

Una familia de experimentos para analizar las preferencias de los usuarios en el diseño de IGUs a partir de modelos BPMN

A Family of experiments to analyze the users' preferences when designing GUIs from BPMN models

Eduardo Díaz

Departamento Escola
Tècnica Superior
d'Enginyeria, Departament
d'Informàtica / Universidad
de Valencia, Avenida de la
Universidad, s/n, 46100,
Burjassot, València, España
diazsua@alumni.uv.es

José Ignacio Panach

Departamento Escola
Tècnica Superior
d'Enginyeria, Departament
d'Informàtica / Universidad
de Valencia, Avenida de la
Universidad, s/n, 46100,
Burjassot, València, España
joigpana@uv.es

Silvia Rueda

Departamento Escola
Tècnica Superior
d'Enginyeria, Departament
d'Informàtica / Universidad
de Valencia, Avenida de la
Universidad, s/n, 46100,
Burjassot, València, España
silvia.rueda@uv.es

Oscar Pastor

Centro de Investigación en
Métodos de Producción de
Software / Universidad
Politécnica de València,
Camino de Vera s/n, 46022,
Valencia, España
opastor@pros.upv.es

Recibido: 05.05.2020 | Aceptado: 31.07.2020

Palabras Clave

Modelo BPMN
Interfaces gráficas de usuario
Estudio empírico
Atractividad
Correctitud
Usabilidad

Resumen

Contexto: Diseñar las Interfaces Gráficas de usuario (IGUs) a partir de la Notación de Modelos de Proceso de Negocio (BPMN) no es trivial, hay gran variedad de alternativas gráficas para una misma primitiva BPMN. **Objetivo:** Una de las contribuciones es la elaboración de recomendaciones para seleccionar las alternativas gráficas a partir de primitivas BPMN que optimizan la usabilidad. Otra es la ejecución de una familia de experimentos para analizar las preferencias de los sujetos en el diseño de IGUs a partir de BPMN y ver si estas preferencias coinciden con nuestras recomendaciones. **Método:** Las recomendaciones de usabilidad fueron extraídas de propuestas anteriores. La familia de experimentos consiste en 2 réplicas de 43 y 47 sujetos donde se estudia la Atractividad (porcentaje de sujetos que prefieren cada alternativa de diseño) y la Correctitud (porcentaje de sujetos cuya preferencia coincide con la recomendación que proponemos). **Resultados:** La Atractividad muestra que 10 alternativas (de 14) de diseño de IGUs superaron el 50% de los votos de los sujetos. La Correctitud muestra que 8 alternativas de diseño (de 14) coinciden con nuestras alternativas recomendadas. **Conclusiones:** Las alternativas con mejor Atractividad y Correctitud son report, datagrid, combobox, timer, hyperlink, Textbox, checkbox y buttons.

Keywords

BPMN model
Graphical User Interfaces
Empirical study
Attractiveness
Correctness
Usability

Abstract

Context: Design Graphical User Interfaces (GUIs) from the Business Process Models Notation (BPMN) is not a trivial process, since there is a wide variety of alternatives to represent the same primitive BPMN in the GUI. **Objective:** One of the contributions is the elaboration of recommendations to select the graphical alternatives from BPMN primitives that optimize usability. Other contribution is the conduction of a family of experiments to analyze the preferences of the subjects in the design of GUIs from BPMN models and check whether such preferences agree with our recommendations. **Method:** Usability recommendations were extracted from previous proposals. Regarding the family of experiments, it consists in 2 replications of 43 and 47 subjects where we study Attractiveness (percentage of subjects who prefer each GUI design alternative) and Correctness (percentage of subjects whose preference agrees with the usability recommendation that we propose). **Results:** Attractiveness shows that 10 design alternatives (out of 14) of GUIs exceeded 50% of the votes of the subjects. Correctness shows that 8 design alternatives (out of 14) agree with our recommended

alternatives. **Conclusions:** The alternatives with the best Attractiveness and Correctness are report, datagrid, combobox, timer, hyperlink, textbox, check box and buttons.

