Evaluación heurística	<ul style="list-style-type: none"> • Fácil de conducir. • Amplia aplicabilidad. • Permite identificar un buen número de problemas de usabilidad (mayores y menores), priorizarlos e interpretar su criticidad. • Se pueden pasar por alto problemas específicos del dominio del sistema en caso de no contar con un conjunto de heurísticas específicas. • Alta inversión de tiempo en el análisis de resultados.
Recorrido cognitivo	<ul style="list-style-type: none"> • Detecta problemas relacionados con la facilidad de aprendizaje. • Requiere cierto entrenamiento por parte de los evaluadores.
Experimentos formales	<ul style="list-style-type: none"> • Objetividad. • Permite recolectar información cuantitativa. • Evaluación enfocada a escenarios. • Alta inversión de tiempo.
Cuestionarios	<ul style="list-style-type: none"> • Simples y económicos. • Baja inversión de tiempo. • Proporcionan información cuantitativa y cualitativa. • Fiabilidad de la información subjetiva.
Interacción constructiva	<ul style="list-style-type: none"> • Detecta una gran cantidad de problemas de usabilidad en el sistema. • Es más natural para los usuarios verbalizar conjuntamente. • Bajo tiempo en relación a otros métodos de interacción directa con el sistema. • La interacción de los usuarios no representa el uso del sistema en condiciones normales. • No se obtiene información cuantitativa.
Entrevistas	<ul style="list-style-type: none"> • Simples y económicas. • Baja inversión de tiempo. • Proporcionan información cuantitativa y cualitativa. • Buen complemento para la obtención de información subjetiva. • El evaluador (o entrevistador) puede controlar el curso de la entrevista. • Fiabilidad de la información subjetiva.
Método del conductor	<ul style="list-style-type: none"> • Permite analizar el modelo mental del usuario. • El evaluador (o conductor) puede controlar el curso de la prueba. • Bajo tiempo en relación a otros métodos de interacción directa con el sistema. • No representa el uso del sistema en condiciones normales (los sistemas evaluados no son comúnmente usados en compañía de una persona guía).

5.2 Metodología para la evaluación colaborativa de la usabilidad

5.2.1 Discusión

El camino más común para hacer evaluaciones de usabilidad es combinar métodos de inspección con métodos de prueba, dependiendo del escenario que se presente [65]. En ese sentido, la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos* claramente deberá incluir, por lo menos, uno de los métodos de inspección y uno de los métodos de prueba de usabilidad.

En primer lugar, respecto a los métodos de inspección, fueron estudiados los métodos: evaluación heurística y recorrido cognitivo. La evaluación heurística permite identificar un buen número de problemas de usabilidad mediante inspecciones al sistema en evaluación;

su capacidad para encontrar problemas de distintos niveles (mayores y menores) y que incumplen distintos principios de usabilidad apoya esta inclusión. En cuanto al recorrido cognitivo, considerando las características *detección de problemas de usabilidad y tiempo*, este no presentó medidas destacables en ninguna de las dos características, excepto en la métrica *Tiempo empleado para completar el análisis de resultados (TEA)*, puesto que empleó (en promedio) el tercer menor tiempo en el análisis de resultados. Con base en lo anterior, los resultados positivos respecto a la *detección de problemas de usabilidad* favorecen a la evaluación heurística. Por tal razón, dicho método de inspección será parte de la metodología de evaluación como el primero de los métodos a ejecutar.

La evaluación heurística se destaca por su gran capacidad para ayudar a encontrar problemas, mientras que su desventaja está en la influencia del dominio del sistema [87]. Los evaluadores pueden tener experiencia en este tipo de evaluaciones, pero si no conocen las reglas del negocio o no disponen de un conjunto de heurísticas específicas, probablemente varios problemas de usabilidad no podrán ser identificados. Una ventaja importante de la evaluación heurística es su facilidad de llevar a cabo en comparación con otros métodos de inspección [88], sin embargo, una desventaja importante radica en el tiempo requerido para su ejecución y análisis de resultados (ver sección 4.3.1.3.2). En la presente investigación fueron considerados solamente 2 métodos de inspección, pero la literatura eleva las capacidades de la evaluación heurística por sobre otros métodos, además de ser el método de inspección más utilizado para la detección de problemas [41].

En segundo lugar, respecto a los métodos de prueba, la metodología de evaluación debe incluir por lo menos uno de los MEU que realizan una interacción directa del usuario con el sistema, es decir, los experimentos formales, la interacción constructiva o el conductor. Este aspecto será discutido más adelante.

Entre los métodos de prueba que no consideran una interacción directa con el sistema se encuentran las entrevistas y cuestionarios. Estos métodos de interrogación permitieron identificar problemas relacionados al diseño, navegación y otros aspectos de los sistemas evaluados, además de obtener información relacionada a la satisfacción subjetiva de los usuarios luego de interactuar con el sistema. Esta última característica es clave al momento de proponer posibles combinaciones de MEU, debido a que dichos métodos de interrogación funcionan como complemento a los métodos de prueba que realizan la interacción directa con el sistema. Adicional a lo anterior, las entrevistas y cuestionarios resultan apropiados para complementar la ejecución de otros métodos de prueba pues su preparación, ejecución y análisis no demandan una cantidad significativa de tiempo (ver sección 4.3.1.3.2).

Las entrevistas y cuestionarios se destacan por su bajo costo para llevarlos a cabo [10]. En cuanto a sus desventajas, al no realizar interacciones directas con el sistema en evaluación, la información se torna ciertamente subjetiva [48]. De esto pueden librarse dichos métodos de interrogación si son realizados luego de que los usuarios hayan interactuado con el sistema mediante el desarrollo de otro de los métodos de prueba, tal como ocurrió en la presente investigación.

Considerando lo anterior, la discusión estará enfocada en la definición de él o los métodos de prueba intermedios que deben realizarse para obtener la mayor cantidad de información relevante acerca de la usabilidad del sistema interactivo evaluado (considerando tiempos razonables y obtención de problemas importantes). Esto se debe a que ha sido establecido como método inicial la evaluación heurística y como métodos finales las entrevistas o

cuestionarios.

En cada una de las áreas de aplicación objeto de estudio, los experimentos formales, la interacción constructiva y el método del conductor han permitido confirmar, de forma general, problemas críticos identificados en la evaluación heurística, además de encontrar algunos otros problemas. Sin embargo, hay algunas diferencias marcadas como las siguientes. Las pruebas realizadas en las diferentes áreas de aplicación indican que la interacción constructiva detectó una mayor cantidad de problemas de usabilidad respecto a los demás métodos de prueba estudiados, además, permitió confirmar empíricamente la mayor parte de los problemas críticos identificados en las evaluaciones heurísticas. Ahora bien, los experimentos formales y el método del conductor también lograron lo mismo que la interacción constructiva, pero a menor escala. Entre los dos métodos antes mencionados, el conductor permitió identificar una mayor cantidad de problemas en dos de las áreas de aplicación objeto de estudio (web transaccional y aplicaciones móviles). En términos generales, el método del conductor permitió identificar la mayor cantidad de problemas entre aquellos que utilizan una lista de tareas específicas. Por lo cual, sería apropiado utilizar dicho método cuando el objetivo es evaluar una serie de tareas, funcionalidades o áreas específicas de un sistema interactivo.

Respecto a la característica *tiempo* para los 3 métodos de prueba antes mencionados, los resultados obtenidos indican que el conductor y la interacción constructiva emplean menos tiempo, en total, que los experimentos formales. La interacción constructiva no necesita de un diseño de tareas para su realización, mientras que los experimentos formales y el método del conductor sí lo requieren, lo cual incrementa el tiempo empleado en la preparación de dichos métodos. Ahora, los experimentos formales demandan una mayor cantidad de tiempo en la *etapa de análisis de resultados*, principalmente por los estudios estadísticos que son realizados, el análisis de las acciones de los usuarios en cada tarea, revisión del tiempo demorado en alcanzar cada tarea, entre otras actividades. De esta manera, la interacción constructiva y el método del conductor se destacan entre los métodos que incluyen una interacción directa entre el usuario y el sistema en evaluación.

Otro factor importante al establecer la *metodología para la evaluación colaborativa de la usabilidad de sistemas interactivos* consiste en el alcance de la evaluación, es decir, las áreas/funcionalidades del sistema que serán evaluadas. La interacción constructiva no limita a la pareja de usuarios a enfocarse en funcionalidades específicas del sistema, por el contrario permite realizar un análisis completo de estas, por lo cual son detectados muchos más problemas de usabilidad que los experimentos formales y método del conductor. Claramente, no ocurre lo mismo con los experimentos formales y el conductor. El listado de tareas asociado a estos métodos restringe la interacción a las partes que el diseño de tareas sugiere, lo cual no es una desventaja, pero evidentemente no permite una evaluación completa (global) del sistema, sino de un conjunto de funcionalidades específicas de este. Cabe mencionar que el listado de tareas asociado al método del conductor es más flexible que el de los experimentos formales, puesto que el evaluador (o conductor) puede controlar el curso de la prueba según cómo el usuario complete las tareas propuestas.

5.2.2 Metodología propuesta

Con base en los argumentos antes presentados, la metodología de evaluación no se plantea como una solución que tenga que adoptarse como un todo, por el contrario, se trata de una propuesta que incluye un conjunto de combinaciones de MEU que en algunos casos (o escenarios) pueden resultar de utilidad. A continuación es presentada la *metodología*

propuesta para la evaluación colaborativa de la usabilidad de sistemas software interactivos, la cual está conformada por 3 combinaciones distintas de MEU, con base en los objetivos de la evaluación.

5.2.2.1 Evaluación global: alta detección de problemas

Si el objetivo de la evaluación es analizar un sistema de manera completa, la metodología para la evaluación global de la usabilidad de sistemas interactivos incluye los métodos: evaluación heurística, interacción constructiva y métodos de interrogación (entrevistas y/o cuestionarios). La Figura 23 presenta la secuencia de MEU para realizar la evaluación global de la usabilidad.

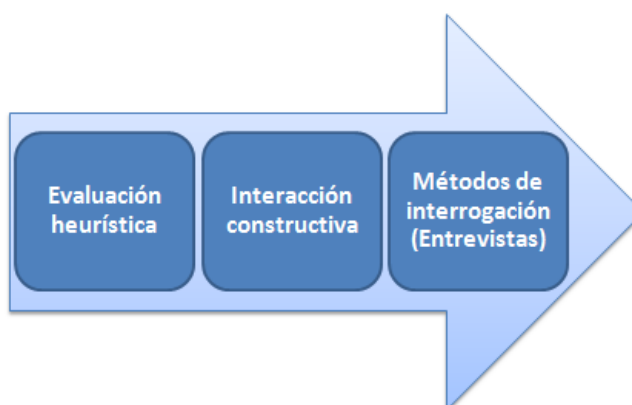


Figura 23. Evaluación global de la usabilidad de sistemas interactivos.

En un principio, la evaluación heurística permite identificar gran parte de los problemas de usabilidad con la participación de evaluadores expertos. Luego, la interacción constructiva permite confirmar varios de esos problemas con la participación de usuarios representativos que exploran libremente las funcionalidades del sistema, por lo que, al igual que la evaluación heurística, se evalúa el sistema completamente. Finalmente, los métodos de interrogación permiten recolectar subjetivamente información acerca de la experiencia de los usuarios, funcionando como un eficiente complemento. En esta combinación de MEU se sugiere utilizar las entrevistas estructuradas como método de interrogación, puesto que permiten confirmar problemas generales identificados en la evaluación heurística. Además, conviene ejecutar las entrevistas después de haber realizado la interacción constructiva teniendo en cuenta el carácter subjetivo de estos métodos.

En esta evaluación la interacción constructiva posee dentro de sus ventajas la posibilidad de obtener resultados con la participación de pocos (6 o más) usuarios representativos. Esto tiene su fundamento por la expresión de las impresiones de los usuarios mientras realizan la interacción. Al trabajar los usuarios en parejas la fluidez de los comentarios es mayor, en consecuencia, podría identificarse una mayor cantidad de problemas de usabilidad. Un factor importante también está en el tiempo, la interacción constructiva requiere menor cantidad de tiempo que los experimentos formales para su preparación y análisis de resultados. Por otra parte, la interacción constructiva permite identificar las razones/causas de los problemas, lo cual es una ventaja importante ya que contribuye en la confirmación de problemas críticos detectados en la evaluación heurística.

La combinación de MEU, propuesta para la evaluación global de un sistema interactivo, coincide con aquellas combinaciones planteadas en los trabajos de investigación [1] y [47],

en los cuales se establecen propuestas metodológicas para evaluar la usabilidad de aplicaciones web transaccionales y de TDi, respectivamente. Esto confirma que la evaluación heurística y la interacción constructiva son métodos que permiten una alta detección de problemas de usabilidad.

La Tabla 43 presenta un resumen de la combinación: *evaluación global*, con el fin de que practicantes de la usabilidad identifiquen la aplicabilidad de dicha combinación.

Tabla 43. Resumen de la combinación: evaluación global.

Objetivo de la evaluación	Evaluación global de un sistema interactivo.
Métodos	Evaluación heurística + Interacción constructiva + Entrevistas
Ventajas	Permite una alta detección de problemas de usabilidad correspondientes a diferentes funcionalidades del sistema.
Requerimientos	Prototipo funcional.
Etapas del proceso de desarrollo	La combinación de métodos puede ser aplicada en distintas etapas del desarrollo de un sistema, tanto en etapas tempranas como en etapas intermedias y finales, por lo que no es necesaria la disponibilidad de una versión final del sistema a evaluar.
Número de evaluadores	Se recomienda que el número mínimo de evaluadores sea 3, máximo 5.
Número de usuarios	Se recomienda ejecutar los métodos de prueba con 10 o más usuarios.
Especificación colaborativa	Evaluación heurística: ver sección 3.4.2. Interacción constructiva: ver Anexo H.5. Entrevistas: ver Anexo K.5.

5.2.2.2 Evaluación específica: reducción de tiempo

Si el objetivo es evaluar ciertos escenarios o funcionalidades de un sistema, y considerando que el factor tiempo es comúnmente crítico, el método del conductor sería el más adecuado puesto que el evaluador (quien asume el rol de conductor) asegurará la evaluación empírica sobre lo deseado, mediante una lista de tareas diseñada especialmente para ello. Los experimentos formales también hacen uso de un diseño de tareas, sin embargo, requieren una alta inversión de tiempo para su preparación y análisis de resultados. Si el sistema tiene un gran número o variedad de funcionalidades, la interacción constructiva probablemente no alcance a cubrir todo el sistema y podría pasarse por alto lo que realmente se quiere evaluar. En ese sentido, la metodología para la evaluación específica de la usabilidad de sistemas interactivos incluye los métodos: evaluación heurística, conductor y métodos de interrogación (entrevistas y/o cuestionarios). Para esta combinación de MEU se sugiere realizar cuestionarios con el fin de obtener información acerca de la percepción general de los usuarios sobre la experiencia en el uso del sistema evaluado. La Figura 24 presenta la secuencia de MEU para realizar la evaluación específica.



Figura 24. Evaluación específica de la usabilidad de sistemas interactivos.

