

# Proceso de evaluación de usabilidad en aplicaciones colaborativas con interfaces de usuario multimodal

Gabriela Sánchez Morales<sup>1</sup>, Carmen Mezura-Godoy<sup>1</sup>, Itzel Alessandra Reyes Flores<sup>1</sup>,  
Edgard Benítez-Guerrero<sup>1</sup>

[gabriesanchez@uv.mx](mailto:gabriesanchez@uv.mx), [cmezura@uv.mx](mailto:cmezura@uv.mx), [itreyes@uv.mx](mailto:itreyes@uv.mx), [edbenitez@uv.mx](mailto:edbenitez@uv.mx)

<sup>1</sup> Facultad de Estadística e Informática, Av. Xalapa, S/N, 91020, Xalapa, Veracruz. México.

DOI: 10.17013/risti.26.68–82

**Resumen:** Las aplicaciones colaborativas (AC), pueden ser evaluadas desde varias dimensiones: contexto grupal, usabilidad, colaboración e impacto cultural. La usabilidad en AC se mide en términos de eficiencia, eficacia y satisfacción, en la que un grupo de usuarios pueden alcanzar objetivos y metas en común. Para la evaluación de usabilidad se han propuesto métodos y heurísticas de evaluación que pueden aplicarse en las diferentes fases del desarrollo de una AC. En este trabajo, presentamos una propuesta de evaluación de usabilidad para aplicaciones colaborativas con interfaces de usuario multimodal, aplicada a los cuatro elementos de una AC: comunicación, colaboración, coordinación y regulación. La propuesta contempla la evaluación en tres etapas: planeación, ejecución y análisis y reporte. La propuesta de evaluación se aplicó a una AC con interfaz multimodal que permite la elaboración de diagramas de flujo.

**Palabras-clave:** usabilidad; evaluación; multimodal; aplicación colaborativa.

## *Usability evaluation of collaborative applications with multimodal user interface*

**Abstract:** Collaborative applications (CA) can be evaluated from several dimensions: group context, usability, collaboration and cultural impact. Usability in CA is measured in terms of efficiency, effectiveness and satisfaction, in which a group of users can achieve common objectives and goals. In order to evaluate groupware usability, several methods and evaluation heuristics have been proposed, they can be applied in the different phases of the development of an AC. In this paper, we present a proposal for usability evaluation for collaborative applications with multimodal user interfaces applied to the four elements of a CA: communication, collaboration, coordination and regulation. The proposal contemplates the evaluation in three stages: planning, execution and analysis and report. The evaluation proposal was applied to an AC with multimodal interface that allows the elaboration of flow diagrams

**Keywords:** Usability, evaluation, multimodal, collaborative applications

## 1. Introducción

El CSCW<sup>1</sup> Lab ha propuesto cuatro dimensiones a considerar al evaluar sistemas colaborativos: contexto grupal, usabilidad, colaboración e impacto cultural (Araujo, 2002). El trabajo realizado por Pinelle y Gutwin (2002) propone que la usabilidad de groupware<sup>2</sup> se define como “la medida en que un sistema de software colaborativo permite que el trabajo en equipo ocurra de manera eficaz, eficiente y satisfactoria para un grupo de usuarios y una actividad en particular”.

Para evaluar la efectividad, eficiencia y satisfacción del software se han desarrollado métodos de evaluación que comprenden un conjunto de actividades bien definidas (Fernandez, 2011). Estos métodos evalúan características específicas dependiendo de las aplicaciones, tales como: Web (Fernandez, 2011; Saltiveri, 2013), móviles (Molano, 2015) y colaborativas (Herskovic, 2007, Quiñonez, 2016). Por ejemplo, en la evaluación de las aplicaciones Web se considera: mantener actualizada la información, direccionar cada liga al elemento correspondiente, entre otros aspectos. En el caso de las aplicaciones móviles se considera: el tamaño de la pantalla, navegación por voz, etc. Para las Aplicaciones Colaborativas (AC o groupware) se han considerados variables de contexto (aspectos personales, situacionales, estructura del grupo, actividades), proceso del grupo (interacción entre los miembros del grupo, comunicación) y las salidas del sistema (relacionadas con las tareas del grupo, como con el grupo mismo) (Pinsonneault, 1989).

Las AC constituyen un software que apoya a los usuarios a alcanzar una meta u objetivo en común. Para ello, las AC proveen los medios para que los usuarios se comuniquen, coordinen sus actividades, colaboren y establezcan reglas para el trabajo en grupo, es decir regulen su actividad (Mezura-Godoy, 2008). Para hacer que estas aplicaciones sean más usables, se han dotado de Interfaces de Usuario Multimodales (IUM), es decir que cuenten con más de una modalidad en la interacción del usuario con el sistema.