1. Introducción

En la actualidad los diseñadores deben seleccionar muy bien qué componentes gráficos de las Interfaces Gráficas de Usuario (IGUs) se visualizarán en el software. Las IGUs son críticas, ya que un software con IGUs no usables no será aceptado por los usuarios. El cómo diseñar IGUs a partir de los requisitos es uno de los retos a los que se enfrentan los desarrolladores de software. Hay modelos de requisitos, como por ejemplo los modelos de procesos de negocio (BPMN, 2013), que se centran en capturar requisitos no funcionales y funcionales sin abordar elementos de interacción. El cómo se transforman estos requisitos en IGU depende de los desarrolladores, que aplican reglas subjetivas en base a su experiencia particular. En la literatura, existen enfoques que proponen la generación de IGUs a partir de BPMN a través de reglas de transformación que permitan sistematizar el proceso lo máximo posible, como: (Díaz, Panach, Rueda, & Pastor, 2018b), (Díaz, Panach, Rueda, & Pastor, 2018a), (Brambilla, Fraternali, & Vaca, 2011), (Sousa, Mendonça, & Vanderdonckt, 2007), (Han, Zhao, & Yang, 2016), (Yongchareon et al., 2018).

La contribución de este artículo es doble. En primer lugar se han definido recomendaciones para indicar cuál de todas las alternativas de diseño para una misma primitiva BPMN optimiza la usabilidad. Esta lista de recomendaciones se ha elaborado juntando resultados extraídos de trabajos previos de la literatura. En segundo lugar, se ha lanzado una familia de experimentos con dos réplicas de 43 sujetos (41 hombres y 2 mujeres) y 47 sujetos (40 hombres y 7 mujeres) para ver las preferencias de estos sujetos a la hora de diseñar las IGUs a partir de modelos BPMN. Todos los sujetos son estudiantes del Grado de Ingeniería Informática de la Universidad de Valencia (España) del Curso 2018/2019 y 2019/2020 respectivamente. En este experimento analizamos dos variables: Atractividad (Azuma, 2008) de las alternativas de diseño de IGUs (porcentaje de sujetos que han votado por cada alternativa de diseño de IGU), y la Correctitud (IEEE, 2010) de recomendaciones de usabilidad (porcentaje de las preferencias de los sujetos que coincide con nuestras recomendaciones que optimizan la usabilidad). Ambas variables se miden a través de un cuestionario desarrollado por los propios autores. El cuestionario contiene 14 preguntas, donde cada una muestra una primitiva BPMN y pregunta qué alternativa de diseño de IGU prefiere el usuario para presentarla visualmente.

Los resultados de la familia de experimentos se puede resumir en los siguientes puntos: (1) con respecto a la variable Atractividad de las alternativas de diseño de las IGUs, tenemos que 10 alternativas de diseño (de un total de 14) superaron el 50% de los votos de los sujetos, lo que muestra un alto valor de consenso. (2) con respecto a la Correctitud de las recomendaciones de usabilidad, tenemos que 8 alternativas de diseño (de un total de 14 recomendaciones evaluadas) elegidas por los sujetos superan el 50% de los votos y coinciden con nuestras alternativas que recomendamos. Esto demuestra que el 58% de las recomendaciones de usabilidad evaluadas coinciden con las preferencias de los usuarios. El estudio de si las recomendaciones de usabilidad mejoran en términos de eficiencia, efectividad y satisfacción (Iso/iec, 2005) está fuera de objetivo de este artículo, que se centra sólo en las preferencias del usuario. El artículo también explica posibles justificaciones por las que hay alternativas de diseño con mayor dispersión y recomendaciones de usabilidad que no hay correlación con las preferencias de los sujetos.

Los elementos gráficos de las IGUs con mayor votos por los sujetos fueron las siguientes Report, Datagrid, Timer, Hyperlink, Textbox, Checkbox y Buttons. Estos elementos gráficos son considerados más usables para cada contexto determinado. Esto permite al diseñador contar con una ayuda para tener una referencia de qué elementos gráficos usar. Esta investigación permite el avance científico para el desarrollo de IGUs usables a partir de un modelo BPMN.