En esta combinación de MEU la evaluación heurística identifica problemas de diseño o detalles de presentación que pueden impedir el progreso de los usuarios al desarrollar una tarea, el cual es un enfoque más específico respecto a la evaluación heurística planteada en la *evaluación global*. Mediante el método del conductor es posible identificar las diferencias existentes entre el modelo conceptual del sistema y el modelo mental de los usuarios. En este método los usuarios de la prueba desarrollan las tareas solicitadas siguiendo su modelo mental, lo que, en muchas ocasiones, genera dichas diferencias por el incorrecto modelado del sistema, que impide o dificulta la realización de las tareas por parte de los usuarios.

Un aspecto positivo del método del conductor es que puede ser controlado por el *evaluador supervisor*, de tal manera que pueda enfocarse en las acciones que resultan complejas para los usuarios. Además, durante la ejecución del método, es posible discutir con el usuario acerca de acciones/reacciones del sistema no esperadas. Por otro lado, en este método se tiene una interacción directa con los usuarios lo cual permite obtener una realimentación “más” completa, con lo que sería posible identificar una mayor cantidad de problemas. Con base en lo anterior, el conductor es un método que resultaría beneficioso para los practicantes de la usabilidad puesto que permite obtener realimentación (en vivo) de los usuarios, así como también información acerca del cumplimiento de las tareas.

La Tabla 44 presenta un resumen de la combinación: *evaluación específica*, con el fin de que practicantes de la usabilidad identifiquen la aplicabilidad de dicha combinación.

Tabla 44. Resumen de la combinación: evaluación específica.

Objetivo de la evaluación	Evaluación de un conjunto de funcionalidades/áreas de un sistema interactivo.
Métodos	Evaluación heurística + Método del conductor + Cuestionarios
Ventajas	Evaluación empírica sobre los aspectos deseados. Reducción de tiempo en comparación a la ejecución de experimentos formales.
Requerimientos	Prototipo funcional.
Etapas del proceso de desarrollo	La combinación de métodos puede ser aplicada en distintas etapas del desarrollo de un sistema, tanto en etapas tempranas como en etapas intermedias y finales, por lo que no es necesaria la disponibilidad de una versión final del sistema a evaluar.
Número de evaluadores	Se recomienda que el número mínimo de evaluadores sea 3, máximo 5.
Número de usuarios	Se recomienda ejecutar los métodos de prueba con 10 o más usuarios.
Especificación colaborativa	Evaluación heurística: ver sección 3.4.2. Método del conductor: ver Anexo I.5. Cuestionarios: ver Anexo J.5.

5.2.2.3 Evaluación enfocada a tareas específicas: sin restricciones de tiempo

Finalmente, se presenta una combinación de MEU que permitiría obtener información precisa y de carácter objetivo sobre la usabilidad de un conjunto de tareas específicas de un sistema interactivo. Sin embargo, dicha combinación está determinada por un factor importante que es la disponibilidad de tiempo. La *evaluación enfocada a tareas específicas*, a diferencia de la combinación: *evaluación específica*, considera que no hay restricciones de tiempo fuertes para realizar las evaluaciones, es decir, se dispone del tiempo necesario para realizar evaluaciones con el fin de mejorar en buena medida la usabilidad de una serie de tareas de un sistema interactivo.

En el caso de que haya una cantidad de tiempo considerable para realizar las evaluaciones (pero no infinita, claramente), la experiencia considera que deben ejecutarse los siguientes métodos: recorrido cognitivo, experimentos formales y métodos de interrogación (entrevistas y/o cuestionarios). El recorrido cognitivo encontraría problemas orientados a tareas. Los experimentos formales se basan en una lista de tareas que los usuarios deben desarrollar

durante la interacción con el sistema, por tal razón, dicho método permite un análisis objetivo y cuantitativo. Este análisis cuantitativo hace que los experimentos formales demanden más tiempo que los demás métodos de prueba. Adicionalmente, en esta combinación de MEU resulta conveniente utilizar los cuestionarios como método de interrogación puesto que, por su carácter objetivo, complementan de forma apropiada los experimentos formales. Se sugiere ejecutar los cuestionarios antes (pre-test) y/o después (post-test) de que los usuarios realicen los experimentos formales. La Figura 25 presenta la secuencia de MEU para realizar la evaluación enfocada a tareas específicas.



Figura 25. Evaluación enfocada a tareas específicas.

Mediante la ejecución de esta combinación de MEU se obtendrá información precisa sobre la usabilidad de un conjunto de tareas específicas, debido a que serán realizadas inspecciones por parte de evaluadores expertos, podrán realizarse análisis estadísticos con base en las acciones de los usuarios mediante los experimentos formales, y finalmente, se obtendrá información sobre las percepciones subjetivas (cualitativas y cuantitativas) de los usuarios representativos.

La Tabla 45 presenta un resumen de la combinación: *evaluación enfocada a tareas específicas*, con el fin de que practicantes de la usabilidad identifiquen la aplicabilidad de dicha combinación.

Tabla 45. Resumen de la combinación: evaluación enfocada a tareas específicas (sin restricciones de tiempo).

Objetivo de la evaluación	Evaluación de un conjunto de tareas específicas de un sistema interactivo.
Métodos	Recorrido cognitivo + Experimentos formales + Cuestionarios
Ventajas	Permite recolectar información cuantitativa. Objetividad en el análisis de resultados.
Requerimientos	Prototipo funcional.
Etapas del proceso de desarrollo	La combinación de métodos puede ser aplicada en distintas etapas del desarrollo de un sistema, tanto en etapas tempranas como en etapas intermedias y finales, por lo que no es necesaria la disponibilidad de una versión final del sistema a evaluar.
Número de evaluadores	Se recomienda que el número mínimo de evaluadores sea 3, máximo 5.
Número de usuarios	Se recomienda ejecutar los métodos de prueba con 10 o más usuarios.
Especificación colaborativa	Recorrido cognitivo: ver Anexo F.5. Experimentos formales: ver Anexo G.5. Cuestionarios: ver Anexo J.5.

5.2.3 Aplicabilidad de la metodología

Las combinaciones de métodos que conforman la metodología de evaluación pueden ser realizadas en, prácticamente, cualquier etapa del ciclo de desarrollo de un sistema

interactivo, aunque probablemente se adaptan mejor en etapas tempranas, cuando se tenga un prototipo funcional (no necesariamente una versión final) que permita efectuar pruebas con usuarios reales. Se sugiere que el sistema a evaluar tenga cierto grado de avance o funcionalidad, para que los expertos puedan evaluarlo de una mejor manera y puedan obtenerse resultados más completos.

Por otro lado, como fue mencionado en la sección 2.1.1, los sistemas interactivos (SI) se dividen en dos grupos: SI software y SI hardware. Estos tienen diferentes características o componentes. Los componentes de un sistema interactivo conforman todas las partes del sistema, ya sean partes software o hardware [17]. Los SI software son aquellos que incluyen algún tipo de software tales como: aplicaciones de escritorio, sitios web, interfaces de telefonía móvil, entre otros. Debe tenerse en cuenta que todos los SI hardware incluyen piezas de software, por ejemplo, un cajero electrónico o cierto aparato eléctrico, tienen componentes de hardware importantes, pero estos también nos ofrecen una interfaz de software que no puede olvidarse cuando se quiera llevar a cabo una evaluación de usabilidad del sistema.

De acuerdo a las evaluaciones de usabilidad realizadas en las áreas: TDi, web transaccional y aplicaciones móviles, una serie de componentes han sido identificados los cuales son susceptibles a una evaluación de usabilidad. Con base en el trabajo relacionado [17], las Tablas 46 y 47 presentan un conjunto de componentes software y hardware que pueden encontrarse en diferentes sistemas interactivos. En ese sentido, la metodología de evaluación puede ser utilizada para evaluar la usabilidad de sistemas interactivos que incluyen cualquiera de los componentes software y hardware listados en dichas tablas.

Tabla 46. Componentes software de un sistema interactivo.

Grupo general	Componentes específicos
Tipo de contenido	Formularios, tablas, listas, fechas, tiempos, valores numéricos, signos de moneda.
Información	Imágenes, noticias, gráficas, texto, formato, URL, abreviaciones, audio, nomenclaturas, colores, iconos.
Gestión de datos	Transmisión de información, formulario de registro, formulario de inicio de sesión, actualización de información, validación de información, recuperación y/o copia de seguridad de información.
Búsqueda	Formulario de búsqueda, resultados de búsqueda.
Área de navegación	Páginas, ventanas, títulos, cursor, atajos.
Salidas de emergencia	Sin componentes asociados.

Las interfaces hardware, como cualquier dispositivo físico, nos permiten interactuar con ellas. Algunos ejemplos son: paneles de ascensores, cajeros automáticos, cámaras fotográficas, grabadoras de vídeo, hornos microondas, etc. Cabe resaltar que dichos sistemas físicos son sistemas interactivos completos con un gran conjunto de componentes y una parte software. Con base en lo anterior, las interfaces hardware se clasifican en [17]: sistemas operativos, navegador, dispositivos de entrada/salida, ayudas técnicas de asistencia, audio, indicadores y ayuda y documentación. Así, cada categoría está dividida en varios componentes (ver Tabla 47) para permitir una clasificación completa de un SI hardware con el objetivo de facilitar su evaluación.

Tabla 47. Componentes hardware de un sistema interactivo.

Grupo general	Componentes específicos
Sistema operativo	Sin componentes asociados.
Navegador	Sin componentes asociados.

Grupo general	Componentes específicos
Dispositivos de entrada/salida	Impresora, tarjeta electrónica, mouse, pantalla, teclado, micrófono, escáner, kinect, control remoto, control de videojuego, joystick.
Ayudas técnicas de asistencia	Iconos, hardware de asistencia.
Audio	Volumen
Indicadores	Sin componentes asociados.
Ayuda y documentación	Sin componentes asociados.

Estas clasificaciones de componentes no son mutuamente excluyentes, se trata de una clasificación complementaria debido a que los componentes hardware y software de un sistema podrían encontrarse en un sistema interactivo concreto.

Adicional a los componentes software y hardware antes mencionados, la metodología de evaluación podría ser aplicada a sistemas interactivos que ofrecen un conjunto determinado de funcionalidades. Las funcionalidades corresponden con la descripción de todas las tareas que pueden realizarse mediante la interfaz del sistema interactivo y que pueden ser objeto para obtener un buen nivel de usabilidad [17]. A modo de ejemplo, se entiende como funcionalidad de una interfaz la realización de tareas como: buscar información, realizar una reserva, compartir un archivo, llenar formularios, imprimir comprobantes de una compra, etc.

Capítulo 6

No todo lo que cuenta puede ser cuantificado, y no todo lo que puede ser cuantificado cuenta. Albert Einstein

Evaluación de la metodología

La metodología propuesta para la evaluación colaborativa de la usabilidad de sistemas software interactivos fue evaluada considerando las redes sociales como área de aplicación. En primer lugar, fue realizada una encuesta con el fin de identificar, según la experiencia y conocimiento de un grupo de expertos, la combinación de MEU más apropiada a utilizar según un escenario específico de evaluación. Luego, en segundo lugar, fueron ejecutados los MEU que conforman la combinación sugerida por los expertos. En tercer lugar, fue ejecutada la combinación de MEU que hace parte de la metodología propuesta de acuerdo al escenario de evaluación. Finalmente, en cuarto lugar, fueron comparados los resultados obtenidos en las evaluaciones para determinar si la combinación propuesta (como parte de la metodología de evaluación) permite obtener resultados más adecuados que los métodos sugeridos por los expertos. La Figura 26 presenta la estructura del presente capítulo.

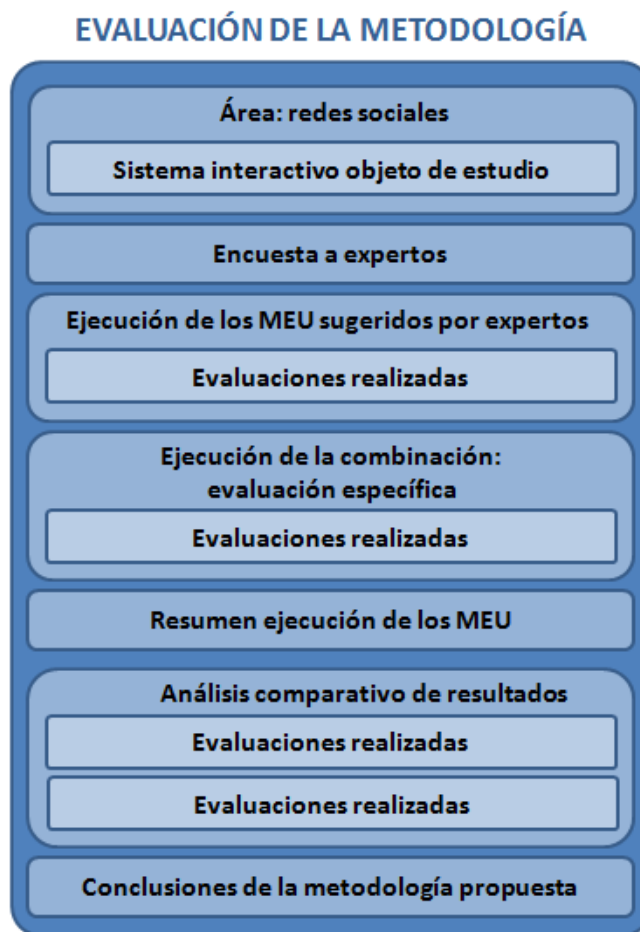


Figura 26. Estructura del capítulo 6.

6.1 Área: redes sociales

Una red social es definida como un servicio que permite a los individuos (1) construir un perfil público o semipúblico dentro de un sistema delimitado, (2) articular una lista de otros usuarios con los que comparten una conexión, y (3) ver y recorrer su lista de las conexiones y de las realizadas por otros dentro del sistema [89]. Las redes sociales se caracterizan por promover y generar contactos y oportunidades de comunicación, aprendizaje, trabajos, negocios, etc. Tienen como objetivo interconectar personas entre sí, de modo que cada uno de los miembros va formando su propia red de contactos de forma independiente, con los que puede comunicarse y compartir noticias e intereses [90]. Algunos ejemplos de redes sociales son: Facebook, Twitter, LinkedIn, Flickr, YouTube, entre otras.

Las redes sociales han sido seleccionadas como objeto de estudio puesto que en la actualidad tienen una importancia significativa, están creciendo y evolucionando a un ritmo vertiginoso. Estas se han convertido en sitios de interacción social y en herramientas promocionales para empresas, artistas, marcas, etc.

6.1.1 Sistema interactivo objeto de estudio

Considerando los criterios definidos en la sección 4.2, el sistema seleccionado es la red académica Researchgate.net, la cual está dirigida a la comunidad investigadora. Este sistema está disponible de forma gratuita, es de fácil acceso y presenta una cantidad apropiada de funcionalidades con buen nivel de navegabilidad.

Esta red académica pone a disposición del usuario una biblioteca de información que dispone de un motor de búsqueda semántica que navega por las principales bases de datos y repositorios universitarios. Pone al alcance más de 35 millones de artículos de interés, a menudo a texto completo, de revistas científicas. Los usuarios tienen la posibilidad de hacer sus propias contribuciones subiendo sus propias publicaciones y recomendándolas a otros usuarios. ResearchGate permite darles una mayor difusión y con ello conseguir un mayor impacto. La red también ofrece la opción de contactar con otros investigadores del mismo campo, debatir con ellos, acceder a blogs, la creación y participación de grupos de discusión privados o abiertos, acceder a información sobre eventos del mundo de la ciencia, una bolsa de trabajo internacional, etc. Esta información se encuentra disponible en el sitio web de la red social: www.researchgate.net.