La modalidad corresponde a entradas o salidas del sistema por medio de: tacto, voz, gestos, entre otros. Una IUM utiliza diferentes dispositivos de entrada y salida de la computadora, haciendo uso de diferentes canales sensoriales (Moller, 2009). La modalidad puede ser entonces representada por la tupla <lenguaje de interacción, dispositivo físico>. En donde el lenguaje de interacción es un conjunto de expresiones bien formadas que transmiten un significado, mientras que el dispositivo físico se refiere a los periféricos de la computadora que dan la entrada o la salida de información (Vanderdonck, 2008). Las IUM han demostrado tener muchas ventajas comparadas con interfaces monomodales, en cuanto a que facilitan el acceso a la información y a la funcionalidad del sistema a través de una gama de diferentes canales de entrada y salida (Dix, 2009), proveen robustez a la interfaz, ayudan al usuario a corregir errores o recuperarse más fácilmente, y añaden métodos de comunicación alternativos a diferentes situaciones y ambientes (Jaimes, 2007). Ejemplos de AC con IUM pueden ser: 1) DISCIPLE (Dorohonceanu, 2000), esta aplicación apoya a los usuarios en la toma de decisiones en un contexto militar, las modalidades de entrada corresponden a reconocimiento de voz, manipulación directa mediante un guante táctil y rastreador

<sup>1</sup> Computer Support Collaborative Work

<sup>2</sup> Término en inglés para denominar a las aplicaciones o software colaborativo.

de mirada, mientras que las de salida corresponde a voz. 2) TowerWorld, diseñado para compartir archivos entre los usuarios con una interfaz con modalidades de reconocimiento de voz y manipulación directa con mouse y teclado, y 3) AudioSpace Invaders, que es un juego colaborativo para personas ciegas que posee una interfaz de audio que permite a los usuarios especificar contextos que les interesa.

En el proceso de desarrollo de software se hace necesario incluir tareas tendientes a la evaluación no sólo funcional, no funcional o de interfaz, sino también en el caso del software colaborativo es necesario evaluar aspectos sociales tales como: conformación del grupo, la interacción entre sus miembros, la comunicación, la colaboración, las reglas del trabajo en grupo, la presencia del grupo, etc. Estos trabajos nos han llevado a incluir en el componente de interfaces diversas modalidades. En trabajos previos hemos propuesto modelos y arquitecturas que contemplen la definición y desarrollo de AC (Mezura-Godoy, 2008; Montané-Jimenez, 2012; Decouchant, 2013) y de AC con interfaces multimodales (Reyes, 2016). Particularmente, en este trabajo nos interesamos en la evaluación de las interfaces de usuario multimodales de las AC. Existen trabajos que proponen la evaluación de la usabilidad de las AC en la fase de diseño de las interacciones (Herskovic, 2007), otros sugieren evaluar no sólo la usabilidad sino otros aspectos como es el impacto cultural (De Araujo, 2004). Aun cuando estos trabajos consideran evaluar varias características de las AC, no se contemplan a las AC con IUM. Por ello, en este trabajo hacemos una propuesta de evaluación de usabilidad de las AC con IUM.

Este artículo se encuentra organizado de la siguiente manera: en la sección 2 se presenta el estado de arte acerca de los métodos de evaluación de las AC. En la sección 3, se describe la propuesta de este trabajo para evaluar la usabilidad de las AC con IUM mediante pruebas de usabilidad con usuarios y basada en heurísticas; en la sección 4, se presenta la aplicación de las pruebas de evaluación que se propone, en una AC que apoya el diseño colaborativo de diagramas de flujo. Finalmente, la sección 5 se presenta las conclusiones y trabajo futuro.

## **2. Métodos de evaluación de AC**

Las aplicaciones colaborativas (AC) son “sistemas basados en computadoras que soportan grupos de personas que participan en una tarea común y que proveen una interfaz para un ambiente compartido” (Ellis, 1991). Las AC apoyan a los grupos de personas en: la comunicación, la coordinación, la colaboración (3C's) y la regulación (Mezura-Godoy, 2008). La arquitectura de este tipo de sistemas comprende cinco componentes: i) funcional (funcionalidades propias de la aplicación), ii) social (reglas de colaboración, interacciones sociales, roles, objetos a manipular, etc.), iii) individual (características, habilidades, preferencias, etc.), iv) no funcional (seguridad, conectividad, manejo de versiones, etc.) y v) de interfaz de usuario y de grupo. Particularmente, la interfaz en este tipo de sistemas es muy importante, ya que tiene que dar cabida a aspectos particulares del usuario (gustos, preferencias, habilidades, etc.), así como aspectos del trabajo en grupo (¿quién está haciendo qué?, ¿cómo lo está haciendo?, ¿qué objetos manipula?, ¿cuándo los usuarios participaron?), entre otros.

Las AC pueden ser evaluadas bajo un enfoque de usabilidad teniendo en cuenta la efectividad, eficiencia y satisfacción. La efectividad puede ser medida en términos de

la precisión y completitud con la que el usuario alcanza sus objetivos, la eficiencia se refiere al tiempo y esfuerzo para alcanzar los objetivos, y la satisfacción es medible en términos de la ausencia de incomodidad y una actitud positiva hacia el uso de la aplicación (Gutwin, 2000).

Se ha propuesto incorporar varios métodos para evaluar las AC a lo largo de las diferentes etapas del desarrollo (análisis, diseño e implementación, previo a la liberación del producto) como se observa en la Tabla 1. En el análisis se requiere medir la viabilidad del sistema mediante los métodos (Herskovic, 2007): Heurísticas de Evaluación de Colaboración (HEC), y Valor Percibido (Herskovic, 2007). En el diseño, los métodos HEC, Walkthrough Colaborativo (WC), Análisis de Usabilidad de la Colaboración (UAC), Pruebas de Usuario Colaborando (OPUC) y Etnografía rápida y sucia (ERS) ayudan a determinar la correcta implementación de los mecanismos de las AC (Herskovic, 2007). Se evalúan las AC liberadas o en etapas previas a la liberación, para conocer la relación de los usuarios con sus necesidades, así como la efectividad y eficiencia; algunos de los métodos sugeridos corresponden a HEC, WC, AUC y OPUC. Para Herskovic (2007), no es recomendable una combinación específica de las tres etapas indicadas, en su caso dejan al evaluador que determine un método dado el tiempo y el costo de cada una de ellas (Ver Tabla 1).