Este documento está estructurado de la siguiente manera: La sección 2 muestra trabajos relacionados. La sección 3 describe recomendaciones de usabilidad que proponemos a partir de trabajos de la literatura. La sección 4 define el diseño del experimento. La sección 5 discute las amenazas a la validez. La sección 6 muestra los resultados. La sección 7 discute la interpretación de resultados. Finalmente, la sección 8 presenta algunas conclusiones relevantes y trabajos futuros.

2. Estado del Arte

En esta sección revisamos trabajos relacionados que hayan hecho experimentos para analizar las variables Atractividad y Correctitud, además de trabajos relacionados con recomendaciones de usabilidad. Realizamos una Revisión de Literatura Dirigida (RLT), una revisión de literatura no sistemática, profunda e informativa destinada a mantener sólo las referencias significativas maximizando la rigurosidad y minimizando el sesgo de selección. Para este propósito, la pregunta semántica para esta investigación se traduce en la

siguiente consulta sintáctica utilizada como una cadena de búsqueda en la biblioteca digital de Scopus (ver <https://www.scopus.com/home.uri>):

"GUI" AND ("BPMN" OR "attractiveness" OR "correctness" OR "usability")

Las referencias resultantes de esta búsqueda se clasificaron en tres categorías Atractividad, Correctitud, y Recomendaciones de usabilidad, que se analizan en mayor detalle en las siguientes sub secciones.

2.1 Atractividad

En este grupo tenemos el trabajo de Dahl et al. (Dahl & Svendsen, 2011), que describe un estudio preliminar de factores que influyen en la usabilidad de interfaces para entornos inteligentes. Se probaron 3 prototipos de IGUs en un laboratorio de usabilidad, se analizaron los comentarios de los sujetos, y se definió la variable atractividad para las IGUs. Los resultados muestran que los sujetos consideran que diseñar interfaces atractivas es importante para la aceptación del usuario. Otro trabajo es el de Schrepp et al. (Schrepp, Held, & Laugwitz, 2006) que presenta un experimento empírico acerca de la satisfacción de los sujetos. Se estudia qué aspectos de interfaces influyen en la usabilidad y la atractividad. Los sujetos evaluaron 3 interfaces diferentes a través del cuestionario AttrakDiff2 (Hassenzahl, Burmester, & Koller, 2003). Además, clasificaron las IGUs por preferencia personal. Los resultados muestran que cuanto más atractiva es una interfaz, mayor preferencia tienen los sujetos por ella. Otro trabajo es de Hartmann et al. (Hartmann, Sutcliffe, & De Angeli, 2007), que presenta un marco teórico para evaluar la atractividad de sitios web basado en la teoría adaptativa de toma de decisiones (Payne, Payne, Bettman, & Johnson, 1993). Este marco se desarrolló en un cuestionario y se evaluaron 3 sitios web (web HCI group, web Design Division and web DSchool de la Universidad de Stanford) que compartían el mismo tema, pero que diferían en el diseño estético. El resultado muestra que el sitio web preferido fue DSSchool, teniendo un estilo más interactivo, alta estética y usabilidad. Otro trabajo es el de Lee et al. (Lee & Yurchisin, 2011) que presenta un estudio que examina los antecedentes de la intención de los consumidores de compra en línea, incluyendo la atractividad del sitio web, identificación del consumidor y la confiabilidad del sitio web. Utilizaron un modelo de ecuación estructural derivado de la identidad social y la investigación previa sobre el comportamiento del consumidor en línea. El resultado indicó que la atractividad de un sitio web influye directa y positivamente en el grado que se identifican con ese sitio web.

Como conclusión de este grupo de trabajos existentes, podemos afirmar que la atractividad se ha evaluado a través

de los comentarios de sujetos (Dahl & Svendsen, 2011) (Hartmann et al., 2007), de cuestionarios (Schrepp et al., 2006), y de antecedentes de los sujetos en el uso de un sitio web (Lee & Yurchisin, 2011). Una de las principales limitaciones del trabajo de (Schrepp et al., 2006) es que evaluaron 3 interfaces de usuario diferentes del mismo tema y completaron un cuestionario definido que mide la calidad del producto de software y una parte de atractividad de esas interfaces de usuario, pero no una evaluación directa a cada elemento gráfico. Otro inconveniente detectado está en el trabajo de (Hartmann et al., 2007), donde se amplían conceptos de alta estética en IGUs que no tienen una relación clara con la atractividad del sitio web. En nuestra propuesta, nos basamos en un cuestionario para medir la atractividad, de forma que sea fácilmente entendible por los sujetos y que puedan responderlo directamente.