6.2 Encuesta a expertos

En el proceso de evaluar la metodología propuesta, el primer paso consistió en realizar una encuesta (ver Anexo Q) entre un grupo de expertos para identificar la combinación de MEU que sería utilizada en un escenario específico de evaluación. En dicha encuesta se solicitaba a los participantes seleccionar un conjunto de métodos de evaluación (de inspección y de prueba) a ejecutar según el siguiente escenario:

*Usted es el evaluador en jefe de un proyecto para evaluar la usabilidad de una red social como es **ResearchGate.net**. La red ofrece diversas funcionalidades a los usuarios, sin embargo, a usted se le ha solicitado evaluar únicamente 6 funcionalidades específicas, esto con el objetivo de obtener información detallada acerca de las acciones de los usuarios en dichas funcionalidades.*

La encuesta fue diligenciada por 13 participantes con experiencia en evaluación de

usabilidad de sistemas interactivos. A continuación, la Tabla 48 presenta los porcentajes de votación obtenidos sobre los MEU consultados.

Tabla 48. Porcentajes de votación.

Método de evaluación	Porcentaje (%) de votación
Métodos de inspección	
Evaluación heurística	46
Recorrido cognitivo	61
Recorrido pluralista	7
Inspección de estándares	15
Análisis de acciones	23
Métodos de prueba	
Experimentos formales	15
Pensando en voz alta	61
Interacción constructiva	0
Grabación del uso	30
Método del conductor	7
Medida de prestaciones	7
Test retrospectivo	15
Cuestionarios	46
Entrevistas	30

Según los resultados de la Tabla 48, los MEU con porcentajes de votación más altos son: recorrido cognitivo, pensando en voz alta y cuestionarios. Por lo cual, dichos métodos son ejecutados (ver sección 6.3) para luego contrastar los resultados con los obtenidos por una de las combinaciones de MEU que hacen parte de la metodología propuesta.

Considerando los comentarios de los participantes de las encuestas, aproximadamente el 61% de ellos coincide en que el recorrido cognitivo permite a los evaluadores (desde la perspectiva de usuarios novatos) obtener una serie de problemas orientados a tareas. Este método combinado con el pensamiento en voz alta supone una buena combinación para intentar comprender el modelo mental de los usuarios mientras intentan realizar las tareas asociadas a funcionalidades específicas. Adicionalmente, entre los métodos de interrogación, los cuestionarios obtuvieron un porcentaje de votación más alto respecto a las entrevistas, así que estos serán ejecutados como complemento del método de prueba pensando en voz alta. La siguiente sección presenta la ejecución de los MEU mencionados.

6.3 Ejecución de los MEU sugeridos por expertos

Esta sección presenta la ejecución de los métodos de evaluación: recorrido cognitivo, pensando en voz alta y cuestionarios (ver Figura 27), los cuales han sido sugeridos por un grupo de expertos y corresponden a aquellos métodos que obtuvieron los porcentajes de votación más altos en la encuesta antes realizada (ver Tabla 48).

Para la ejecución de los métodos: recorrido cognitivo y cuestionarios, es utilizada la especificación colaborativa descrita en los Anexos F.5 y J.5, respectivamente. Sin embargo, para el caso del método pensando en voz alta no se cuenta con la especificación colaborativa, razón por la cual las actividades de la *etapa de planeación* se establecieron con base en el método del conductor, fueron definidas las actividades propias de la *etapa de ejecución*, y fue utilizado el proceso colaborativo definido para actividades recurrentes destacadas (ver Anexo D) en la *etapa de análisis de resultados*. Lo anterior obedece a que los métodos de prueba, en este caso el pensamiento en voz alta, siguen un proceso similar, lo que cambia principalmente es la dinámica de ejecución del método.

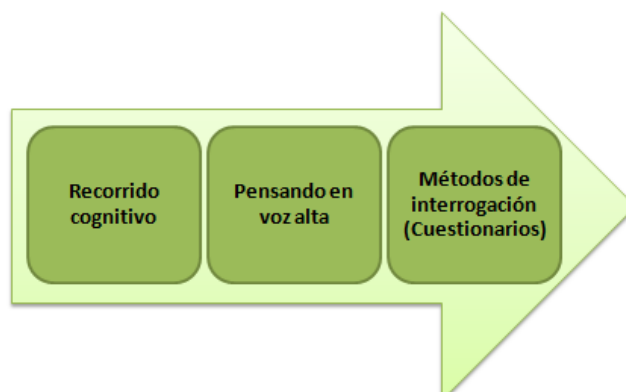


Figura 27. Secuencia de métodos sugerida por expertos.

6.3.1 Evaluaciones realizadas

Por restricción de extensión del documento, la información relacionada a la ejecución y análisis de resultados de los MEU sugeridos por expertos, es presentada en el Anexo R.

6.4 Ejecución de la combinación: *evaluación específica*

Teniendo en cuenta el escenario de evaluación bajo el cual fue realizada la encuesta a expertos (ver sección 6.2), la combinación de MEU a ejecutar es la *evaluación específica* (ver sección 5.2.2.2), la cual está conformada por los métodos: evaluación heurística, método del conductor y cuestionarios. En el escenario de evaluación se solicita evaluar un número específico de funcionalidades, por tal razón, y considerando que el tiempo es un factor clave para obtener resultados, se considera que la combinación *evaluación específica* es la más apropiada a ejecutar. Cabe destacar que el método del conductor obtuvo uno de los porcentajes de votación más bajos según la Tabla 48, lo cual es un aspecto relevante a considerar en el análisis comparativo que es realizado más adelante respecto a la combinación de métodos sugerida por expertos.

En la combinación *evaluación específica* la evaluación heurística está enfocada a detectar problemas que pueden impedir el progreso de los usuarios al utilizar alguna de las funcionalidades de interés. El método del conductor asegurará la evaluación empírica de las funcionalidades deseadas mediante una lista de tareas diseñada para ello. Por último, los cuestionarios van a permitir obtener información (cuantitativa) acerca de la satisfacción subjetiva de los usuarios respecto al uso de las funcionalidades evaluadas.

6.4.1 Evaluaciones realizadas

Por restricción de extensión del documento, la información relacionada a la ejecución y análisis de resultados de los MEU que conforman la combinación: *evaluación específica*, es presentada en el Anexo S.

6.5 Resumen ejecución de los MEU

La Tabla 49 presenta, en resumen, información acerca de la ejecución de los MEU (sugeridos por expertos y que conforman la *evaluación específica*) en el área de redes sociales. El análisis de resultados es presentado más adelante en la sección 6.6.

Tabla 49. Resumen de la ejecución de los MEU en el área de redes sociales.

Combinación	Método de evaluación	# Evaluadores	# Usuarios	# Problemas	# Problemas críticos confirmados	# Problemas no detectados en la evaluación heurística
Sugerida por expertos	Recorrido cognitivo	4	No aplica	18	11	9
	Pensando en voz alta	4	10	13	8	0
	Cuestionarios	4	10	0	0	0
Evaluación específica	Evaluación heurística	5	No aplica	30	11	No aplica
	Método del conductor	4	10	21	9	1
	Cuestionarios	4	10	0	0	0

Las siguientes secciones presentan el análisis de resultados considerando las características: *detección de problemas de usabilidad* y *tiempo* (definidas en la sección 4.1).

6.6 Análisis comparativo de resultados

Con base en la ejecución de los MEU (sugeridos por expertos y que conforman la *evaluación específica*) en el área de redes sociales, y considerando las condiciones de evaluación bajo las cuales estos fueron ejecutados, esta sección tiene como objetivo realizar un análisis comparativo de los resultados obtenidos en las combinaciones de métodos sugeridas. En primer lugar, son analizadas las métricas asociadas a la característica *detección de problemas de usabilidad*, y en segundo lugar, son analizadas aquellas relacionadas a la característica *tiempo*. Este análisis de resultados permitirá confirmar o desechar si la combinación *evaluación específica*, en el marco de la metodología de evaluación propuesta, permite obtener información más completa acerca de la usabilidad de un sistema interactivo.

El análisis de resultados es realizado directamente sobre las medidas de las métricas. Esto es, no son utilizadas tablas de normalización (como en la sección 4.4.4) debido a que los resultados corresponden a una misma área de aplicación, con lo cual no es necesario llevar las medidas a una misma escala de valores.

6.6.1 Característica: *detección de problemas de usabilidad*

Los problemas de usabilidad identificados por los MEU (sugeridos por expertos y que conforman la *evaluación específica*), se tornan como un factor importante a considerar para realizar las comparaciones necesarias y así confirmar o desechar una de las combinaciones propuestas en el marco de la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos*. En ese sentido, la Tabla 50 presenta las medidas de las métricas asociadas a la característica: *detección de problemas de usabilidad*.

Tabla 50. Medidas de las métricas asociadas a la característica: *detección de problemas de usabilidad*.

Combinación	Método de evaluación	Cantidad total de problemas (CTP)	Cantidad de problemas críticos (CPC)	Cantidad de problemas frecuentes (CPF)
Sugerida por expertos	Recorrido cognitivo	18 (16)	11	13
	Pensando en voz alta	13	8	10
	Cuestionarios	0	0	0

Combinación	Método de evaluación	Cantidad total de problemas (CTP)	Cantidad de problemas críticos (CPC)	Cantidad de problemas frecuentes (CPF)
Evaluación específica	Evaluación heurística	30 (23)	11	13
	Método del conductor	21	9	11
	Cuestionarios	0	0	0

En primer lugar, respecto a los métodos de inspección ejecutados, la evaluación heurística identificó un mayor número de problemas (30) en comparación al recorrido cognitivo (18). Por la naturaleza de la evaluación heurística de evaluar diferentes aspectos de un sistema, fueron identificados 7 problemas que no están completamente relacionados con las funcionalidades objeto de estudio (según el escenario de evaluación descrito en la sección 6.2). De esta manera, contemplando el escenario de evaluación, la evaluación heurística identificó un total de 23 problemas. De igual forma, en el recorrido cognitivo fueron detectados 2 problemas que no están relacionados con las funcionalidades evaluadas, por lo cual se considera que dicho método identificó un total de 16 problemas de usabilidad. En ese sentido, respecto a la métrica CTP, la evaluación heurística permitió la detección de 7 problemas adicionales en comparación al recorrido cognitivo.

Considerando las métricas CPC y CPF, es destacable que el recorrido cognitivo detectó los problemas críticos y frecuentes identificados en la evaluación heurística, lo cual pone en evidencia que el recorrido cognitivo detecta problemas de usabilidad relevantes. Sin embargo, su realización está sujeta altamente a la experiencia de los evaluadores pues estos requieren conocimientos y habilidades para detectar problemas de aprendizaje asociados a las funcionalidades o tareas evaluadas. En contraste, la evaluación heurística fue más sencilla de realizar ya que está sujeta a las heurísticas utilizadas y grado de detalle en los elementos (sub-heurísticas) que hacen parte de la lista de chequeo.

Con base en lo anterior, la evaluación heurística permitiría obtener información más completa (y de forma más sencilla) acerca de la usabilidad de un sistema software interactivo en comparación al recorrido cognitivo. Así, el proceso de evaluación revela que el método de inspección (evaluación heurística) incluido en la combinación *evaluación específica*, permite obtener resultados apropiados cuando es enfocado a un conjunto de tareas concretas.

En segundo lugar, respecto a los métodos de prueba, el método del conductor permitió identificar una mayor cantidad de problemas (valor más alto en la métrica CTP) en comparación al método pensando en voz alta. Sin embargo, respecto a las métricas CPC y CPF, la diferencia entre estos dos métodos es mínima. Esto sugiere que ambos métodos funcionan de forma adecuada confirmando problemas críticos del sistema identificados previamente por un método de inspección. En ese sentido, otros aspectos son analizados tal como se presenta a continuación.

Los métodos pensando en voz alta y conductor tienen en común que el usuario verbaliza sus impresiones durante la interacción con el sistema mediante una lista de tareas previamente diseñada. Ahora bien, el proceso de evaluación confirmó que el principal problema del pensamiento en voz alta es que las verbalizaciones del usuario interfieren de manera importante en el uso normal del sistema. Adicionalmente, este método depende en gran medida de la espontaneidad del usuario para expresar todas sus impresiones al realizar las tareas sugeridas. Es así como el método del conductor supera estas desventajas, puesto que durante su realización se establece una conversación más “normal” con el usuario. Sin

embargo, el punto en contra consiste en que, generalmente, no son dos usuarios los que interactúan con un sistema. El punto positivo está en que la realimentación obtenida mediante el método del conductor supone una mayor detección de problemas, puesto que el evaluador (quien asume el rol de conductor) tiene mayor control sobre el desarrollo de las pruebas, enfatizando sobre las tareas que no han podido completarse con éxito.

Teniendo en cuenta lo anterior, se estima que la realimentación obtenida con el método del conductor es de significativa relevancia en comparación a los registros obtenidos mediante el pensamiento en voz alta. En ese sentido, el proceso de evaluación revela que el método de prueba (conductor), incluido en la combinación *evaluación específica*, permitiría realizar análisis e interpretaciones más apropiadas acerca de las acciones e impresiones de los usuarios, con lo cual podría identificarse una mayor cantidad de problemas de usabilidad.

A partir de los métodos de inspección ejecutados, se destaca que la evaluación heurística puede ser realizada a un bajo costo, que depende del número de evaluadores “expertos” que realicen el proceso (de 3-5 como es sugerido para detectar la mayoría de problemas de usabilidad). Por tal razón, el costo resultaría menor que otros MEU que requieren de usuarios finales para ser llevados a cabo. Por otro lado, en cuanto a los métodos de prueba, es importante resaltar que el método del conductor obtuvo uno de los porcentajes de votación más bajos según la Tabla 48 (al igual que el método medida de prestaciones), mientras que el pensamiento en voz alta obtuvo el porcentaje de votación más alto. Esto indica que el método del conductor no es comúnmente utilizado en la práctica, por lo cual los resultados obtenidos en la presente investigación destacan el buen funcionamiento de dicho método y promueven su uso.

Por último, en tercer lugar, los cuestionarios no detectaron problemas de usabilidad, sin embargo, obtuvieron información complementaria acerca de la satisfacción subjetiva de los usuarios respecto al uso y realización de las tareas en el sistema. Razón por la cual, el proceso de evaluación revela que este método de interrogación funciona como buen complemento de los métodos de prueba para capturar datos cuantitativos y cualitativos.

6.6.2 Característica: *tiempo*

Respecto a la característica *tiempo* en el área de redes sociales, cabe mencionar que las actividades colaborativas que conforman los MEU ejecutados, fueron realizadas de forma virtual (utilizando documentos compartidos de Google Docs) debido a que los evaluadores se encontraban distribuidos geográficamente. Por lo cual, las contribuciones por parte de los evaluadores fueron recolectadas en un periodo de tiempo considerable (4 días en promedio).

La Tabla 51 presenta una aproximación del tiempo empleado (en minutos) por los participantes de los grupos en realizar las actividades que conforman las etapas de planeación, ejecución y análisis de resultados, de los MEU ejecutados.