Método	Etapas	Objetivo
H-E y Valor Percibido	Análisis	Medir impacto organizacional y las expectativas de los diseñadores y usuarios
HEC, WC, UAC, OPUC y ERS	Diseño y desarrollo	Verificar correcta implementación de mecanismos de las AC y encontrar defectos
HEC, WC, AUC y OPUC	Liberado o previo a la liberación	Medir la eficiencia y la efectividad y la relación entre funcionalidad, usuarios y las necesidades de la organización

Tabla 1 – Métodos de evaluación por etapas de desarrollo software

Para evaluar las AC, trabajos previos han modificado los métodos usados para evaluar aplicaciones mono-usuarios (heurísticas de evaluación, walkthrough cognitivo, observación de las pruebas de usuario, análisis de tarea, etc.). Algunos de los métodos que permiten evaluar las AC son los siguientes (Herskovic, 2007):

- HEC evalúa la comunicación de los usuarios distribuidos y la colaboración con artefactos a través de un espacio visual compartido. HEC se basa en 8 heurísticas relacionadas con los medios para la comunicación (verbal intencional, gestual intencional, adecuada, acciones corporales de los usuarios, artefactos compartidos, entre otras), la protección al trabajo y la administración de las transiciones de la colaboración (estrecha y ligeramente acoplada); finalmente parte de estas 8 heurísticas brindan apoyo a las personas para coordinar sus acciones, buscar colaboradores y contactarlos (Baker, 2002).
- En WC se observa a usuarios para identificar episodios de colaboración y construir escenarios como una descripción de la actividad o conjunto de tareas. Se incluye al usuario, resultados planeados y circunstancias alrededor. Un

evaluador puede tomar uno o todos los roles y seguir una tarea recordando cada problema encontrado.

- Pruebas de Usuario Colaborando (OPUC): Algunos métodos existentes parten de observar la colaboración entre los usuarios, ya sea para analizar algún problema en especial (Herskovic, 2007).
- ERS se basa en la etnografía para buscar problemas de aceptabilidad y usabilidad en sistemas actuales que puedan ser corregidos en sistemas futuros o dar simplemente una idea general a los diseñadores de cómo se lleva a cabo el trabajo colaborativo (Herskovic, 2007);
- Análisis de Usabilidad de la Colaboración (AUC): método que se basa en un análisis jerárquico de tareas individuales y grupales, permite ejecutar una tarea usando rutas alternativas y tareas opcionales; las tareas grupales deben relacionarse con mecanismos colaborativos que puedan verse reflejados en la interfaz de usuario (Pinelle, 2003).
- Humano-Ejecución (H-E): divide la interfaz en espacios compartidos y definen varios escenarios críticos de las actividades colaborativas, con la finalidad de comparar la ejecución de estos (Herskovic, 2007).

Es importante señalar que para llevar a cabo una evaluación se debe seguir un plan de trabajo bien definido. Para ello, De Araujo (2004), Yu (2013), y Rogers (2011) coinciden en establecer los siguientes puntos para realizar un plan de evaluación, que consiste en: establecer un objetivo, seleccionar el método de evaluación, determinar el escenario de interacción, seleccionar y reclutar el grupo de participantes, programar las tareas y fechas, recolectar y analizar resultados.

### **3. Evaluación de AC-IUM mediante pruebas de usabilidad y heurísticas**

La propuesta de evaluación de AC-IUM se basa en el proceso general de la evaluación de usabilidad en tres etapas (Sánchez, 2018): 1) planeación para definir los objetivos de la prueba, el método de evaluación, la selección de los usuarios, el laboratorio, entre otros aspectos; 2) ejecución es llevar a cabo todo lo planeado y 3) análisis y reporte de los datos obtenidos. En este trabajo estas tres etapas son específicas para evaluar Aplicaciones Colaborativas con Interfaces de Usuario Multimodal (AC-IUM), mediante pruebas de usabilidad y basadas en heurísticas. En esta propuesta se considera que las pruebas de usabilidad son las que se llevan a cabo con usuarios, mientras que las basadas en heurísticas son ejecutadas por evaluadores expertos en usabilidad.

Algunas actividades en el proceso de evaluación pueden usarse tanto en las pruebas de usabilidad como en las basadas en heurísticas, pero otras no; por tal razón, en esta propuesta, se especificará cuales actividades pueden ser aplicadas para ambos casos o solo para uno. Además de lo anterior, en este trabajo se ofrecerán sugerencias con la finalidad de ayudar a personas con poca experiencia en la realización de dichas pruebas. A continuación se explican con detalle cada una de las tres etapas mencionadas para la evaluación de usabilidad de las AC-IUM.

### 3.1. Etapa de planeación

En la planeación de la evaluación se consideran 4 aspectos: 1) objetivos y preguntas de apoyo, 2) tareas, 3) recursos y 4) plan de ejecución. Se recomienda realizar formatos desde las tareas, con la finalidad de tener una mejor organización.

#### 1) *Objetivos y preguntas de apoyo*

Definir los objetivos de la evaluación, los cuales frecuentemente responden a qué tan usable es la aplicación considerando cada una de las modalidades empleadas; mientras que las preguntas de apoyo ayudan a resolver dichos objetivos; por ejemplo, ¿podrán los usuarios enterarse del ingreso de un usuario mediante la modalidad de audio? o ¿podrán los usuarios colaborar para el lograr el objetivo deseado mediante la modalidad textual?