2.2 Correctitud

El trabajo de Sreedhar et al. (Sreedhar, 2016) describe que existen muchas pautas de diseños y métricas para la evaluación de sitios web y aplicaciones, aunque la mayoría carece de un marco de especificaciones. El resultado muestra un conjunto de técnicas de optimización para mejorar la correctitud del sitio web. Otro trabajo es el de Stavnycha et al. (Stavnycha, Yin, & Römer, 2015) que presenta una investigación que tiene por objetivo analizar los efectos de 9 prácticas de desarrollo preseleccionadas sobre la correctitud de aspecto de calidad de versiones de software. Se recopilaron datos de desarrolladores de software de todo el mundo, utilizando una encuesta en línea. Analizaron 1006 respuestas válidas con ayuda de métodos estadísticos. El resultado mostró que 4 de 9 prácticas de desarrollo tienen efectos significativos sobre la correctitud del software.

Como conclusión de este grupo de trabajos existentes, podemos decir que el trabajo (Sreedhar, 2016) se basa en un conjunto de técnicas, y su limitación es que no muestra de una manera clara y práctica estas técnicas para mejorar la correctitud. El trabajo de (Stavnycha et al., 2015) propone que la evaluación de la correctitud del software se basa en un cuestionario acerca del proceso de desarrollo de software, pero no en las IGUs, como en nuestro enfoque.

2.3 Usabilidad

El trabajo de Molich et al. (Molich, Jeffries, & Dumas, 2007) presenta un estudio que compara los comentarios de usabilidad escritos por diferentes autores describiendo problemas de usabilidad similares del sitio web del Hotel Pennsylvania en New York. Estos comentarios fueron proporcionados por 17 equipos de profesionales que evaluaron de forma independiente. El resultado muestra que hay 84 comentarios estudiados, 14 comentarios abordan 6

problemas de usabilidad que contenían recomendaciones que fueron útiles y utilizables, 14 recomendaciones no fueron útiles, y 16 no fueron utilizables. El resultado de este documento sugiere características para recomendaciones útiles, es decir recomendaciones para resolver problemas de usabilidad que conducen a cambios que mejoran eficientemente la usabilidad de un producto. Otro trabajo es el de Kous et al. (Kous, Pušnik, Heričko, & Polančič, 2020) que presenta una evaluación de usabilidad de cómo los diferentes tipos de usuarios finales (N=25) (estudiantes, población en general, personas mayores e investigadores) responden al sitio web de una biblioteca para medir su eficacia, eficiencia y satisfacción. Realizaron pruebas formales de usabilidad, análisis de registros y cuestionarios. Los resultados muestran que diferentes grupos de usuarios finales alcanzan diferentes niveles de efectividad y eficiencia, mientras no hay una diferencia significativa entre los grupos a nivel de satisfacción. Además, indica que los usuarios no alcanzaron el nivel para un sitio web utilizable, por lo que presentan recomendaciones para mejorar la utilidad de un sitio web especialmente para usuarios no experimentados. Otro trabajo es el de Osada et al. (Osada, Muke, Piwowarczyk, Telec, & Trawiński, 2020) que presenta un análisis comparativo y examina la usabilidad de métodos y patrones seleccionados de entrada de datos en sistemas y sitios web. Desarrollaron una herramienta experimental en forma de una aplicación web que contenía un amplio repertorio de 36 patrones de entrada de datos para realizar pruebas de usabilidad. Los resultados recopilados proporcionaron una base para comparar las medidas de eficacia y eficiencia de los patrones probados y los métodos de entrada de datos. Además recomienda los mejores patrones y métodos de entrada de datos en un contexto de uso definido. Otro trabajo es el de Park et al. (Park, Jeong, & Kim, 2018) que presenta una evaluación de usabilidad de interfaces de menú en la pantalla táctil de relojes inteligentes a través de dos experimentos. En el primer experimento proporcionan 40 elementos en vista de cuadrícula y estilos de diseño de vista de lista. El resultado muestra que la eficiencia fue mayor para el diseño de vista de lista, sin embargo la satisfacción fue mayor para el diseño de la vista cuadrícula. El segundo experimento muestra un menú eficiente con elementos jerárquicos que pueden agruparse en categorías superior e inferior. El resultado muestra valores positivos en eficiencia y satisfacción al proporcionar el menú por categorización.