Tabla 51. Medidas de las métricas asociadas a la característica: *tiempo*.

Combinación	Método de evaluación	Tiempo empleado para completar la etapa de planeación (TEP)	Tiempo empleado para completar la etapa de ejecución (TEE)	Tiempo empleado para completar el análisis de resultados (TEA)
Sugerida por expertos	Recorrido cognitivo	219	90	60
	Pensando en voz alta	151	34	92

Combinación	Método de evaluación	Tiempo empleado para completar la etapa de planeación (TEP)	Tiempo empleado para completar la etapa de ejecución (TEE)	Tiempo empleado para completar el análisis de resultados (TEA)
	Cuestionarios	82	10	62
Evaluación específica	Evaluación heurística	304	123	96
	Método del conductor	180	31	83
	Cuestionarios	79	10	60

En términos generales, respecto a los métodos de inspección, el recorrido cognitivo empleó menor tiempo que la evaluación heurística puesto que los valores de las métricas TEP, TEE y TEA son menores que los obtenidos en dicho método. El TEP es la métrica que presenta la mayor diferencia entre estos dos métodos, ya que la planeación de la evaluación heurística incluye algunas actividades que demandan un poco más de tiempo (acercamiento al sistema a evaluar, selección de heurísticas específicas adecuadas y elaboración del documento guía para los evaluadores). Con base en lo anterior, convendría seleccionar el recorrido cognitivo si el tiempo para realizar la evaluación es limitado, sin embargo, el tiempo empleado para realizar la planeación de la evaluación heurística podría reducirse debido a que en la literatura actual se han propuesto una serie de heurísticas específicas para evaluar diferentes sistemas interactivos, tales como: mundos virtuales [91], aplicaciones web transaccionales [77], aplicaciones móviles basadas en tecnología táctil [92], aplicaciones grid computing [93], aplicaciones de televisión digital interactiva [94], entre otros. De esta manera, si el sistema a evaluar corresponde a alguno de los mencionados, se estima que la planeación resulta menos demorada y se incrementa la posibilidad de identificar un mayor número de problemas de usabilidad (sin necesidad de tener expertos de alto nivel).

Siguiendo con las métricas TEE y TEA en los métodos de inspección, el proceso de evaluación nuevamente revela que el método más demorado en la *etapa de ejecución* es la evaluación heurística. Este método, a pesar de ser simple para desarrollar, necesita de un período de tiempo considerable para que los evaluadores inspeccionen el sistema tranquilamente, realicen sus contribuciones con base en el análisis de la información (rankings de criticidad, severidad y frecuencia), entre otras actividades. Sin embargo, los resultados obtenidos por este método podrían ser un factor importante para la persona responsable de la evaluación, haciendo parecer que el factor tiempo no fuera determinante.

Ahora bien, respecto a los métodos de prueba, la métrica TEP presenta la mayor diferencia entre el método del conductor y el pensamiento en voz alta, beneficiando a este último. Esto se debe principalmente a que en el método del conductor (según la especificación colaborativa definida para este método) el *evaluador supervisor* (que asume el rol de conductor) planea un escenario específico para la evaluación y se familiariza con la secuencia de acciones asociadas a las tareas que realizará el usuario, para luego cuestionarse sobre aquellas acciones que son realizadas durante la prueba.

Por otro lado, respecto a las métricas TEE y TEA, la diferencia entre los dos métodos de prueba (pensando en voz alta y conductor) es bastante cerrada. La duración de una sesión de prueba en el pensamiento en voz alta depende en gran medida de la espontaneidad del usuario para expresar sus impresiones. De esta manera, si el evaluador supervisor no motiva al usuario a verbalizar la mayor cantidad de información, los resultados no serían apropiados para analizar. En contraparte, si el usuario verbaliza la mayor cantidad de información,

incrementaría el tiempo de la sesión de prueba, y en consecuencia, los evaluadores podrían emplear una mayor cantidad de tiempo analizando e interpretando las impresiones del usuario (con el objetivo de comprender su modelo mental). Teniendo en cuenta lo anterior, en el método del conductor el análisis de una sesión de prueba demandaría menor cantidad de tiempo debido a que los evaluadores se enfocan en las acciones que el usuario no pudo completar en la consecución de una tarea.

En cuanto a los cuestionarios, como método de interrogación complementario a los métodos de prueba, estos emplearon la menor cantidad de tiempo en las etapas que conforman el proceso de evaluación (planeación, ejecución y análisis de resultados). En general, la preparación de los cuestionarios consta de pocas actividades, disponen de la información inmediatamente, con lo cual es posible analizar rápidamente las respuestas de los usuarios. Cabe mencionar que el poco tiempo empleado en el análisis de resultados estuvo relacionado a la disponibilidad limitada de los evaluadores para participar en las actividades colaborativas. Por tal razón, en el análisis e interpretación de la información el *evaluador supervisor* tuvo una mayor participación.

6.7 Conclusiones de la metodología propuesta

En el capítulo 5 fueron propuestas 3 combinaciones de MEU (global, específica y enfocada a tareas) las cuales conforman la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos*. Luego (en el capítulo 6), mediante la evaluación de usabilidad de un sistema correspondiente al área de redes sociales, se procedió a evaluar una de las combinaciones propuestas teniendo en cuenta un escenario específico de evaluación, intentando verificar que la combinación de MEU (en este caso la *evaluación específica*) permitiera obtener resultados más adecuados en comparación a una combinación de métodos sugerida por expertos en el tema de evaluación de usabilidad de sistemas interactivos. Finalmente, fue realizado un análisis comparativo de resultados (ver sección 6.6) con el objetivo de confirmar o desechar la combinación de MEU en estudio.

La metodología de evaluación ha sido propuesta considerando las características *detección de problemas de usabilidad y tiempo*. Una serie de métricas, correspondientes a cada característica, fueron estudiadas para observar el comportamiento de los MEU objeto de estudio en diferentes áreas de aplicación. Como trabajo futuro, si se desea proponer combinaciones adicionales de MEU considerando otros factores/características (como por ejemplo: costo económico), se sugiere realizar un proceso como el que fue llevado a cabo en la presente investigación. El cual consiste básicamente en estudiar el comportamiento de los MEU según un conjunto de métricas asociadas a la nueva característica en diferentes áreas de aplicación, esto con el fin de intentar generalizar los resultados obtenidos.

Las combinaciones de MEU propuestas fomentan el desarrollo sistemas interactivos evolutivos, con lo cual dichas combinaciones pueden ser utilizadas de forma iterativa en, prácticamente, cualquier etapa del ciclo de desarrollo del sistema cuando se tenga un prototipo funcional (no necesariamente una versión final) que permita efectuar pruebas con usuarios reales, aunque el grado de formalidad necesaria varía, según el prototipo disponible para evaluar en cada momento, y según las limitaciones de tiempo y recursos. Tal como se menciona en la sección 5.2.3, lo sugerido es que el sistema a evaluar tenga cierto grado de avance o funcionalidad, para que los expertos puedan evaluarlo de una mejor manera y puedan obtenerse resultados más completos. Esto porque la evaluación de una versión muy temprana del sistema (como podrían ser prototipos en papel), no permitiría obtener

resultados satisfactorios.

La primera combinación de MEU propuesta: *evaluación global*, enfocada a analizar un sistema de manera completa, incluye los métodos: evaluación heurística, interacción constructiva y entrevistas. Esta combinación se estima que funciona correctamente cuando es requerido un análisis de tipo global, mediante el cual será identificado un buen número de problemas de usabilidad, tanto por parte de los evaluadores como del análisis de la interacción de los usuarios representativos. Las entrevistas, como método de interrogación complementario, van a permitir obtener información adicional/complementaria acerca de la percepción de los usuarios respecto al sistema evaluado, con lo cual también sería posible confirmar problemas críticos identificados por los dos métodos anteriormente realizados.

La segunda combinación de MEU propuesta: *evaluación específica*, enfocada a evaluar ciertos escenarios o funcionalidades de un sistema, incluye los métodos: evaluación heurística, método del conductor y cuestionarios. Esta combinación, única evaluada entre las 3 combinaciones propuestas, resulta ser de gran utilidad para evaluar funcionalidades específicas debido a que la información obtenida mediante el método del conductor permite detectar problemas en aquellos puntos donde el usuario solicita ayuda/información al evaluador (conductor). Esto es, en los puntos donde hay comunicación entre el usuario y conductor es muy probable que haya necesidades de información en el sistema. En cuanto a la evaluación de esta combinación, la *evaluación específica* obtuvo resultados más adecuados que una serie de métodos propuestos por expertos, teniendo en cuenta un escenario específico de evaluación.

En las dos primeras combinaciones propuestas (*evaluación global y específica*), los métodos: interacción constructiva y conductor, posiblemente son llevados a cabo en un contexto que no es el más real, ya que por ejemplo, en un ambiente de TDi o aplicaciones móviles, no interactúan dos usuarios al mismo tiempo con una aplicación (a través del mando a distancia o dispositivo móvil). Sin embargo, son métodos que funcionan muy bien y permiten recolectar buena cantidad de información acerca de las percepciones subjetivas de los usuarios, puesto que ellos se sienten con más confianza para expresar sus impresiones a viva voz. Con lo cual, en áreas de aplicación diferentes a las estudiadas en la presente investigación, dichos métodos podrían obtener resultados más apropiados respecto a otros métodos "clásicos" en el tema de evaluación de usabilidad (como por ejemplo: el pensamiento en voz alta).

La tercera combinación de MEU propuesta: *evaluación enfocada a tareas específicas*, tiene como objetivo analizar tareas concretas de un sistema interactivo. Esta combinación incluye los métodos: recorrido cognitivo, experimentos formales y cuestionarios. En esta combinación los tres métodos aportan sus importantes características, pero son los experimentos formales los que marcan la diferencia respecto a la *evaluación específica*. Los experimentos formales permiten realizar un eficiente análisis de las tareas de interés. Estos tienen un buen nivel de objetividad y se complementan de forma adecuada con los cuestionarios (pre-test y post-test), que también presentan buena objetividad y permiten obtener información cuantitativa. Así, al ejecutar los 3 métodos que conforman esta combinación, se estima que la información obtenida acerca de la usabilidad de las tareas sería completamente objetiva. Sin embargo, esta combinación conviene utilizarla cuando la disponibilidad de tiempo es alta.

Teniendo en cuenta las 3 combinaciones propuestas, las actividades que conforman la *etapa*

de planeación de los MEU podrían realizarse de forma paralela. Sin embargo, la *etapa de ejecución* deberá ejecutarse de manera secuencial. Esto es, en cada combinación hay métodos de prueba intermedios, por lo cual estos deben ser ejecutados una vez sea realizado el análisis e interpretación de los resultados y sean identificados los problemas críticos en los métodos de inspección. De igual forma, los métodos de interrogación (entrevistas y cuestionarios) deben ejecutarse una vez haya terminado el método de prueba intermedio, excepto en el caso de los cuestionarios pre-test que son realizados antes de los experimentos formales.

Los métodos de interrogación (entrevistas y cuestionarios) están presentes al final de las 3 combinaciones propuestas. El proceso de evaluación ha permitido confirmar que con estos métodos se obtienen resultados cuantitativos y cualitativos que confirman problemas identificados por otros métodos de evaluación, claro está, si son bien diseñados. Razón por la cual, en la presente investigación se ha propuesto la especificación colaborativa para cada uno de estos métodos. Adicionalmente, las entrevistas y cuestionarios también permiten obtener información subjetiva que con otros métodos difícilmente es posible recolectar, lo cual hace que funcionen muy bien como complemento para otros métodos de prueba.

Las combinaciones de MEU permitirían cubrir los puntos más críticos para obtener la medida de la usabilidad con un nivel de precisión aceptable. Primero, existe, por lo menos, un método de inspección y un método de prueba. En segundo lugar, incluyen evaluaciones que realizan análisis cuantitativos y cualitativos, objetivos y subjetivos, y de evaluación global y evaluación específica del sistema. Por esto, es posible decir que se están cubriendo todos los factores necesarios para evaluar, en buena medida, la usabilidad de un sistema software interactivo.

Las conclusiones derivadas del proceso de evaluación con la combinación de MEU: *evaluación específica*, han sido exitosas en el sentido que permitió la detección de una mayor cantidad de problemas respecto a una combinación sugerida por personas con experiencia en la evaluación de usabilidad de sistemas interactivos. El método de prueba (conductor), que hace parte de la *evaluación específica*, ha confirmado un buen número de los problemas críticos identificados por el método de inspección (evaluación heurística). La forma en cómo se obtiene esa información es distinta respecto a otros métodos de prueba, lo importante es que ha cumplido con lo esperado. Ahora bien, respecto al factor tiempo, la combinación *evaluación específica* emplea mayor tiempo en la *etapa de planeación* que la combinación sugerida por expertos, sin embargo, los tiempos empleados en la ejecución y análisis de resultados no presentan una diferencia significativa.

En resumen, la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos* ha tenido una base firme para su postulación mediante la realización de casos de aplicación en diferentes áreas (televisión digital interactiva, web transaccional y aplicaciones móviles), y una de las combinaciones de MEU que la conforman ha sido evaluada exitosamente mediante un caso de aplicación final. Así, el resultado de la investigación consiste en 3 combinaciones de MEU (especificados de forma colaborativa) que conforman la metodología de evaluación, en la cual una de las combinaciones fue evaluada empíricamente según un escenario específico de evaluación. De esta manera, la metodología de evaluación está conformada por las siguientes combinaciones de MEU:

- **Evaluación global (alta detección de problemas):** Evaluación heurística – Interacción constructiva – Entrevistas.

- **Evaluación específica (reducción de tiempo):** Evaluación heurística – Método del conductor – Cuestionarios.
- **Evaluación enfocada a tareas específicas (sin restricciones de tiempo):** Recorrido cognitivo – Experimentos Formales – Cuestionarios.

Capítulo 7

Para el investigador no existe alegría comparable a la de un descubrimiento, por pequeño que sea. Sir Alexander Fleming

Experiencias y recomendaciones

Este capítulo incluye las experiencias y recomendaciones obtenidas gracias a la realización de la presente investigación.

Especificación colaborativa de los métodos de evaluación de usabilidad

Las actividades generales que conforman un proceso conviene agruparlas o clasificarlas, por ejemplo, las actividades que conforman los MEU diseñados de forma colaborativa, fueron agrupadas en 3 etapas: la primera etapa (planeación) incluye las actividades de preparación que son realizadas antes de ejecutar el método, la segunda etapa (ejecución) incluye las actividades realizadas por los evaluadores (y usuarios, en el caso de los métodos de prueba) una vez se encuentran en el lugar donde será ejecutado el método, y la tercera etapa (análisis de resultados) incluye las actividades de análisis e interpretación de la información recolectada durante la ejecución del método. La clasificación de las actividades permite diseñar de forma más estructurada y ejecutar de forma más ordenada un proceso especificado de forma colaborativa. Ahora, las actividades iniciales (de la *etapa de planeación*, por ejemplo) deben ser realizadas cuidadosamente, ya que la continuación y el progreso de las demás actividades se encuentran sujetos a estas, a las condiciones de tiempo y recursos.