#### 2) *Tareas*

Consiste en determinar las tareas individuales y grupales considerando la (las) modalidad(es) de interacción; dichas tareas frecuentemente se relacionan con las preguntas de apoyo o con los mecanismos de colaboración (comunicación, coordinación, colaboración y regulación). Algunas tareas individuales pueden ser configurar el entorno personal o realizar una búsqueda mediante navegación por voz o textual. Mientras que las tareas grupales pueden relacionarse con los mecanismos de colaboración (comunicación, colaboración, coordinación y regulación según la clasificación de (Sánchez, 2017)) y sus diferentes modalidades.

En las *pruebas de usabilidad*, una actividad relacionada con las tareas es establecer el tiempo de cada una; con la finalidad de poner tiempo límite para realizar estas y su vez reportar cuales fueron realizadas en el tiempo estimado. Sugerencias: 1) considerar tiempo para que los usuarios exploren la aplicación; 2) realizar un formato para los usuarios, en donde se incluyan las tareas a realizar, un escenario por cada tarea y preguntas relacionadas con la satisfacción, y 3) otro formato para los observadores con las tareas seleccionadas y los eventos que se sugieren observar.

#### 3) *Recursos*

En los recursos se consideran los **humanos (RH)**, los **físicos** y los de **evaluación**.

Los **RH** incluyen el equipo de pruebas y la selección de usuarios o evaluadores expertos. El equipo de las pruebas se refiere a las personas encargadas de realizar las pruebas, cuyos roles corresponden a: Moderador, Observador y Técnico. El Moderador conducirá toda la prueba que abarca desde dar la bienvenida a los participantes hasta finalizar las pruebas. El Observador es empleado en las *pruebas de usabilidad* para registrar: tiempo empleado en cada tarea, dificultades encontradas y el estado de ánimo de los usuarios al interactuar con el sistema en cuestión, etc. El Técnico verificará las instalaciones físicas requeridas y de software, la conexión a internet, etc.; en el caso de las *pruebas de usabilidad*, una actividad más es preparar lo requerido al momento de interactuar con diferentes modalidades (micrófono y audífonos para interactuar por



voz, etc.). *Sugerencia en las pruebas de usabilidad:* 1) establecer el perfil deseado en la selección de usuarios, por ejemplo el nivel de experiencia, rango de edades, género, preferencias de las modalidades de interacción, etc. 2) Formar subgrupos de usuarios heterogéneos respecto a sus preferencias de modalidades.

Los **recursos físicos** consideran el hardware, el software y el lugar para realizar las pruebas. El hardware corresponde a los dispositivos en donde se ejecutará la aplicación a evaluar; en el caso de las *pruebas de usabilidad* se incluyen lo necesario para grabar a los usuarios, así como equipo auxiliar que permitan la interacción con diferentes modalidades. Mientras que el software se refiere a la aplicación a evaluar; para las *pruebas de usabilidad* se deben considerar las herramientas de grabación (interacciones y el rostro del usuario). El recurso físico también incluye el lugar en donde se realizaran las pruebas, como son laboratorios de usabilidad o habitaciones acondicionadas. El recurso físico en el caso de *pruebas de usabilidad* es aconsejable designar a cada usuario en cuartos diferentes, en el caso que la modalidad de interacción sea por voz.

El **recurso de evaluación** se refiere a seleccionar el método de evaluación (Sánchez, 2017) para determinar la usabilidad en términos de efectividad, eficiencia y satisfacción, considerando las heurísticas ligadas a los aspectos de las AC (comunicación, coordinación, colaboración y regulación), mismos que se observan en la Tabla 2. Dichos métodos deben considerar cada una de las modalidades, por ejemplo, entrada de voz, o salida de audio.

En el aspecto de comunicación se considera 3 heurísticas: la heurísticas 1 que provee los medios para la comunicación verbal intencional y apropiada, la heurística 2 proporciona los medios para la comunicación gestual intencional y adecuada, mientras que la heurística 3 provee comunicación originada por las acciones corporales de los usuarios. En general, las heurísticas de este aspecto pueden considerar modalidades como entrada de voz, salida de audio, entrada de gráfico, etc. Por tal razón, es importante evaluar la usabilidad de las heurísticas considerando sus diferentes modalidades.

Mecanismos de las AC	Heurística	Método		
		Efectividad	Eficiencia	Satisfacción
Comunicación	1,2,3			
Coordinación	4,5	OPUC, WC, UAC	HEC, OPUC, WC	Cuestionarios de satisfacción considerando las modalidades
Colaboración	6,8			
Regulación	7, 9			

Tabla 2 –Métodos de evaluación a partir de los mecanismos de las AC y las heurísticas

El aspecto de coordinación considera las heurísticas 4 y 5 (Ver Tabla 2), la heurística 4 proporciona comunicación consecuente de artefactos compartidos, a su vez la heurística 5 provee protección al trabajo. Es importante mencionar que dadas las diferentes modalidades que se pueden emplear en estas heurísticas, es necesario medir la usabilidad que ofrecen. Por ejemplo, no es lo mismo comunicar por audio los cambios en un artefacto compartido que comunicarlo de manera textual.

Mientras que la colaboración está compuesta por las heurística 6 que se encarga de administrar las transiciones entre la colaboración estrechamente y ligeramente acoplada; adicionalmente, la heurística 8 facilita la búsqueda de colaboradores y la manera de establecer contacto.