Como conclusión de este grupo de trabajos existentes podemos decir que se analizaron recomendaciones de usabilidad mediante comentarios de profesionales (Molich et al., 2007), análisis de métodos y patrones (Osada et al., 2020), evaluación de usabilidad (Kous et al., 2020) (Park et al., 2018). Una de las principales limitaciones de estos trabajos es que

proponen varias recomendaciones influyentes de usabilidad pero relativamente muestran poca orientación de cómo deben usarse estas interfaces de usuario en la práctica. En el trabajo de (Kous et al., 2020) evaluaron solo a una pequeña muestra de 25 usuarios. Una de nuestras contribuciones con el presente trabajo es la definición de recomendaciones de usabilidad para diseñar IGUs a partir de modelos BPMN de una forma muy clara y práctica, y la existencia de estas recomendaciones no es común en el área de BPMN ya que no es un enfoque muy común.

Analizando los trabajos relacionados podemos destacar que no existen trabajos previos que hayan tratado cómo optimizar la usabilidad de las interfaces a partir de modelos BPMN. El presente trabajo pretende cubrir precisamente esta carencia. Además de proponer las recomendaciones de usabilidad, el presente trabajo también aporta una validación de dichas recomendaciones. Son pocos los trabajos previos que han experimentado sobre transformaciones a partir de modelos BPMN.

3. Recomendaciones de Usabilidad

Esta sección muestra una de las contribuciones de este artículo: una lista de recomendaciones de usabilidad a la hora de diseñar IGUs a partir de modelos BPMN. Estas recomendaciones son una recopilación de recomendaciones ya existentes en la literatura: Galitz's usability guidelines (Galitz, 2007), Johnson's GUI Bloopers (Johnson, 2008), Microsoft MSDN (Microsoft, 2018), SAP Fiori (Fiori, 2018), U.S. Web Design (U. S. W. Design, 2018), Material Design (M. Design, 2018) y Van Welie (Welie & Traetteberg, 2000). Las recomendaciones se han definido para cada una de las reglas de generación de IGUs a partir de primitivas de modelado de BPMN. Estas reglas están fuera de la contribución del presente artículo y se pueden consultar en (Díaz et al., 2018b) (Díaz et al., 2018a). Las reglas tienen como entrada elementos BPMN para representar comportamiento y elementos del diagrama de clases de UML para representar la estructura de datos. Cada regla de generación puede tener diversas alternativas, con las recomendaciones de usabilidad pretendemos que de todas estas alternativas, el desarrollador tenga a su disposición una recomendación. Por ejemplo, la Regla 2, si varias tareas de tipo usuario están en el mismo carril y tienen dependencia entre ellas, las alternativas de diseño son wizard, group box y tabcontrol. Y el contexto de uso recomendado para wizard sería cuando cada tarea de tipo de usuario tiene más de 8 atributos del diagrama de clases.

La Tabla 1 muestra las primitivas del diagrama de clases y las primitivas del modelo BPMN a partir de las cuales se pueden generar los componentes gráficos, la alternativa de diseño que genera, y el contexto de uso en el cual se recomienda la

alternativa para optimizar la usabilidad. Junto a cada contexto de uso recomendado aparece la fuente de la literatura utilizada para su definición. A partir de los atributos del diagrama de clases UML se pueden generar los siguientes componentes gráficos: text box, combobox, listbox, checkbox, y radiobutton. Además, a partir de las primitivas del modelo BPMN se pueden generar los siguientes componentes gráficos: wizard, tabcontrol, groupbox, report, datagrid, hyperlink, timer, y messagebox. Por ejemplo, a partir de varias Tareas de Tipo Usuario en BPMN se puede generar un wizard, y esta alternativa de generación se recomienda cuando hay varias Tareas Tipo Usuario y existen más de ocho atributos del diagrama de clases