Durante el diseño de las actividades colaborativas, en la tercera fase del *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software (Establecer tareas colaborativas)* inicialmente fueron asociados a las actividades colaborativas tres o cuatro patrones de colaboración (entre los cuales están: *generación, reducción, organización y evaluación*). Dado que los expertos involucrados en el proceso de evaluación muchas veces tienen restricciones de disponibilidad, es posible ignorar el *thinklet* asociado al patrón de colaboración: *evaluación*. Esto ocurre porque en la ejecución de los *thinklets* asociados a los patrones de colaboración: *generación, reducción u organización*, estaría implícito el proceso de evaluación. La recomendación anterior aplica para actividades colaborativas que tienen asociados varios patrones de colaboración (incluido el patrón de colaboración: *evaluación*) y en su realización participan personas con un alto nivel de experticia.

Respecto a la segunda y tercera fase del modelo, *Descomposición de la técnica y Establecer tareas colaborativas*, respectivamente, es posible integrar estas fases en la medida que sean detalladas las actividades que conforman un proceso determinado. Esto es, si en la segunda fase son identificadas minuciosamente las actividades que conforman el proceso, no sería necesario definir las subactividades en la tercera fase.

En la cuarta fase del modelo, *Relación con proceso colaborativo*, algunas actividades colaborativas tuvieron asociados los *thinklets*: *FreeBrainstorm* y *FastFocus*, en los patrones de colaboración: *generación y reducción*, respectivamente. Esta combinación de *thinklets* resulta demorada, y en la mayoría de los casos los evaluadores disponen de tiempo limitado, por lo cual se recomienda utilizar la combinación de *thinklets*: *OnePage* y *Pin the Tail on the*

Donkey, también pertenecientes a los patrones de colaboración: *generación* y *reducción*, respectivamente. Estos *thinklets* permiten lograr el objetivo de la actividad en menor tiempo, aunque puede que no todos los miembros del grupo participen en la misma medida.

Respecto a la sexta fase del modelo, *Validación*, en varios casos se llegó a la conclusión que en las actividades que tenían asociado el patrón de colaboración: *evaluación*, y que adicionalmente tenían asociados los patrones: *generación* y *reducción*, los resultados obtenidos no presentaban gran relevancia con el uso del patrón *evaluación*, en comparación al esfuerzo requerido para ejecutar el *thinklet* asociado a dicho patrón. De igual forma, en actividades que tenían asociados los patrones: *generación* y *reducción*, el patrón *reducción* fue eliminado, ya que durante el proceso de generación se tiende a obtener información definitiva por parte de los evaluadores.

La validación del diseño colaborativo de los MEU mediante la *Simulación* y *Prueba Piloto*, permitió probar la lógica del diseño, corroborar que a cada paso se le crearán los entregables requeridos y verificar si la ejecución de los métodos puede llevarse a cabo con los recursos definidos. Así, una vez terminado el proceso de evaluación, fueron realizados ajustes a la especificación colaborativa de los MEU estudiados en la presente investigación. Cabe mencionar que este fue un proceso transversal a la ejecución de los MEU en las áreas de aplicación objeto de estudio, ya que luego de terminar las evaluaciones, eran realizados ajustes a la especificación colaborativa de los métodos.

Los integrantes del grupo de trabajo de la presente investigación consideraron que las dos formas de validación del diseño seleccionadas: *Simulación* (inspección del proceso) y *Prueba Piloto* (ejecución del proceso), se complementan de forma apropiada y que con la aplicación de estas dos, es suficiente para la validación preliminar de la especificación colaborativa de los MEU que hacen parte de la metodología de evaluación propuesta.

Trabajo colaborativo en tiempo real

La conformación de equipos no es una tarea fácil, cuantas más personas participan en la realización de las actividades, mayor tiende a ser la dispersión de ideas y la comunicación resulta más difícil. Adicionalmente, la coordinación de los equipos suele ser dispendiosa, debido a las diferencias entre los distintos modelos mentales que confluyen en un mismo entorno [7]. Sin embargo, fomentar el trabajo colaborativo entre los integrantes de un grupo para analizar la información recolectada a partir de la ejecución de un conjunto de MEU, genera ciertas ventajas, tales como: identificar un mayor número de problemas de usabilidad, disminuir el esfuerzo de una sola persona al analizar e interpretar la información y evitar que ese análisis esté sesgado a la perspectiva individual, se aprende de los aportes y experiencia de los miembros del grupo.

En un contexto real no es tarea fácil conseguir evaluadores con experiencia en la evaluación de usabilidad de software, que conozcan las características del dominio del sistema y estén disponibles. Por tal razón, un problema frecuente durante la realización del presente proyecto se presentaba en el momento de realizar las actividades que conforman la *etapa de análisis de resultados* de los MEU diseñados de forma colaborativa. Este problema consistió en que algunos evaluadores disponían de tiempo limitado para participar en las actividades de análisis e interpretación de la información, lo cual ocasionó retrasos en la ejecución de las actividades. Además, esto provocó que en algunos casos no todos los evaluadores del grupo participaran en la misma medida. A lo anterior, se le suma el hecho de que todos los

evaluadores actuaron como voluntarios sin remuneración.

Durante la ejecución de los MEU en el área de TDi, los evaluadores tuvieron una participación activa en el análisis de la información recolectada. La remuneración y participación en un proyecto de investigación (asociado al presente trabajo) los motivaba a participar para obtener resultados a corto plazo, lo cual no sucedió en las demás áreas de aplicación objeto de estudio. En las áreas web transaccional y aplicaciones móviles el grado de compromiso de los evaluadores no fue el ideal, lo cual se vio reflejado en las escasas contribuciones asociadas a la información recolectada en los MEU. Lo anterior pudo ocurrir por la poca disponibilidad de tiempo de los evaluadores para participar en las actividades colaborativas, además que no hubo un componente adicional (financiero, por ejemplo) que los motivara a participar en mayor medida.

Respecto a las actividades de análisis e interpretación de la información recolectada en el área de TDi, cabe resaltar que las discusiones (pasos del *thinklet* asociado al patrón de colaboración *clarificación*) sugeridas a los evaluadores fueron realizadas de forma síncrona (en el laboratorio de TDi de la Universidad del Cauca). Con lo cual, para el evaluador supervisor resultó más sencillo coordinar la sesión con el objetivo de motivar la participación de todos los evaluadores y obtener resultados más rápidamente. Sin embargo, en las otras áreas de aplicación objeto de estudio las discusiones fueron coordinadas de forma asíncrona, lo cual resultó más complicado por la poca disponibilidad de los evaluadores. Para los objetivos de la presente investigación, esto no afecta los resultados obtenidos en las evaluaciones pues interesaba mayormente la cantidad de problemas de usabilidad detectados en los diferentes MEU. El análisis de la información, así como proponer soluciones a los problemas de usabilidad detectados, sobrepasan los límites de la investigación. Sin embargo, en un caso real dicha información es de vital importancia para la organización que realiza las evaluaciones de usabilidad.

Teniendo en cuenta lo anterior, el tiempo para completar las actividades colaborativas que conforman los MEU puede variar según el número de evaluadores que conformen el grupo. Lo recomendable es que para la realización de las actividades colaborativas, los integrantes del grupo trabajen de forma síncrona, lo cual se considera, facilitará la coordinación del proceso de comunicación entre los evaluadores y disminuirá en gran medida el tiempo de ejecución de las actividades. Los resultados serían más ricos en contenido si se tiene un grupo de evaluadores (contratado y/o remunerado) comprometido con las evaluaciones para lograr un trabajo colaborativo productivo.

Durante la ejecución de los MEU en las diferentes áreas de aplicación, el autor de la presente investigación actuó exclusivamente como *evaluador supervisor* (coordinador) de los equipos de trabajo, revisando las contribuciones de los mismos, sin influir en la evaluación de los sistemas interactivos estudiados.

Uso de tecnología

En la evaluación de usabilidad es de vital importancia contar con herramientas tecnológicas que permitan registrar la interacción del usuario con el sistema en evaluación, esto con el fin de no perder detalle alguno en las acciones que realiza el usuario. De esta manera, en las áreas de aplicación: web transaccional y redes sociales, fue utilizada la herramienta software MORAE diseñada para facilitar el proceso de evaluación de usabilidad. En el área de aplicaciones móviles el proceso de evaluación estuvo apoyado de la aplicación Phone

Screen Sharing (de Samsung) para observar la pantalla del dispositivo móvil en un computador. En el área de TDi no fue utilizada una herramienta software (instalada en el STB) para registrar las acciones de los usuarios durante la interacción con las aplicaciones, sin embargo, todas las sesiones de prueba fueron registradas con una cámara de video. Pero, en general, en las 4 áreas de aplicación que fueron la base de experimentación, la observación minuciosa del *evaluador supervisor* fue la principal herramienta para realizar las anotaciones correspondientes acerca de las acciones del usuario.

Teniendo en cuenta lo anterior, existen herramientas software y hardware que dan soporte al proceso de evaluación de usabilidad, sin embargo, la adquisición de algunas de ellas puede resultar costosa. Por lo cual, en la presente investigación fueron utilizadas herramientas software que están al alcance de cualquier practicante de la usabilidad para apoyar el proceso de evaluación y tener la información (grabaciones, anotaciones, etc.) de forma digital, para luego poder distribuirla entre los evaluadores participantes.

Mediante la presente investigación se ofrece a los practicantes de la usabilidad la primera versión de la herramienta software *ECUSI – Evaluación Colaborativa de la Usabilidad de Sistemas Interactivos* (disponible temporalmente en la URL: <http://ecussi.juancamiloceron.com/>), la cual fue concebida con el objetivo de guiar la ejecución de las actividades que conforman un conjunto de métodos de evaluación de usabilidad, esto haciendo uso de documentos compartidos en Google Docs que permiten la participación de diferentes personas involucradas en el proceso de evaluación. Para una segunda versión de la herramienta conviene incluir requerimientos relacionados a la automatización de actividades que conforman la evaluación heurística (tales como: la calificación de severidad y frecuencia, generación de rankings de criticidad y severidad, entre otras) e inclusión de mecanismos que promuevan la comunicación entre evaluadores.

Método de inspección: evaluación heurística

Considerando que existen algunos pocos MEU, que por naturaleza, presentan características que los distinguen como métodos colaborativos, ya que en su ejecución participan varios evaluadores, como por ejemplo: la evaluación heurística; el tiempo demorado en ejecutarlos de la forma tradicional en comparación a la forma “colaborativa” no presenta una diferencia significativa. Sin embargo, es destacable que en la forma “colaborativa” el proceso para llevarlos a cabo es detallado en mayor medida, de tal manera que pueden identificarse claramente los entregables correspondientes a cada actividad que lo conforma. Además, en las actividades que son colaborativas por la naturaleza de los métodos, se establece claramente (mediante la asociación de *thinklets*) el proceso de comunicación entre los evaluadores que participan en el desarrollo de dichas actividades.

Para realizar una evaluación heurística es altamente necesario contar con un conjunto de heurísticas específicas apropiadas para evaluar el sistema de interés, esto con el objetivo de minimizar la probabilidad de ignorar problemas específicos del dominio. Así, en la presente investigación han sido utilizadas una serie de heurísticas específicas para aplicaciones de TDi, web transaccional y aplicaciones móviles soportadas en dispositivos con tecnología táctil, sin embargo, en la literatura actual no fueron encontradas heurísticas específicas para redes sociales. Por tal razón, fue utilizada la herramienta software Open-HEREDEUX la cual sugiere una serie de heurísticas basadas en componentes, funcionalidades, características y facetas, según el sistema a evaluar.

Métodos de prueba

En todos los casos de las áreas de aplicación, en las evaluaciones participaron mismos evaluadores en diferentes métodos, lo cual pudo haber influenciado los resultados obtenidos puesto que conocían de antemano el sistema y problemas presentados. Esta situación se presentó por dificultades en conseguir diferentes evaluadores que tuvieran disponibilidad para participar en los métodos aplicados.

Dado que han sido identificadas una serie de actividades recurrentes destacadas en el proceso de evaluación de usabilidad, es posible utilizar el proceso colaborativo definido en la ejecución de MEU de prueba diferentes a los estudiados en la presente investigación, tal como sucedió con el método pensando en voz alta. Las actividades de la *etapa de planeación* de dicho método fueron definidas con base en el método del conductor, fueron definidas las actividades propias de la *etapa de ejecución*, y en la *etapa de análisis de resultados* fue utilizado el proceso colaborativo definido para actividades recurrentes destacadas.

En la realización de los métodos de prueba se encuentran usuarios con diferentes personalidades, por lo que es conveniente ejecutar un método de evaluación con un número considerable de usuarios representativos (tal como se sugiere en la Tabla 5 del Anexo B.3), esto con el fin de obtener información variada ya que se encuentran usuarios que proveen muy buena realimentación para mejorar el sistema que se esté evaluando. Sin embargo, son los propios evaluadores supervisores o responsables del proceso de evaluación quienes deciden el tamaño de la muestra de usuarios.

En la realización de los métodos de prueba deben contemplarse aspectos relacionados a las restricciones físicas de la interacción con el sistema a evaluar, por lo que la persona encargada de la evaluación debe cerciorarse de que el entorno en el cual es realizada la prueba, sea lo más parecido a la realidad. En el caso de los MEU realizados en la presente investigación, estos fueron llevados a cabo en instalaciones de la Universidad del Cauca, específicamente en el laboratorio experimental de TDi y la sala de doctorado del IPET – Instituto de Posgrados en Ingeniería Electrónica y Telecomunicaciones. El laboratorio tiene la infraestructura tecnológica básica para el despliegue de las aplicaciones de TDi, sin embargo, no ofrece un entorno completamente amigable a los usuarios como es el de una sala de una casa, que es el lugar donde normalmente se ve la televisión. La sala de doctorado ofreció condiciones apropiadas (de iluminación, muebles, conexión a internet) para la realización de los métodos de prueba correspondientes a las áreas: web transaccional, aplicaciones móviles y redes sociales.

Contexto de móvil

En trabajos relacionados como [95][96][97], se destaca que un factor influyente al momento de medir el grado de usabilidad de una aplicación móvil es el ambiente o contexto real donde el usuario utiliza la aplicación. Estas pruebas de campo dan lugar a diferentes interrupciones, distracciones u otras variables que afectan la usabilidad, los usuarios pueden estar de pie, caminando, o sentados en un ambiente oscuro o luminoso mientras usan la aplicación. Así, las pruebas de usabilidad en un contexto real resultan difíciles de realizar debido a las complicaciones para la recolección de datos, además que limitan el control sobre estas y sobre el usuario, ya que este se encuentra físicamente en movimiento con un número de variables desconocidas que afectan potencialmente la evaluación. Por tales razones, en la presente investigación los métodos de prueba en el área de aplicaciones móviles fueron

ejecutados en un ambiente controlado (en una sala de estudio), que no es un laboratorio de usabilidad, con el fin de obtener datos cualitativos como cuantitativos y tener mayor control sobre las pruebas y el usuario mismo. En consecuencia, no fue evaluada la usabilidad de la aplicación estudiada en un contexto real de uso. La anterior decisión obedece a que en un buen número de casos se obtienen datos más relevantes cuando las pruebas son realizadas en un ambiente controlado (en relación a las pruebas de campo), tal como se confirma en estudios relacionados como [98][99][100].