En el aspecto de regulación de una AC se consideran la heurística 7 que se relaciona con la asignación de tareas a los participantes de la actividad, y cómo éstas se tienen que llevar a cabo para alcanzar la tarea final. La heurística 9 que provee a los usuarios los medios para la definición de la regulación social [5].

Tanto en el aspecto de colaboración como de regulación se debe considerar las modalidades que permiten cada uno de ellos.

A partir de los mecanismos de las AC y las heurísticas de evaluación propuestas se propone evaluar la usabilidad considerando la: efectividad, eficiencia y satisfacción.

- Efectividad: se refiere a la completitud de las tareas individuales y grupales, considerando cada una de las modalidades establecidas, por ejemplo: colaborar en la redacción de un documento mediante la modalidad de entrada de voz. La completitud puede definirse mediante WC, AUC u OPUC; ya que la primera está relacionada con el recorrido de las tareas, la segunda con el desglose de las tareas; mientras que la última se basa en observar la colaboración.
- Eficiencia: es reportar las tareas individuales o grupales terminadas dentro o fuera del tiempo esperado, considerando la complejidad que puede estar asociada a una modalidad, por ejemplo: la modalidad de audio asociada al *awarness*. La eficiencia puede calcularse mediante HEC, OPUC o WC; las dos primeras pueden servir para determinar el tiempo; también el segundo y el tercer modelo pueden ser empleados para saber el esfuerzo requerido por los usuarios.
- Satisfacción: se obtiene al hacer cuestionarios o entrevistas que involucren las modalidades utilizadas en cada una de las heurísticas consideradas en cada aspecto de las AC.

#### 4) Plan de ejecución

Se considera 2 actividades: 1) Redactar un guión de ejecución con los siguientes puntos: agradecimiento, explicación de la aplicación a evaluar, objetivo de la evaluación, actividades a realizar; indicar si serán observados al momento de interactuar, y 2) Realizar el programa de actividades de acuerdo al guión. En esta parte se sugiere: 1. Hacer un formato para el guión y establecer el tiempo por cada actividad; si las realizaciones de las tareas duran más de 25 minutos se debe tener un descanso. 2. Hacer un formato tipo *checklist* que considere actividades a realizar antes, durante y después de la prueba (Sánchez, 2018).

### 3.2. Etapa de ejecución

El día de la prueba se debe tener a la mano todos los formatos en el orden que serán usados, llegar con anticipación al lugar de la evaluación para preparar lo necesario. Durante la prueba se debe seguir el guion o el *checklist* correspondiente.



En esta etapa se sugiere que para *las pruebas de usabilidad*: 1) El moderador debe emplear tiempo antes de la prueba para platicar con el usuario con la finalidad de relajarlos y hacerlos sentir en confianza, 2) Evitar horarios usuales de comida, ya que los usuarios no estarán concentrados en la actividad, 3) El equipo de evaluación debe ser amable con los usuarios, no sorprenderse o emitir sonidos cuando cometan errores los usuarios, evitar distracciones (teléfonos, visitas personales, platicar, etc.), 4) Al finalizar las pruebas invierte tiempo con los usuarios, puesto que ellos en ocasiones proveen más información, 5) Reúnete con tu equipo de evaluación para hablar de mejoras en las pruebas, y 6) Al finalizar cada sesión o todas las pruebas platica con los observadores para recopilar sus observaciones, detectar problemas graves sobre la aplicación.

### **3.3. Etapa de análisis y reporte**

Al finalizar las pruebas se debe hacer un reporte que incluya una: explicación breve acerca de la aplicación a evaluar, una descripción por un lado de la planeación siguiendo el punto 3.1, y por otro del proceso que se llevó cabo en la ejecución de la prueba. Finalmente, el reporte incluye el análisis del resultado.

El análisis de los resultados obtenidos incluye el total de los usuarios o evaluadores expertos, cuántos son novatos, intermedios y expertos. Los resultados relacionados con la efectividad, eficiencia y satisfacción se representa usando tablas, gráficos, los valores mínimos y máximos, porcentaje o el número de éxitos del total de usuarios. En donde la:

- La afectividad debe ser reportada al describir la completitud de las tareas individuales y grupales, considerando cada una de las modalidades establecidas.
- La eficiencia: es reportar el tiempo consumido en las tareas individuales o grupales terminadas. En el caso de las pruebas de usuario se reportan las tareas terminada en el tiempo esperado o fuera de tiempo.
- La satisfacción es reportado al mencionar la preferencia de la modalidad de cada tarea reportada por los usuarios finales, dichos resultados son obtenidos de los cuestionarios o entrevistas.

## **4. Evaluando Co-flows usando la metodología propuesta**

En esta sección se describe la aplicación de la propuesta de evaluación en una AC denominada “Co-flows”. Esta AC permite a un grupo de usuarios, ubicados en diferentes espacios físicos, colaborar al mismo tiempo en la creación de diagramas de flujo. La aplicación en su componente funcional ofrece medios para la colaboración, comunicación, coordinación y regulación.

Para la colaboración se ofrecen funcionalidades que permitan la creación, modificación y guardado de diagramas de flujo; inserción, modificación y eliminación de objetos del diagrama tales como: círculos, rombos, rectángulos, entre otros y acceso a un área de dibujo común.