Tabla 1: Recomendaciones de Usabilidad

Primitivas diagrama de clases UML / BPMN	Alternativa de diseño	Contexto de uso recomendado para la alternativa
Atributos del diagrama de clases	Textbox	(1) El atributo puede ser una cadena o un número corto en una sola línea (Fiori, 2018). (2) El atributo puede ser una contraseña, URL, número de teléfono o dirección de correo electrónico (Fiori, 2018). (3) El atributo puede ser cualquier valor que el sistema no pueda predecir, como la respuesta de un usuario. Existe una gran variabilidad en las respuestas de los usuarios (U. S. W. Design, 2018).
	Combobox	(4) El atributo puede ser una larga lista de elementos (mínimo 13, máximo 200 entradas) (Fiori, 2018). (5) Los valores en la lista de opciones son información secundaria y no es necesario que se muestren inmediatamente (Fiori, 2018).
	Listbox	(6) El atributo puede tomar valores de una o más opciones (Microsoft, 2018). (7) Una lista lo suficientemente grande como para mostrar de 3 a 8 elementos cuando se abre (Microsoft, 2018). (8) Para elegir una o varias opciones de una lista.
	Checkbox	(9) Los atributos pueden tomar valores de un grupo o una lista de opciones que se pueden seleccionar independientemente uno del otro o ninguno (Fiori, 2018). (10) Los atributos pueden tomar valores de una lista de opciones que se muestran inmediatamente sin interacción del usuario (también en casos de solo lectura) (Fiori, 2018).

	Radiobutton	(11) El atributo puede tomar un valor de un conjunto de opciones mutuamente excluyentes (U. S. W. Design, 2018). (12) El usuario debe elegir rápidamente entre al menos dos opciones claramente diferentes (Fiori, 2018).
BPMN: Tarea tipo usuario o Compuerta paralela	Wizard	(13) El usuario necesita un asistente de navegación secuencial (Microsoft, 2018). (14) Cada tarea de tipo usuario tiene más de 8 atributos del diagrama de clases. (15) Hay 3 o más pasos con dependencia en el modelo BPMN (Welie & Traetteberg, 2000).
	Tabcontrol	(16) El usuario debe mostrar la información en páginas clasificadas por separado (M. Design, 2018).
	Groupbox	(17) El usuario necesita agrupar elementos gráficos dentro de un formulario (Microsoft, 2018). (18) La tarea de tipo usuario tiene menos de 8 atributos.
BPMN: Tarea tipo servicio	Report	(19) La tarea de tipo servicio muestra información lista para imprimir.
	Datagrid	(20) El usuario necesita mostrar datos con múltiples propiedades en filas y columnas (cuadrícula de datos) (Galitz, 2007). Una ventaja de la cuadrícula de datos es que los usuarios pueden ordenar y filtrar los datos (Johnson, 2008).
BPMN: Compuerta decisión exclusiva	Radiobutton	(21) Cada rama de la compuerta que conecta a una tarea tiene una opción, donde el usuario solo puede elegir una (U. S. W. Design, 2018).
BPMN: Evento tipo simple	Hyperlink	(22) El usuario necesita una navegación entre interfaces u otro evento.
BPMN: Evento tipo temporizador	Timer	(23) El usuario necesita un evento que requiera un cronómetro.
	Messagebox	(24) El usuario debe ser interrumpido durante la ejecución de una acción (Fiori, 2018). (25) El sistema necesita mostrar mensajes de error, mensajes de advertencia, mensajes de éxito, mensajes de confirmación o mensajes de información. (26) El usuario debe tomar una decisión.

4. Experimento

Esta sección describe el diseño de la familia de experimentos para estudiar las preferencias de los sujetos y validar las recomendaciones de usabilidad extraídas de la literatura a la hora de diseñar IGUs a partir de un modelo BPMN. El experimento se realiza desde una perspectiva de

investigación interesada en las preferencias de los sujetos al diseñar IGUs a partir de modelos BPMN. Esta sección está estructurada de la siguiente manera: Primero muestra los sujetos del experimento, segundo las preguntas de investigación y se definen sus hipótesis, tercero el cuestionario donde validaremos las variables, cuarto el procedimiento del experimento, y finalmente los resultados.