Teniendo en cuenta lo anterior, tal como se menciona en [99], lo recomendable es realizar pruebas de campo en las fases finales de desarrollo, con la condición de haber realizado previamente, como mínimo, una prueba en un ambiente controlado y/o laboratorio en las fases de desarrollo anteriores o antes de las pruebas de campo. Esto con el fin de descartar la mayor cantidad de problemas y haber podido obtener datos cuantitativos acerca del uso de la aplicación y/o dispositivo.

Métricas utilizadas

Para analizar los MEU objeto de estudio fue utilizada, en un momento dado, una calificación (único valor) que permitiera evaluar o estimar globalmente las características: *detección de problemas de usabilidad y tiempo*, de dichos métodos. Esto con el fin de obtener un indicador global que pudiera ser utilizado como base de comparación entre los MEU desarrollados en las áreas de aplicación objeto de estudio. Para el cálculo de la calificación global, la asignación de los pesos correspondientes a las características fue equivalente, es decir, que se asignó igual importancia a cada característica y, del mismo modo, a las métricas que conforman cada una de ellas. Este criterio fue empleado bajo la premisa de que cada una de las características influye de manera importante en la decisión para ejecutar un método de evaluación frente a otro. Sin embargo, este único valor fue descartado debido a que no está correctamente justificada la síntesis de las métricas en una calificación global por las siguientes razones: (i) las métricas son de diversas naturalezas y son difíciles de combinar, (ii) tienen más valor las métricas separadas que agregadas, (iii) no está justificado el peso de cada característica y cada métrica, (iv) la dificultad no está en elegir los pesos con los que ponderar, sino en poder asegurar que tiene sentido combinar dos métricas para producir otra que representa un atributo deseado, en este caso la calidad de un MEU.

Con base en lo anterior, fue planteada la siguiente recomendación: dependiendo de la necesidad de información y del contexto de evaluación, los evaluadores podrían acordar la asignación de diferentes pesos para las características (y de las métricas que incluyen), como una forma de reflejar la mayor importancia de unas respecto de otras en relación a los objetivos del caso de estudio. Además, se podrían realizar encuestas a los implicados (representante de la organización, expertos/evaluadores, usuarios) para determinar y ponderar los pesos de las características, tal como propone [101].

Capítulo 8

El conocimiento no es una vasija que se llena, sino un fuego que se enciende. Plutarco

Conclusiones y trabajo futuro

8.1 Conclusiones

La metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos propuesta en la presente investigación ha sido planteada mediante la realización de 3 casos de aplicación en diferentes áreas (televisión digital interactiva, web transaccional y aplicaciones móviles). La metodología está conformada por 3 combinaciones de MEU que pueden ser utilizadas en diferentes escenarios, esas son: (a) *evaluación global* – alta detección de problemas, (b) *evaluación específica* – reducción de tiempo y (c) *evaluación enfocada a tareas específicas* – sin restricciones de tiempo. Esto considerando las condiciones bajo las cuales fueron ejecutados los métodos de evaluación objeto de estudio. Luego, una de las combinaciones de MEU fue evaluada exitosamente mediante un caso de aplicación final en un área (redes sociales) diferente a las antes contempladas. Así, dependiendo de los objetivos de evaluación de cierto sistema, se estima que cada combinación de MEU que conforma la metodología propuesta permitiría obtener resultados de la medida de la usabilidad de manera confiable.

En un buen número de trabajos han sido ejecutados MEU y se recomienda el uso de ellos, sin embargo, escasos trabajos sugieren combinaciones de métodos de evaluación que puedan ser potencialmente aplicables para evaluar la usabilidad de sistemas de diferentes áreas de aplicación. Ahora bien, las combinaciones de MEU que conforman la metodología propuesta no presentan un alto grado de complejidad, con lo cual es posible potenciar su comprensibilidad y aplicabilidad en diferentes sistemas interactivos. Lo ideal sería ejecutar los MEU en un laboratorio de usabilidad, sin embargo, se considera que los métodos de evaluación que hacen parte de las combinaciones propuestas pueden lograrse con una evaluación semi-formal (en un contexto controlado). Esto es, con una inversión relativamente pequeña podría obtenerse una mejora significativa en la usabilidad de un sistema software interactivo. Lo anterior no limita a los evaluadores a utilizar otras tecnologías durante el proceso de evaluación (como por ejemplo, el eye-tracking) con el objetivo de complementar la evaluación y resultados obtenidos.

Teniendo en cuenta que en los MEU tradicionales no se han definido de manera clara y específica los roles y las funciones para los diferentes actores involucrados en el proceso de evaluación, en esta investigación la especificación colaborativa de los MEU objeto de estudio provee una secuencia de actividades bien definidas, especificación de entregables, descripción de los diferentes participantes del proceso de evaluación y especificación del proceso de comunicación entre los participantes de la evaluación. Por otro lado, se tiene escasa documentación (guías o lineamientos) acerca de cómo ejecutar evaluaciones colaborativas de usabilidad de sistemas interactivos. Con base en lo anterior, surgió la idea de desarrollar esta investigación, cuyos resultados se pretende sean utilizados por practicantes de la usabilidad o personas responsables de estructurar el proceso de

evaluación de la usabilidad de diferentes sistemas interactivos. La finalidad perseguida en la elaboración de una guía de evaluación colaborativa es brindar herramientas a la comunidad para facilitar el desarrollo de una cultura de evaluación de la usabilidad de software.

En esta investigación fue ejecutado un conjunto de MEU, donde la televisión digital interactiva, web transaccional, aplicaciones móviles y redes sociales, fueron las áreas de aplicación consideradas como base experimental. Luego de la ejecución de los métodos fue realizado el análisis de los resultados obtenidos mediante una serie de métricas asociadas a las características: *detección de problemas de usabilidad y tiempo*. Es importante recalcar que no existe un “mejor método”, todos tienen fortalezas y debilidades y están enfocados a evaluar aspectos específicos de usabilidad, por tal razón, según [10], combinarlos es el procedimiento más apropiado. De esta manera, la presente investigación propone 3 combinaciones de MEU que pueden servir como base para evaluar un sistema interactivo según un escenario específico. Este es un punto de partida considerado relevante, ya que como se afirma en [102], existe una variedad de MEU y son necesarios trabajos que los comparen y analicen su eficiencia en diferentes áreas y contextos de aplicación, desde un enfoque de *usabilidad formativa*.

Dado que desde la perspectiva de los usuarios la usabilidad puede hacer la diferencia entre desarrollar una tarea completamente o no durante la interacción con un sistema, y disfrutar el proceso o sentirse frustrado, mediante esta propuesta de investigación se intenta aumentar en gran medida la posibilidad de éxito de los sistemas software. Ahora bien, se tiene que uno de los principales atributos de calidad de los sistemas interactivos es la usabilidad. Esta resulta ser un aspecto crítico en el desarrollo de sistemas que puedan declararse exitosos, ya que el hecho de que un sistema sea usable implica que permite a sus usuarios trabajar de manera eficaz, eficiente y satisfactoria. Por lo cual, la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos* contribuiría en ese aspecto.

La Ingeniería de Colaboración está enfocada en el diseño de prácticas de trabajo colaborativo para tareas recurrentes destacadas. Ahora bien, los métodos de prueba siguen un mismo proceso, en donde es necesaria la preparación de la prueba a realizar, definición de los perfiles de usuarios, realización de una prueba piloto, detección de problemas de usabilidad, análisis e interpretación de la información recolectada, entre otros aspectos. Por tal razón, la especificación colaborativa de una serie de actividades recurrentes (como las descritas en el Anexo D) puede ser reutilizada para definir de forma colaborativa otros MEU. De esta manera, si posteriormente se desea experimentar con otro método de evaluación, este puede especificarse de forma colaborativa (con base en los diseños existentes) y combinarse con otros métodos que conforman la metodología de evaluación. En ese sentido, la metodología propuesta está abierta a incluir otros MEU, además que el diseño colaborativo de estos no sería una tarea complicada al adoptar los procesos predefinidos.

La realización de un proceso diseñado de forma colaborativa incrementa en gran medida la posibilidad de obtener resultados más completos y ricos en contenido, respecto a un proceso que no integra aspectos de trabajo colaborativo. Mediante la presente investigación ha sido identificado un conjunto de MEU adecuados para ejecutar en diferentes entornos (televisión digital interactiva, web), los cuales pueden ser realizados de la forma tradicional (como se han definido), obteniendo resultados significativos. Sin embargo, la realización de dichos métodos diseñados de forma colaborativa, permitiría obtener resultados más apropiados respecto al número de problemas identificados y tiempo empleado en el análisis de la información, esto considerando las ventajas del trabajo colaborativo. Cabe resaltar que la

colaboración se hace apropiada principalmente en la *etapa de análisis de resultados*. En la *etapa de planeación* es opcional trabajar de forma colaborativa.

El proceso para definir las combinaciones de *thinklets* (asociados a los *patrones de colaboración*) no fue sencillo. En la medida en que fueron ejecutados los MEU fueron realizados ajustes a las combinaciones considerando factores como el tiempo, el cual es determinante puesto que los evaluadores disponen de tiempo limitado para hacer sus contribuciones. Combinaciones iniciales de *thinklets*, como por ejemplo: *freebrainstorm* (*thinklet* asociado al patrón *generación*) y *fastfocus* (*thinklet* asociado al patrón *reducción*) fueron descartadas por razones de tiempo y cantidad de información obtenida. De esta manera, los casos de aplicación (realización de la forma de validación *prueba piloto* [8]) permitieron ajustar poco a poco la especificación colaborativa de los MEU. Cabe resaltar que el trabajo realizado (evaluaciones de usabilidad) va más allá de identificar una serie de problemas de usabilidad, lo interesante fue observar las dinámicas de grupo, y cómo los *thinklets* contribuyen a un proceso de comunicación “más” ordenado.

Las dinámicas del trabajo colaborativo contribuyen a obtener resultados rápidamente cuando las personas trabajan de forma síncrona (como sucedió en el caso del área de TDi). En las áreas web transaccional y aplicaciones móviles los evaluadores estaban distribuidos geográficamente, por lo cual las contribuciones se obtuvieron en un periodo de tiempo considerable (3-4 días en promedio). Cabe destacar que en todos los casos la información fue distribuida, por lo que el análisis no queda sujeto a la percepción de una sola persona, y el esfuerzo (carga de trabajo) es distribuido entre varias personas. Trabajar de forma colaborativa asíncrona es un opción que toma el *evaluador supervisor* siendo consciente que no todas las personas tienen igual disponibilidad, por lo cual la obtención de resultados se va a ver retrasada. Esto implica hacer una planeación ajustada de tiempo para dar un espacio prudente y puedan contribuir todos los evaluadores.

Dado que en el Modelo de Facilitación del Proceso (MFP) han sido identificadas una serie de carencias relacionadas a la representación de información, la propuesta de notación para el modelado de procesos colaborativos extendiendo la notación HAMSTERS (ver sección 3.3.3), resulta apropiada para enriquecer la representación gráfica del MFP. Considerando los resultados obtenidos en el caso de estudio (ver sección 3.3.4), mediante las extensiones realizadas a la notación se contribuye a la comprensión, entendimiento y uso del modelo.

La notación extendida (MFP+HAMSTERS) es aplicable para representar actividades de diversos procesos relacionados a diferentes sistemas software interactivos y facilita el trabajo de aquellos practicantes de la Ingeniería de Colaboración que requieran una herramienta para modelar procesos colaborativos. Sin embargo, conviene realizar otros casos de estudio y refinado de la notación extendida con el fin de ofrecer una herramienta confiable a los diseñadores de procesos colaborativos.

Respecto a la tecnología utilizada, para la realización de las actividades colaborativas fueron consideradas herramientas software existentes, ya que es una tarea complicada “obligar” a los evaluadores a utilizar un sistema o plataforma desconocida para ellos y de la cual tienen altas expectativas, teniendo en cuenta tecnologías existentes en la actualidad (como las de Google por ejemplo). Por tal razón, fueron seleccionadas herramientas con las cuales los evaluadores están familiarizados y utilizan constantemente durante sus actividades diarias, además que no requieren de configuraciones complejas para la realización de las actividades. Así, fueron utilizados los documentos compartidos de GoogleDocs como

herramienta ofimática colaborativa y herramientas de comunicación como Hangout y Skype. Teniendo en cuenta lo anterior, como producto externo a esta investigación, se obtuvo la primera versión de una herramienta software que guía la ejecución de los MEU, la cual trata de integrar tecnologías conocidas para los evaluadores, como son las de Google.

La usabilidad es un factor diferencial, y Nielsen [44] analiza la relación costo-beneficio de evaluar la usabilidad y lo plantea como una gran inversión de acuerdo a las altas tasas de retorno de inversión. Teniendo en cuenta lo anterior, las organizaciones desarrolladoras de sistemas software interactivos podrían beneficiarse con la metodología de evaluación propuesta en la presente investigación, para evaluar de forma colaborativa la usabilidad de sus sistemas, ya que al integrar varias personas de diferentes áreas de conocimiento y con diferentes niveles de experticia en el proceso de evaluación, estas comparten conocimiento y recursos, garantizando así la obtención de sistemas más usables para los usuarios finales. Esto se vería reflejado en un aumento de la cuota de mercado y, en consecuencia, un aumento de los beneficios generales de la organización, puesto que conseguiría mayor satisfacción de los usuarios/clientes.

Ejecutar los métodos colaborativos que conforman la metodología de evaluación propuesta, en comparación a los métodos tradicionales, podría traer mayores costos ya que involucran a más personas en la realización de actividades. Sin embargo, la literatura destaca que el trabajo colaborativo en la evaluación de usabilidad presenta diversos beneficios (en cuanto a cantidad de problemas de usabilidad identificados por ejemplo). Ahora, está ampliamente reconocido que el ROI (Retorno de la Inversión) de la usabilidad es mayor que el producido por cualquier otra inversión que pueda realizarse en un proyecto. Esto es, los sistemas que más invertirán en usabilidad serán más fáciles de usar e incrementarán las ventas [7][44]. Teniendo en cuenta lo anterior, los autores de la presente investigación consideran que los sistemas software interactivos desarrollados utilizando la metodología de evaluación propuesta, tendrían una calidad adecuada y el ROI podría ser positivo, a pesar de requerir mayor inversión por la participación de más personas.

Teniendo en cuenta todas las conclusiones anteriores, se logra confirmar la hipótesis definida al inicio de la investigación. Sin lugar a dudas, la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos*, conformada por 3 combinaciones de MEU (que consideran diferentes escenarios de evaluación), brinda la información de la usabilidad de forma más completa que aquellos métodos de evaluación ejecutados de forma independiente.

Tras la finalización de todas las etapas y actividades previstas para establecer la *metodología para la evaluación colaborativa de la usabilidad de sistemas software interactivos*, cabe mencionar que para los integrantes de esta investigación la realización de la misma ha sido una labor gratificante, tanto por los resultados obtenidos como las experiencias vividas a lo largo del proceso.

8.2 Trabajo futuro

Esta tesis de doctorado es un trabajo amplio que deja muchas puertas abiertas para continuar con la investigación realizada. Las principales actividades futuras están relacionadas al refinamiento de la metodología de evaluación propuesta. La utilización de las combinaciones de MEU propuestas (global, específica y enfocada a tareas) en distintos sistemas software interactivos servirá como realimentación para conocer el comportamiento

de las mismas. Es probable que, con el paso del tiempo y la gran velocidad de cambio de las tecnologías de información y comunicación, alguna de las combinaciones propuestas deba ser actualizada o surjan nuevas combinaciones que den soporte a escenarios de evaluación distintos de los contemplados en la presente investigación.