En la comunicación se provee de un chat que permite enviar y recibir mensajes entre los usuarios; dichos mensajes pueden enviarse y recibirse mediante la modalidad gráfica, textual y audio. Los mensajes pueden ser enviados de una modalidad y recibirse en otra

dadas las preferencias y características del ambiente de cada usuario. La comunicación, también ofrece un servicio de notificación para informar de las acciones realizadas en el espacio común y *awareness*; este servicio habilita la modalidad de audio y texto dada las preferencias del usuario y del ambiente de él.

Para la coordinación y regulación, se ha considerado la asignación de roles y actividades a los usuarios, con la finalidad de permitir a los usuarios una mejor coordinación.

Co-flows se ha evaluado en diferentes etapas del desarrollo del software (Sánchez, 2017). En este artículo se muestra el proceso de evaluación de usabilidad de Co-flows en una versión completa.

#### 4.1. Planeación

El *objetivo* de esta prueba es medir qué tan usable es Co-flows al ofrecer diversas modalidades para la interacción.

En esta evaluación se consideran las *tareas grupales* como se observa en la Tabla 3, en donde se describe cada tarea, el tiempo considerado y las modalidades que serán empleadas durante la ejecución de la tarea solicitada.

Respecto a los participantes (*Recursos humanos*) se consideran 5 equipos formados entre 3 y 4 estudiantes de semestres similares, con experiencia en el uso de aplicaciones colaborativas, además de identificarse por su canal sensorial preferente, obtenido de un pre-test.

No.	Tarea	Descripción de la tarea	Modalidades a evaluar	Tiempo estimado
1	Explorar la aplicación	Los usuarios pueden explorar la aplicación colaborativa	ninguna	5 min.
2	Elegir un coordinador	Los usuarios deben coordinarse usando el chat		5 min
3	Realizar un diagrama de flujo	El equipo de manera grupal realiza un diagrama de flujo, que representa el envío de un correo electrónico desde un navegador Web.	Audio, texto	16 min

Tabla 3 – Tareas grupales solicitadas a los usuarios

El *equipo de pruebas* será formado por 2 personas para desarrollar el rol de técnico antes de la prueba; durante la prueba uno será observador y otro moderador. El moderador debe tener experiencia en la realización de pruebas.

El *hardware* requerido: cuatro computadoras personales (Laptop) con webcam para las grabaciones y audífonos. El *software* requerido es un programa para grabar a los usuarios y Co-flows instalado en una computadora. Los *lugares* para realizar las pruebas serán laboratorios de pruebas, cuartos de observación y laboratorio de cómputo de dos universidades públicas, una en Puebla y otra en Veracruz. Se elaboró un formato de listado de actividades relacionadas con la preparación de hardware y software para el

día de la evaluación y un formato de consentimiento por parte de los usuarios para ser video grabados.

En el *recurso de evaluación* se aplicó el método OPUC que midió la efectividad en todos los mecanismos de las AC mientras que la satisfacción se midió usando un cuestionario basado en las 6 heurísticas implementadas. Dicho cuestionario permitirá conocer cuál de todas las modalidades ofrecidas en la comunicación, colaboración, coordinación y regulación es más satisfactoria.

Es importante mencionar que la eficiencia es medible respecto al tiempo y esfuerzo; por lo tanto, en esta evaluación, únicamente se considera en el mecanismo de colaboración; debido a que es muy notorio identificar el tiempo empleado para realizar el diagrama de flujo; mientras que para el resto de los mecanismos no es evidente.

En la planeación se consideraron dos formatos a llenarse: 1) Guion de ejecución y 2) Programa de actividades antes, durante y después de las pruebas

Mecanismos de las AC	Heurística	Método		
		Efectividad	Eficiencia	Satisfacción
Comunicación	1,2,3	OPUC	-	Cuestionarios de satisfacción considerando las modalidades
Coordinación	4		-	
Colaboración	8		OPUC	
Regulación	7		-	

Tabla 4 – Métodos de evaluación considerados en Co-flows

#### 4.2. Ejecución

Las pruebas se llevaron a cabo durante cinco días, un equipo cada día, en cada prueba se siguió el formato “Programa de actividades” y los demás formatos generados.

Tres pruebas se llevaron a cabo en el laboratorio de usabilidad de la universidad de Puebla y en un cuarto de observación; debido a que dos usuarios se ubicaban en el laboratorio y el resto (dos o un participante) en el área de observación. Dos pruebas más se realizaron en un laboratorio de cómputo, en la universidad de Veracruz. Durante la ejecución de las 5 pruebas se colocó a los usuarios en lugares separados cada uno en diferente computadora, con el mayor silencio posible y evitando interrupciones externas.

Debido a que Co-flows permite usar diferentes modalidades, se decidió cambiar estas aleatoriamente, para que los usuarios pudiesen evaluar las diferentes modalidades que ofrece Co-flows. Durante la ejecución de las tareas solicitadas se grabaron las interacciones en pantalla y los rostros de cada usuario, para obtener datos que pudiesen ayudar en determinar la efectividad y la eficiencia.

### 4.3. Reporte y análisis

Participaron 16 estudiantes de carreras afines al área de computación, el 62% eran de 8° semestre, los restantes eran de 10° semestre; todos con experiencia en el manejo de aplicaciones colaborativas. El 56% de los participantes tenían preferencia por el canal sensorial visual, el 25% kinestésico y 19% auditivo.

La efectividad relacionada al mecanismo de comunicación, se puede decir que Co-flows si permitió a los usuarios comunicarse; sin embargo se presentaron algunos errores relacionados a la modalidad de audio, ya que el traductor usado para cambiar de audio a texto no funcionaba bien; otro error fue que el audio se cicló repitiendo el mismo mensaje todo el tiempo o se atrasaban los audios.