4.1 Sujetos

La familia de experimentos consta de 2 réplicas. En ambas réplicas los sujetos son estudiantes del Grado de Ingeniería Informática de la Universidad de Valencia. Todos estudiantes de la asignatura de Entornos de Usuario, donde se ven conceptos de Interacción Persona Ordenador (y en especial de usabilidad). La muestra está formada por 90 sujetos entre las dos réplicas: la primera con 43 sujetos (41 hombres y 2 mujeres, 29 en el rango de 17-20 años de edad, 9 en el rango de 21-25 años, 1 en el rango de 26-30 años, 1 en el rango de 31-35 años, y 3 por encima de los 35 años, $M = 26.20$, $SD = 4.71$), y la segunda con 47 sujetos (40 hombres y 7 mujeres, 30 en el rango de 17-20 años de edad, 13 en el rango de 21-25 años, 1 en el rango de 26-30 años, 2 en el rango de 31-35 años, y 1 por encima de los 35 años, $M = 21.31$, $SD = 4.55$). Los sujetos fueron reclutados en dos años académicos: Curso 2018/2019 (primera réplica) y el Curso 2019/2020 (segunda réplica). La Tabla 2 muestra los datos demográficos de los sujetos. Se puede apreciar que los sujetos no tienen mucho conocimiento en BPMN pero sí un medio y alto conocimiento en el diseño de IGUs. Esta amenaza respecto a la carencia en conocimiento en BPMN se minimizó a través de tutoriales en BPMN que los sujetos debían cursar antes de participar en el experimento. Además, una parte de los sujetos muestra que tiene conocimientos en diagramas de clases UML. Los sujetos participaron voluntariamente en el experimento.

Tabla 2: Conocimiento de Diagrama de clases, modelo BPMN e IGUs en Curso 2018/2019 y 2019/2020

Curso 2018/2019				
Conocimiento	Ninguno	Bajo	Medio	Alto
Clases UML	23%	40%	35%	2%
Modelo BPMN	53%	35%	10%	2%
IGUs	7%	42%	42%	9%
Curso 2019/2020				
Conocimiento	Ninguno	Bajo	Medio	Alto
Clases UML	36%	38%	26%	0%
Modelo BPMN	57%	36%	7%	0%
IGUs	6%	41%	46%	7%

4.2 Preguntas de investigación, hipótesis, variables respuesta, y métricas

El experimento se basa en las siguientes preguntas de investigación:

RQ1: ¿Cuál es el porcentaje de atractividad para cada una de las alternativas de diseño de IGUs en un contexto de uso a partir de un modelo BPMN? La atractividad se considera como el grado en que el producto de software es atractivo para el usuario conforme a la ISO 25010 (Azuma, 2008). Para abordar esta propiedad, queremos estudiar la hipótesis **H1**: "Por cada contexto de uso hay una alternativa de diseño que destaca más del 50% del resto".

RQ2: ¿Cuál es la correctitud de la alternativa de diseño recomendada con respecto a las preferencias de los sujetos? La correctitud se define como el grado en que un sistema o componente está libre de fallos en su especificación, diseño e implementación, de acuerdo con IEEE (IEEE, 2010). Para abordar esta propiedad, queremos analizar la hipótesis **H2**: "Las preferencias de los sujetos en las alternativas de diseño de las IGUs son las mismas que nuestras alternativas recomendadas".

Las variables respuesta son los efectos estudiados en el experimento. En nuestra familia de experimentos tenemos dos variables respuesta:

RQ1: requiere una variable para medir la atractividad de las alternativas de diseño de las IGUs para cada contexto de uso. Hemos desarrollado un cuestionario con 14 preguntas, donde cada pregunta muestra una primitiva de un modelo BPMN y se muestran 4 alternativas de diseño representadas gráficamente como en una IGU. Esta variable se mide como el porcentaje de sujetos que han votado por cada alternativa de diseño de IGU (ver fórmula 1). PA significa Porcentaje de Atractividad. NSVA significa Número de sujetos que votan por cada alternativa. Si el porcentaje de cada alternativa de diseño de la IGU es cercano a 100%, significa que esa alternativa de diseño es la preferida entre todas.