La metodología de evaluación fue propuesta considerando las características: *detección de problemas de usabilidad y tiempo*, en los métodos de evaluación ejecutados. Un conjunto de métricas, correspondientes a cada una de dichas características, fue estudiado para observar el comportamiento de los MEU objeto de estudio en diferentes áreas de aplicación. Como trabajo futuro, si se desea plantear combinaciones adicionales de MEU considerando otros factores/características, conviene realizar un proceso como el llevado a cabo en la presente investigación, que consiste básicamente en estudiar el comportamiento de los MEU con base en el nuevo factor definido en diferentes áreas de aplicación. Por otro lado, sería interesante tener una o más métricas asociadas al impacto de la colaboración en el proceso de uso de los métodos de evaluación analizados.

Dado que en la presente investigación únicamente fue evaluada la combinación de MEU: *evaluación específica*, es conveniente que las combinaciones: *evaluación global* y *evaluación enfocada a tareas concretas*, también sean evaluadas mediante casos de estudio específicos. De igual forma, conviene difundir la metodología de evaluación propuesta (y las combinaciones de métodos que la conforman) a la comunidad académica, con el fin de que sea considerada en investigaciones que tratan temas de trabajo colaborativo y evaluación de usabilidad de software.

Conviene continuar realizando ajustes a la especificación colaborativa de los MEU objeto de estudio. Es posible actualizar las especificaciones colaborativas propuestas teniendo en cuenta el elemento propuesto [103]: *ThinXel*, para representar el flujo de proceso de los *thinklets*. Adicionalmente, es posible realizar una validación adicional al diseño colaborativo mediante las técnicas *Revisión* y *Recorrido* (proporcionadas por la *Metodología para el Desarrollo de Procesos Colaborativos*).

Los equipos de trabajo, responsables del proceso de evaluación de usabilidad de sistemas interactivos, presentan los mismos desafíos que tiene cualquier otro equipo, tales como participantes dominantes o tímidos, falta de consenso, entre otros. No necesariamente, los expertos en usabilidad tienen las habilidades requeridas para coordinar esos desafíos y guiar las pruebas de usabilidad satisfactoriamente [104]. La evaluación de usabilidad de sistemas interactivos es un proceso que requiere de tiempo y experiencia en el área. Es destacable que durante las evaluaciones de usabilidad de los sistemas fueron involucradas varias personas (profesionales directamente implicados en el desarrollo de los sistemas, ingenieros de sistemas, ingenieros electrónicos, usuarios finales y expertos en usabilidad), con el fin de que estas fueran llevadas a cabo de manera más apropiada. Sin embargo, como trabajo futuro conviene realizar evaluaciones de usabilidad con equipos multidisciplinarios buscando hacer frente a desafíos como los antes mencionados.

Considerando que la ejecución de algunas fases que conforman el *Modelo de proceso para el diseño de técnicas colaborativas de evaluación de usabilidad de software* (obtenido a partir de la *Metodología para el desarrollo de procesos colaborativos*) puede resultar una tarea tediosa para los diseñadores de procesos colaborativos por su alto componente teórico, es posible proponer una metodología ágil para el diseño de procesos colaborativos.

En la presente investigación, en las áreas de aplicación estudiadas, excepto en TDi, los

métodos de inspección no fueron ejecutados en las mismas condiciones que los métodos de prueba (en el mismo lugar, usando el mismo dispositivo, entre otras). Por lo cual, resulta conveniente realizar nuevas evaluaciones donde sean aislados aquellos factores que pueden tener influencia en los resultados obtenidos. Esto es considerado como trabajo futuro pues realizar las evaluaciones en ambientes controlados (con iguales condiciones) resulta complicado cuando los evaluadores se encuentran distribuidos geográficamente.

En esta investigación la faceta *usabilidad* ha sido considerada como eje central, por lo que se propone ampliar el alcance del trabajo pasando de “evaluación colaborativa de la usabilidad” a “evaluación colaborativa de la experiencia de usuario”. Así, para un estudio posterior, convendría incluir (o combinar) elementos de otras facetas, como por ejemplo: emotividad, multiculturalidad, entre otras, a la metodología de evaluación propuesta.

Aunque es una tarea complicada, se considera apropiado realizar un estudio detallado de la relación coste/beneficio o del ROI para las distintas combinaciones de MEU propuestas.

Finalmente, en el marco de la evaluación de usabilidad, el enfoque futuro tiende a proponer componentes de evaluación. De esta manera, podría pensarse en hacer una posible integración de los MEU que conforman las diferentes combinaciones propuestas. También pensar en una integración de actividades para que una combinación de MEU sea vista como un único componente de evaluación y no como los métodos que la conforman.

8.3 Publicaciones

Como resultado del trabajo de investigación fueron realizadas algunas publicaciones, no obstante queda pendiente generar más publicaciones donde sean presentados los resultados finales logrados. Las publicaciones realizadas hasta el momento son las siguientes:

Capítulos de libro:

- Collazos, C., Solano, A. Creación de objetos de aprendizaje de contenidos abiertos accesibles desde un enfoque colaborativo. *Objetos de Aprendizaje de Contenidos Abiertos Accesibles: Del Diseño a la Reutilización*, 2014, pp 60-81.
- Collazos, C., Solano, A. Diseño de procesos colaborativos basado en patrones de colaboración. *Ambientes colaborativos para la enseñanza de las ciencias y las tecnologías*, 2013, pp. 83-102.

Revistas:

- Solano, A., Granollers, T., Collazos, C., Arciniegas, J. Experiencias en la especificación colaborativa de métodos de evaluación de usabilidad. *Faz, revista de diseño de interacción*. Número 7/Julio 2014, pp. 175-186. ISSN: 0718-526X.
- Solano, A., Chanchí, G., Collazos, C., Arciniegas, J., Rusu, C. Directrices para el diseño de aplicaciones usables en entornos de Televisión Digital Interactiva. *Revista Ingeniería y Universidad*. Vol. 18, 2014, pp. 103-120. ISSN: 2011-2769. Indexada en la Categoría B.
- Solano, A., Rusu, C., Collazos, C., Arciniegas, J. Evaluating Interactive Digital Television Applications through Usability Heuristics. *INGENIARE, Revista Chilena de*

Ingeniería, Universidad de Tarapacá, Chile. Vol. 21, 2013, pp. 16-29. ISSN: 0718-3305. Homologada en Colombia en la Categoría A1.

- Solano, A., Collazos, C., Arciniegas, J. Evaluando la Usabilidad de Aplicaciones de Televisión Digital Interactiva desde una Perspectiva Colaborativa. *Entre Ciencia e Ingeniería*. Vol. 12, 2012, pp. 43-49. ISSN: 1909-8367. Indexada en la Categoría B.

Conferencias internacionales:

- Solano, A., Granollers, T., Collazos, C., Rusu, C. Modelo de Facilitación del Proceso incluyendo elementos de la notación HAMSTERS. Aceptado en *XV Congreso Internacional de Interacción Persona-Ordenador (Interacción 2014)*. Puerto de la Cruz, Tenerife, España, Septiembre 2014. ISSN 10: 84-697-1072-9.
- Solano, A., Granollers, T., Collazos, C., Rusu, C. Proposing formal notation for modeling collaborative processes extending HAMSTERS notation. Aceptado en *World Conference on Information Systems and Technologies (WorldCIST'14)*. Madeira, Portugal, Abril 2014. ISSN: 2194-5357.
- Solano, A., Masip, L., Granollers, T., Collazos, C., Rusu, C., Arciniegas, J. Setting usability iTV heuristics in Open-HEREDEUX. Aceptado en *6th Latin American Conference on Human Computer Interaction (CLIHIC 2013)*. Guanacaste, Costa Rica, Diciembre 2013. ISBN: 978-3-319-03067-8.
- Solano, A., Granollers, T., Collazos, C., Arciniegas, J. Experiencias en la especificación colaborativa de métodos de evaluación de usabilidad. Aceptado en *XIV Congreso Internacional de Interacción Persona-Ordenador (Interacción 2013)*. Madrid, España, Septiembre 2013.
- Solano, A., Collazos, C., Muñoz, J., Pantoja, W. Propuesta de mecanismo para la elaboración de Objetos de Aprendizaje desde un enfoque colaborativo. Aceptado en la *V Conferencia Conjunta Iberoamericana sobre Tecnologías y Aprendizaje (CCITA 2013)*, Cancún, México, Septiembre 2013.
- Solano, A., Collazos, C., Rusu, C., Merchán, L. Evaluating the Usability of Interactive Digital Television Applications. Aceptado en *10th International Conference on Information Technology: New Generations (ITNG 2013)*, Las Vegas, Nevada, Abril 2013.

Conferencias nacionales:

- Solano, A., Collazos, C., Rusu, C. Estudio de métodos de evaluación de usabilidad colaborativos en el área web transaccional. Aceptado en el *Noveno Congreso Colombiano de Computación (8CCC)*, Pereira, Colombia, Septiembre 2014. ISBN: 978-1-4799-1056-4.
- Solano, A., Collazos, C., Rusu, C. Evaluación colaborativa de la usabilidad de sistemas software interactivos. Aceptado en el *Octavo Congreso Colombiano de Computación (8CCC)*, Armenia, Colombia, Agosto 2013.
- Solano, A., Rusu, C., Collazos, C., Arciniegas, J. Methodological Proposal to Evaluate the Usability of Interactive Digital Television Applications. Aceptado en el *Séptimo Congreso Colombiano de Computación (7CCC)*, Medellín, Colombia, Octubre 2012. ISBN: 978-1-4673-1474-9.

Referencias bibliográficas

- [1] R. Otaiza, "Metodología de evaluación de usabilidad para aplicaciones web transaccionales," Magíster en Ingeniería Informática Tesis de Maestría, Escuela de Ingeniería Informática, Pontificia Universidad Católica de Valparaíso, Valparaíso, Chile, 2008.
- [2] H. Sharp, Y. Rogers, J. Preece, *Interaction Design Beyond Human - Computer Interaction*, 2 ed.: Wiley, John & Sons, Incorporated, 2007.
- [3] L. Masip, M. Oliva, T. Granollers, "User experience specification through quality attributes," in *Human-Computer Interaction-INTERACT 2011*, ed: Springer, 2011, pp. 656-660.
- [4] ISO, "International Software Quality Standard, ISO/IEC 25010," in *Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Systems and software quality models*, ed, 2011.
- [5] J. Nielsen, *Usability engineering*: Morgan Kaufmann Publishers, 1993.
- [6] ISO, "International Standard ISO/IEC 9241," in *Ergonomic requirements for office work with visual display terminals*, ed, 1998.
- [7] T. Granollers, "MPLu+a una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinarios," Tesis Doctoral, Departamento de Sistemas Informáticos, Universidad de Lleida, Lleida, 2007.
- [8] G. Kolfshoten, G.-J. D. Vreede, "The Collaboration Engineering Approach for Designing Collaboration Processes," in *International Conference on Groupware: Design, Implementation and Use*, 2007, pp. 38-54.
- [9] R. Otaiza, C. Rusu, S. Roncagliolo, "Evaluating the usability of transactional Web Sites," in *Third International Conference on Advances in Computer-Human Interactions (ACHI'10)*, Saint Maarten, 2010, pp. 32-37.
- [10] X. Ferré, "Marco de integración de la usabilidad en el proceso de desarrollo software," Tesis Doctoral, Lenguajes y Sistemas Informáticos e Ingeniería del Software, Universidad Politécnica de Madrid, Madrid, 2005.
- [11] X. Ferré, N. Bevan, "Usability planner: a tool to support the process of selecting usability methods," in *Human-Computer Interaction-INTERACT 2011*, ed: Springer, 2011, pp. 652-655.
- [12] T. Granollers, "MPLu+a una metodología que integra la ingeniería del software, la interacción persona-ordenador y la accesibilidad en el contexto de equipos de desarrollo multidisciplinarios," Tesis Doctoral, Departamento de Sistemas Informáticos, Universidad de Lleida, Lleida, 2004.
- [13] M. Obeso, "Metodología de Medición y Evaluación de la Usabilidad en Sitios Web Educativos," Tesis Doctoral, Universidad de Oviedo, Oviedo, 2005.
- [14] C. A. Ellis, S. J. Gibbs, G. Rein, "Groupware: some issues and experiences," *Communications of the ACM*, vol. 34, pp. 39-58, 1991.
- [15] Y. Méndez, "Modelo de Proceso para el Diseño de Técnicas Colaborativas de Evaluación de Usabilidad de Software," Tesis de Maestría, Facultad de Ingeniería Electrónica y Telecomunicaciones, Departamento de Sistemas, Universidad del Cauca, Popayán, 2011.
- [16] L. Masip, M. Oliva, T. Granollers, "Repositorio Abierto de Heurísticas," in *XIII Congreso Internacional de Interacción Persona-Ordenador (Interacción 2012)*, Elche, Alicante, España, 2012.
- [17] L. Masip, M. Oliva, T. Granollers, "Classification of Interactive System Components Enables Planning Heuristic Evaluation Easier," in *Design, User Experience, and Usability. Theory, Methods, Tools and Practice*, ed: Springer, 2011, pp. 478-486.
- [18] ISO, "9241-210:2008," in *Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems*, ed, 2008.
- [19] E. L.-C. Law, V. Roto, M. Hassenzahl, A. P. Vermeeren, J. Kort, "Understanding, scoping and defining user experience: a survey approach," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2009, pp. 719-728.