Analizando la efectividad de Co-flows en el mecanismo de coordinación; esta aplicación si permitió coordinar a los usuarios sin problemas; ya que retroalimentaba a los usuarios cuando los objetos o herramientas estaban siendo ocupados por alguien más, e informaba de lo que están haciendo los usuarios.

Modalidad	%Satisfacción- Comunicación			%Satisfacción- Coordinación		
	Completa	Media	Nula	Completa	Media	Nula
Audio	64.1	31.3	4.7	68.7	25	6.3
Texto	<b>87.5</b>	12.5	0	<b>81.2</b>	18.8	0
Gráfico	59.4	37.5	3.1	68.7	31.3	0

Tabla 5 – Niveles de satisfacción de la comunicación y coordinación

En la Tabla 5 se muestra la satisfacción en los mecanismos de comunicación y coordinación. Respecto a la comunicación y la coordinación, la modalidad textual proporcionó mayor satisfacción en ambos casos.

El mecanismo de colaboración fue evaluado por la efectividad, eficiencia y satisfacción. Los datos correspondientes a la efectividad mostraron que la colaboración tuvo un poco de problemas, ya que un equipo logró terminar el diagrama, pero los demás quedaron de la siguiente forma: 65%, 56%, 45% y 31% de completitud de la tarea solicitada. Respecto a los problemas encontrados al momento de colaborar, uno de ellos ocurrió al momento de agregar una etiqueta de texto en el área de dibujo, la aplicación agregaba dos, otro problema se relacionó con el equipo que mostró menos avance, ya que no trabajaron en equipo.

La efectividad respecto a la regulación indica que los usuarios pudieron regular sus actividades poco, ya que Co-flows informaba quiénes están conectados pero no era evidente saber si estaban disponibles.

Respecto a la eficiencia en la colaboración, el tiempo para hacer un diagrama variaba entre 15:40 y 11:02 minutos.

La Tabla 6 muestra la satisfacción de la colaboración y la regulación en donde se observa que la modalidad que les dio mayor satisfacción fue la textual.

Al finalizar la prueba, algunos comentarios de los usuarios coincidían en que les parecía interesante utilizar todas las modalidades para interactuar con la aplicación, pero que les gustaría que fuera una combinación de todas, y que ellos pudieran elegir cuándo utilizar una u otra.

Modalidad	%Satisfacción- colaboración			%Satisfacción- regulación		
	Completa	Media	Nula	Completa	Media	Nula
Audio	75	18.8	6.2	62.5	31.2	6.3
Texto	<b>81.2</b>	18.8	0	<b>75</b>	25	0
Gráfico	75	18.7	6.3	68.7	25	6.3

Tabla 6 – Niveles de satisfacción en la colaboración y regulación

## 5. Conclusiones

La evaluación de la usabilidad de un AC con interfaz de usuario multimodal puede resultar una tarea muy complicada para los evaluadores ya sea novatos o expertos. Ya que además de considerar que la AC provea de los mecanismos de comunicación, colaboración, coordinación y regulación, es necesario identificar el apoyo que recibe el usuario al incluir varias modalidades en la interacción con el sistema y los otros usuarios. Por ello, en este trabajo se propone orientar respecto al proceso de evaluación de usabilidad de las AC-IUM; así como sugerir algunos métodos de evaluación. En esta propuesta la evaluación de AC considera la satisfacción, eficiencia y efectividad, mediante la agrupación de heurísticas de acuerdo a los mecanismos de las AC (colaboración, comunicación, coordinación y regulación).

La propuesta de evaluación de AC-IUM se valida aplicándola a Co-flows, un software de construcción de diagramas de flujo. Al evaluar la usabilidad de la aplicación Co-flows, siguiendo el proceso propuesto en este artículo, se observó una mejor organización de la evaluación, incluso se facilitó seleccionar el método de evaluación adecuado ya que se describe de manera general como evaluar la usabilidad mediante estos métodos; por otro lado fue más fácil seleccionar las tareas ya que estas están relacionadas con los mecanismos de las AC (comunicación, colaboración, coordinación y regulación) y a su vez con las heurísticas. Además la propuesta de evaluación, ayudó a determinar cómo calcular la eficiencia, efectividad y satisfacción de las AC-IUM.

Como trabajo futuro se propone proveer multimodalidad a otros componentes de la interfaz de usuario, con objeto de evaluar otros aspectos de la AC a más profundidad. También es deseable, reclutar usuarios con diferentes canales sensoriales para realizar otras pruebas y comparar resultados. Finalmente, se desea validar las heurísticas y los métodos propuestos con otras aplicaciones que están en desarrollo actualmente.