- [20] L. Masip, M. Oliva, T. Granollers, "Concreción de la Experiencia de Usuario mediante Atributos de Calidad," in *XII Congreso Internacional de Interacción Persona-Ordenador (Interacción 2011)*, Lisboa, 2011.
- [21] J. Nielsen, *Designing web usability*. Pearson Deutschland GmbH, 2004.
- [22] N. Bevan, J. Kirakowski, J. Maissel, "What is usability?," in *Human Aspects in Computing: Design and Use of Interactive Systems with Terminals*, 1991.
- [23] J. Nielsen, *The usability engineering life cycle*: IEEE, 1992.
- [24] A. Dix, J. Finlay, G. D. Abowd, *Human-computer interaction*. Prentice hall, 2004.
- [25] X. Ferré, "Integration of usability techniques into the software development process," in *International Conference on Software Engineering (Bridging the gaps between software engineering and human-computer interaction)*, 2003, pp. 28-35.
- [26] R. Fitzpatrick, "Strategies for evaluating software usability," *Articles*, p. 1, 1999.
- [27] S. Riihiaho, "Experiences with usability evaluation methods," Licentiate thesis, Laboratory of Information Processing Science, Helsinki University of Technology, 2000.
- [28] A. Holzinger, "Usability engineering methods for software developers," *Communications of the ACM*, vol. 48, pp. 71-74, 2005.
- [29] W. O. Galitz, *The essential guide to user interface design: an introduction to GUI design principles and techniques*: Wiley Computer Pub., 2002.
- [30] K. Schmidt, L. Bannon, "Taking CSCW seriously," *Computer Supported Cooperative Work (CSCW)*, vol. 1, pp. 7-40, 1992.
- [31] S. Greenberg, *Computer-supported cooperative work and groupware*: Academic Press Ltd., 1991.
- [32] E. Turban, *Decision support and expert systems: management support systems*: Prentice Hall PTR Upper Saddle River, NJ, USA, 1995.
- [33] M. Sapon-Shevin, B. J. Ayres, J. Duncan, "Cooperative learning and inclusion," *Creativity and collaborative learning: A practical guide to empowering students and teachers*, pp. 45-58, 1994.
- [34] G. L. Kolfshoten, R. O. Briggs, G. Vreede, "Definitions in Collaboration Engineering," in *International Conference on System Sciences*, 2006, pp. 58-74.
- [35] G. J. De Vreede, R. O. Briggs, "Collaboration engineering: designing repeatable processes for high-value collaborative tasks," in *38th Annual Hawaii International Conference on System Sciences 2005*, pp. 21-32.
- [36] R. Briggs, G. Kolfshoten, V. Gert-Jan, D. Douglas, "Defining key concepts for collaboration engineering," in *Americas Conference on Information Systems*, 2006, pp. 121-128.
- [37] R. O. Briggs, G. J. De Vreede, J. F. Nunamaker Jr, "Collaboration engineering with ThinkLets to pursue sustained success with group support systems," *Journal of Management Information Systems*, vol. 19, pp. 31-64, 2003.
- [38] G. L. Kolfshoten, R. O. Briggs, J. H. Appelman, G. J. de Vreede, "ThinkLets as building blocks for collaboration processes: a further conceptualization," *Groupware: Design, Implementation and Use*, pp. 137-152, 2004.
- [39] R. O. Briggs, G. J. De Vreede, J. F. Nunamaker Jr, D. Tobey, "ThinkLets: achieving predictable, repeatable patterns of group interaction with group support systems (GSS)," in *34th Hawaii International Conference on Systems Sciences*, 2001, p. 9.
- [40] G. Kolfshoten, G.-J. D. Vreede, "Thinklet Design Support Booklet," 2006, pp. 4-48.
- [41] A. Følstad, E. Law, K. Hornbæk, "Analysis in practical usability evaluation: a survey study," in *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems*, 2012, pp. 2127-2136.
- [42] A. Solano, C. Collazos, J. Arciniegas, "Evaluando la Usabilidad de Aplicaciones de Televisión Digital Interactiva desde una Perspectiva Colaborativa.," *Entre Ciencia e Ingeniería*, vol. 12, pp. 43-49, 2012.
- [43] A. P. Vermeeren, E. L.-C. Law, V. Roto, M. Obrist, J. Hoonhout, K. Väänänen-Vainio-Mattila, "User experience evaluation methods: current state and development needs," in *Proceedings of the 6th Nordic Conference on Human-Computer Interaction: Extending Boundaries*, 2010, pp. 521-530.
- [44] J. Nielsen, "Usability inspection methods," in *Conference companion on Human factors in*

- computing systems*, 1994, pp. 413-414.
- [45] J. S. Dumas, J. C. Redish, *A practical guide to usability testing*: Ablex Publishing Corporation, 1999.
- [46] J. Rubin, D. Chisnell, *Handbook of Usability Testing: Howto Plan, Design, and Conduct Effective Tests*: Wiley Publishing, 2008.
- [47] A. Solano, "Propuesta metodológica para la evaluación colaborativa de la usabilidad de aplicaciones de Televisión Digital Interactiva," Tesis de Maestría, Instituto de Posgrados en Ingeniería Electrónica, Universidad del Cauca, Popayán, Cauca, 2012.
- [48] L. Masip, "Análisis de viabilidad de soluciones para la automatización de la evaluación heurística," Máster Universitario en Interacción Persona-Ordenador Tesis de Maestría, Escola Politècnica Superior, Universitat de Lleida, Lleida, España, 2010.
- [49] T. Granollers, J. Lorés, "Incorporation of users in the Evaluation of Usability by Cognitive Walkthrough," *HCI related papers of Interacción*, pp. 243-255, 2006.
- [50] C. Martinie, P. Palanque, M. Winckler, "Structuring and composition mechanisms to address scalability issues in task models," in *Human-Computer Interaction-INTERACT 2011*, ed: Springer, 2011, pp. 589-609.
- [51] A. Solano, C. Parra, "Métodos de Indagación Colaborativos para la Evaluación de Usabilidad de Software," Tesis de pregrado, Facultad de Ingeniería Electrónica y Telecomunicaciones, Departamento de Sistemas, Universidad del Cauca, Popayán, 2009.
- [52] D. Diaper, *Task analysis for human-computer interaction*: Prentice Hall PTR, 1990.
- [53] F. Paternò, *Model-based Design and Evaluation of Interactive Applications*: Springer, 2000.
- [54] J. Annett, "Hierarchical task analysis," *Handbook of cognitive task design*, pp. 17-35, 2003.
- [55] N. Stanton, *The handbook of task analysis for human-computer interaction*: Routledge, 2004.
- [56] F. Paternò, C. Mancini, S. Meniconi, "ConcurTaskTrees: A diagrammatic notation for specifying task models," in *Human-Computer Interaction INTERACT'97*, 1997, pp. 362-369.
- [57] F. Paternò, "ConcurTaskTrees: an engineered notation for task models," *The handbook of task analysis for human-computer interaction*, pp. 483-503, 2004.
- [58] T. Tullis, B. Albert, *Measuring the user experience: collecting, analyzing, and presenting usability metrics*, Second ed.: Morgan Kaufmann, 2013.
- [59] A. Solano, T. Granollers, C. Collazos, C. Rusu, "Proposing formal notation for modeling collaborative processes extending HAMSTERS notation," in *World Conference on Information Systems and Technologies (WorldCIST'14)*, Madeira, Portugal, 2014.
- [60] A. Solano, C. Parra, C. Collazos, Y. Méndez, "Evaluación de Usabilidad de Software desde una Perspectiva Colaborativa," in *Conferencia Latinoamericana de Medios Audiovisuales en Red - LACNEM 2010*, Cali, Colombia, 2010, pp. 42-47.
- [61] J. Nielsen, R. Molich, "Heuristic evaluation of user interfaces," in *Proceedings of the SIGCHI conference on Human factors in computing systems: Empowering people*, 1990, pp. 249-256.
- [62] L. Masip, M. Oliva, T. Granollers, "OPENHEREDEUX: open heuristic resource for designing and evaluating user experience," in *13th IFIP TC 13 International Conference on Human-Computer Interaction*, Berlin, 2011, pp. 418-421.
- [63] M. Piattini, F. García, J. Garzás, M. Genero, "Medición y estimación del software: técnicas y métodos para mejorar la calidad y productividad del software," *Ra-Ma (ed.)*, pp. 121-127, 2008.
- [64] D. C. Baird, J. C. Peña, *Experimentación: Una Introducción a la teoría de mediciones y al diseño de experimentos*: Prentice-Hall Hispanoamericana, 1991.
- [65] H. R. Hartson, T. S. Andre, R. C. Williges, "Criteria for evaluating usability evaluation methods," *International Journal of Human-Computer Interaction*, vol. 15, pp. 145-181, 2003.
- [66] F. Bellotti, S. Vrochidis, E. Parissi, P. Lhoas, D. Mathevon, M. Pellegrino, G. Bo, I. Kompatsiaris, "A T-learning Courses Development and Presentation Framework," *IEEE Multidisciplinary Engineering Education Magazine*, vol. 3, pp. 69-76, 2008.
- [67] D. Geerts, D. De Grooff, "Supporting the social uses of television: sociability heuristics for social TV," presented at the Computer Human Interaction (CHI '09), 2009.
- [68] G. Chanchí, E. Gabriel, W. Y. Campo, J. P. Amaya, J. L. Arciniegas, "Esquema de servicios para Televisión Digital Interactiva, basados en el protocolo REST-JSON," *Cadernos de Informática*, vol. 6, pp. 233-240, 2011.

- [69] A. Solano, C. Rusu, C. A. Collazos, J. Arciniegas, "Evaluando aplicaciones de televisión digital interactiva a través de heurísticas de usabilidad," *Ingeniare. Revista chilena de ingeniería*, vol. 21, pp. 16-29, 2013.
- [70] P. Aarreniemi, "T-Learning Model for Learning via Digital TV," presented at the Annual Conference on Innovation in Education for Electrical and Information Engineering EAEEIE, 2005.
- [71] M. Rinnetmäki, M. Heikkinen, I. Kosonen, M. Saarijärvi, P. Nykänen, A. Saikanmäki, *A guide for Digital TV Service Producers*: Ministry of Transport and Communications, 2004.
- [72] A. Ahonen, L. Turkki, M. Saarijärvi, M. Lahti, T. Virtanen, "Guidelines for designing easy-to-use interactive television services: experiences from the ArviD programme," *Interactive digital television: technologies and applications*. IGI Publishing, Hershey, PA, pp. 207–223, 2008.
- [73] K. Chorianopoulos, "User interface design principles for interactive television applications," *International Journal of Human-Computer Interaction*, vol. 24, 2008.
- [74] C. A. Collazos, J. L. Arciniegas, V. Mondragón, X. Garcia, "Lineamientos de usabilidad para el diseño y evaluación de la televisión digital interactiva," *Revista Avances en Sistemas e Informática*, vol. 5, pp. 213-218, 2008.
- [75] R. Brecht, T. Kunert, H. Krömker, "User requirements and design guidance for interactive TV news applications," presented at the European Conference on Interactive Television, User-Centred ITV Systems, Programs and Applications (EuroITV 2005), 2005.
- [76] T. Kunert, *User-centered interaction design patterns for interactive digital television applications*: Springer, 2009.
- [77] D. Quiñones, C. Rusu, S. Roncagliolo, "Redefining Usability Heuristics for Transactional Web Applications," in *11th International Conference on Information Technology: New Generations*, 2014, pp. 260-265.
- [78] G. Kappel, B. Pröll, S. Reich, W. Retschitzegger, *Web engineering: The discipline of systematic development of web applications*: John Wiley & Sons, 2006.
- [79] K. A. Hummel, A. Hess, T. Grill, "Environmental context sensing for usability evaluation in mobile HCI by means of small wireless sensor networks," in *Proceedings of the 6th International Conference on Advances in Mobile Computing and Multimedia*, 2008, pp. 302-306.
- [80] B. Ballard, *Designing the mobile user experience*: John Wiley & Sons, 2007.
- [81] G. Zhao, J. Liu, Y. Tang, W. Sun, F. Zhang, X. Ye, N. Tang, "Cloud computing: A statistics aspect of users," in *Cloud Computing*, ed: Springer, 2009, pp. 347-358.
- [82] T. Wasserman, "Software engineering issues for mobile application development," *FoSER 2010*, 2010.
- [83] K. W. Tracy, "Mobile Application Development Experiences on Apple's iOS and Android OS," *Potentials, IEEE*, vol. 31, pp. 30-34, 2012.
- [84] M. Mehta, I. Ajmera, R. Jondhale, "Mobile Cloud Computing," *International journal of Electronics*, 2013.
- [85] A. Jaokar, "Mobile Cloud Computing: Issues and Risks from a Security Privacy Perspective," in *SecureCloud 2010*, 2010.
- [86] W. J. Orlikowski, J. D. Hofman, "An improvisational model for change management: the case of groupware technologies," *MIT Sloan management review*, vol. 38, pp. 11-22, 1997.
- [87] L. Masip, T. Granollers, M. Oliva, "A heuristic evaluation experiment to validate the new set of usability heuristics," in *International Conference on Information Technology: New Generations (ITNG 2011)*, 2011, pp. 429-434.
- [88] M. González, L. Masip, A. Granollers, M. Oliva, "Quantitative analysis in a heuristic evaluation experiment," *Advances in Engineering Software*, vol. 40, pp. 1271-1278, 2009.
- [89] D. M. Boyd, N. B. Ellison, "Social network sites: definition, history, and scholarship," *Engineering Management Review, IEEE*, vol. 38, pp. 16-31, 2010.
- [90] T. O'Reilly, *What is web 2.0? Design Patterns and Business Models for the Next Generation of Software*: O'Reilly Media Inc, 2009.
- [91] C. Rusu, R. Muñoz, S. Roncagliolo, S. Rudloff, V. Rusu, F. A., "Usability Heuristics for Virtual Worlds," in *The Third International Conference on Advances in Future Internet (AFIN2011)*, 2011, pp. 16-19.

- [92] R. Inostroza, C. Rusu, S. Roncagliolo, V. Rusu, "Usability Heuristics for Touchscreen-based Mobile Devices: Update Proceeding Series.," in *First Chilean Conference on Human-Computer Interaction (ChileCHI2013)*, 2013, pp. 24-29.
- [93] C. Rusu, S. Roncagliolo, G. Tapia, D. Hayvar, V. Rusu, D. Gorgan, "Usability Heuristics for Grid Computing Applications," presented at the The Fourth International Conference on Advances in Computer-Human Interactions, 2011.
- [94] A. Solano, C. Rusu, C. Collazos, S. Roncagliolo, J. L. Arciniegas, V. Rusu, "Usability Heuristics for Interactive Digital Television," presented at the Third International Conference on Advances in Future Internet (AFIN 2011), Nice/Saint Laurent du Var, France, 2011.
- [95] C. K. Coursaris, D. J. Kim, "A meta-analytical review of empirical mobile usability studies," *Journal of Usability Studies*, vol. 6, pp. 117-171, 2011.
- [96] M. C. Trivedi, M. A. Khanum, "Role of Context in Usability Evaluations: A review," *Advanced Computing: An International Journal (ACIJ)*, vol. 3, pp. 69-78, 2012.
- [97] E. Beck, M. Christiansen, J. Kjeldskov, N. Kolbe, J. Stage, "Experimental evaluation of techniques for usability testing of mobile systems in a laboratory setting," in *OzCHI 2003*, Brisbane, Australia, 2003.
- [98] J. Kjeldskov, C. Graham, S. Pedell, F. Vetere, S. Howard, S. Balbo, J. Davies, "Evaluating the usability of a mobile guide: the influence of location, participants and resources," *Behaviour & Information Technology*, vol. 24, pp. 51-65, 2005.
- [99] H. B.-L. Duh, G. C. Tan, V. H.-h. Chen, "Usability evaluation for mobile device: a comparison of laboratory and field tests," in *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services*, 2006, pp. 181-186.
- [100] C. P. van Elzakker, I. Delikostidis, P. J. van Oosterom, "Field-based usability evaluation methodology for mobile geo-applications," *The Cartographic Journal*, vol. 45, pp. 139-149, 2008.
- [101] M. C. L. Yip, E. Mendes, "Web usability measurement: Comparing logic scoring preference to subjective assessment," in *Web Engineering*, ed: Springer, 2005, pp. 53-62.
- [102] J. R. Lewis, "Usability: Lessons Learned... and Yet to Be Learned," *International Journal of Human-Computer Interaction*, vol. 30, pp. 663-684, 2014.
- [103] S. W. Knoll, M. Horning, G. Horton, "Applying a thinkLet-and thinXel-based group process modeling language: A prototype of a universal group support system," in *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*, 2009, pp. 1-10.
- [104] G.-J. d. Vreede, A. Fruhling, A. Chakrapani, "A repeatable collaboration process for usability testing," in *38th Annual Hawaii International Conference on System Sciences (HICSS'05)*, 2005, pp. 46-46.