## Referencias

- Araujo, R., Santoro, F., & Borges, M. (2002). The CSCW lab for groupware evaluation. In Proceedings of *CRIWG'02*, (Vol 2440, pp. 222-231). Lecture Notes in Computer Science.
- Baker, K., Greenberg, S., & Gutwin, C. (2002). Empirical development of a heuristic evaluation methodology for shared workspace groupware. In Proceedings of *the 2002 ACM conference on Computer supported cooperative work* (pp. 96-105). ACM.
- Barnum, C. M. (2010). Usability testing essentials: ready, set... test! (pp. 11-13). Elsevier.
- De Araujo, R. M., Santoro, F. M., & Borges, M. R. S. (2004). The CSCW lab ontology for groupware evaluation. In Proceedings of *Computer Supported Cooperative Work in Design, 2004*. The 8th International Conference on (Vol. 2, pp. 148-153). IEEE.
- Dix, A. (2009). Human-computer interaction. In: *Encyclopedia of database systems* (pp. 1327-1331). Springer US.
- Decouchant, D., Mendoza, S., Sánchez, G., & Rodríguez, J. (2013). Adapting groupware systems to changes in the collaborator's context of use, *Expert System with Applications*, 40 (11), 4446–4462.
- Dorohonceanu, B., Sletterink, B., & Marsic, I. (2000). A novel user interface for group collaboration. In Proceedings of *the 33rd Annual Hawaii International Conference on* (pp. 10-pp). IEEE.
- Ellis, C. A., Gibbs, S. J., & Rein, G. (1991). Groupware: some issues and experiences. *Communications of the ACM*, 34(1), 39–58.
- Fernandez, A., Insfran, E., & Abrahão, S. (2011). Usability evaluation methods for the web: A systematic mapping study. *Information and Software Technology*, 53(8), 789–817.
- Gutwin, C., & Greenberg, S. (2000). The mechanics of collaboration: Developing low cost usability evaluation methods for shared workspaces. In *Enabling Technologies: Infrastructure for Collaborative Enterprises, 2000*. (WET ICE 2000). *Proceedings. IEEE 9th International Workshops on* (pp. 98-103). IEEE.
- Herskovic, V., Pino, J. A., Ochoa, S. F., & Antunes, P. (2007). Evaluation methods for groupware systems. In Proceedings of *International Conference on Collaboration and Technology* (pp. 328-336). Berlin, Heidelberg: Springer.
- Jaimes, A., & Sebe, N. (2007). Multimodal human-computer interaction: A survey. *Computer vision and image understanding*, 108(1-2), 116–134.
- Montané-Jiménez, L., G., Benítez-Guerrero, E., & Mezura-Godoy, C. (2014). Towards a Context-Aware Framework for Improving Collaboration of Users in Groupware Systems. *EAI Endorsed Transactions on Context-aware Systems and Applications*, 14(1), e4.
- Mezura-Godoy & C. Montané-Jiménez, L. (2008) Modelling Regulated Social Spaces for Groupware Applications, *Research in Computing Science. Advances in Computer Science and Artificial Intelligence*, 39, 47–60.



- Molano, J. I. R., Yara, E. S., & García, L. K. J. (2015). Model for measuring usability of survey mobile apps, by analysis of usability evaluation methods and attributes. In *Proceedings of Information Systems and Technologies (CISTI), 2015 10th Iberian Conference on* (pp. 1-6). IEEE.
- Moller, S., Engelbrecht, K. P., Kuhnel, C., Wechsung, I., & Weiss, B. (2009). A taxonomy of quality of service and quality of experience of multimodal human-machine interaction. In *Proceedings of Quality of Multimedia Experience, 2009. QoMEX 2009. International Workshop on* (pp. 7-12). IEEE.
- Pinelle, D., Gutwin, C., & Greenberg, S. (2003). Task analysis for groupware usability evaluation: Modeling shared-workspace tasks with the mechanics of collaboration. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 10(4), 281–311.
- Pinelle, D., & Gutwin, C. (2002). Groupware walkthrough: Adding context to groupware usability evaluation. In *Proceedings of the 2002 SIGCHI Conference on Human Factors in Computing Systems* (pp. 455–462). ACM Press.
- Pinsonneault, A., & Kraemer, K. (1989). The impact of technological support on groups: An assessment of the empirical research. *Decision Support Systems*, 5(3), 197–216.
- Quiñonez, Y., Luzardo, G. & Granda, R. (2016). Implementación de un sistema multi-táctil en ambientes educativos para promover y facilitar la evaluación del trabajo colaborativo en el aula. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, 17, 57–65.
- Reyes, I. A., Mezura-Godoy, C., & Sánchez, G. (2016). Hacia un modelo de interfaces multimodales adaptables a los canales de aprendizaje en aplicaciones colaborativas como apoyo a la educación. *Research in Computing Science*, 111, 57–67.
- Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design: beyond human-computer interaction*. Hoboken, NJ, USA: John Wiley & Sons.
- Saltiveri, T. G., Fransi, E. C., & Salla, Y. M.. (2013). Análisis de usabilidad de cooperativas del sector de la fruta y aceite en el área de Lleida. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, 11, 45–59.
- Sánchez, G., Mezura-Godoy, C., Reyes, I. A., & Benítez-Guerrero (2017, October). Evaluación de usabilidad de aplicaciones colaborativas con interfaces de usuario multimodal, Applications in Software Engineering. In *Proceedings of the 6th International Conference on Software Process Improvement (CIMPS)* (pp. 119,127), IEEE.
- Sánchez, G., Mezura-Godoy, C., & Benítez-Guerrero (2018). Proceso de Pruebas de Usabilidad de Software. *Tecnología Educativa Revista CONAIC*, 5(1).
- Vanderdonckt, J., Calvary, G., Coutaz, J., & Stanciulescu, A. (2008). Multimodality for plastic user interfaces: models, methods, and principles. In *Multimodal User Interfaces* (pp. 61-84). Berlin, Heidelberg: Springer.
- Yu, Z., Guo, W., & Zheng, X. (2013, October). A brief review: Groupware evaluation by performance of specific tasks. In *TENCON 2013-2013 IEEE Region 10 Conference* (31194) (pp. 1-4). IEEE.