$$PA = (NSVA / \text{Número de sujetos}) \times 100\% \quad (1)$$

RQ2: requiere una variable para medir la correctitud de las recomendaciones de usabilidad, la métrica en este caso es la comparativa de las preferencias de los sujetos a la hora de elegir la alternativa de diseño versus nuestras alternativas recomendadas para optimizar la usabilidad. Utilizamos el mismo cuestionario que para RQ1, teniendo en cuenta que en cada pregunta hay una alternativa que es la que recomendamos para ese contexto. Medimos esta variable como el porcentaje de las preferencias de los sujetos en las alternativas de diseño de las IGUs que concuerdan con nuestras recomendaciones de usabilidad (ver fórmula 2). Si el resultado es cercano a 100%, significa que las preferencias de los sujetos de la alternativa de diseño de IGU son las mismas que nuestras recomendaciones. PCRU significa Porcentaje de Correctitud de las Recomendaciones de Usabilidad. NSNR

significa Número de Sujetos que han elegido Nuestras Recomendaciones.

$$PCRU = (NSNR / \text{Número de sujetos}) \times 100\% \quad (2)$$

4.3 Cuestionario

El cuestionario fue desarrollado por los investigadores para evaluar la Atractividad de las alternativas del diseño y la Correctitud de las recomendaciones de usabilidad. Este cuestionario cuenta con 14 preguntas, cada pregunta está extraída de un contexto de uso de la Tabla 1 (marcados en negrita). En la pregunta se desea conocer para ese contexto de uso, cuál es la alternativa que prefiere el sujeto. Las preguntas se muestran de forma textual y las alternativas (4 por pregunta) se muestran textual y gráficamente (a modo de ejemplo de cómo sería esa alternativa en una IGU). Cabe destacar que de los 26 contextos de uso definidos en la Tabla 1, se han evaluado 14 (que salen en negrita en la Tabla 1). Esta decisión se ha tomado para evitar la amenaza de cansancio en los sujetos y poder ejecutar el experimento en 40 minutos máximo. La elección de 14 contextos de uso utilizados en el cuestionario ha dependido de las siguientes reglas: cada primitiva BPMN aparecerá al menos en una pregunta; cada elemento gráfico de la IGU debe aparecer en al menos una pregunta; los contextos de uso utilizados deben ser lo más disjuntos posibles. A continuación veremos las preguntas, contexto de uso del que deriva, y las alternativas de diseño.

Pregunta 1 (c0ntexto de uso 14): Si cada Tarea de tipo usuario cuenta con más de ocho atributos.

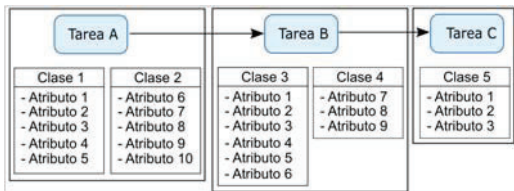


Figura 1: Pregunta 1 y alternativas de diseño

Pregunta 2 (c0ntexto de uso 19): Si la Tarea B de tipo servicio debe mostrar información para imprimir.

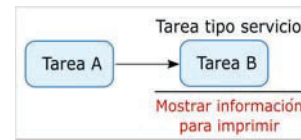


Figura 2: Pregunta 2 y alternativas de diseño.

Pregunta 3 (c0ntexto de uso 15): Si las Tareas A, B y C de tipo usuario existe dependencia en más de 3 tareas, y se deben usar para una navegación de forma secuencial.



Figura 3: Pregunta 3 y alternativas de diseño.

Pregunta 4 (c0ntexto de uso 20): Si la Tarea B de tipo servicio debe mostrar información en tablas organizadas.

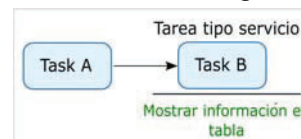


Figura 4: Pregunta 4 y alternativas de diseño.

Pregunta 5 (c0ntexto de uso 4): Si el valor del Atributo 2 de la tarea de tipo usuario debe ser seleccionado de una lista de más de 13 opciones.

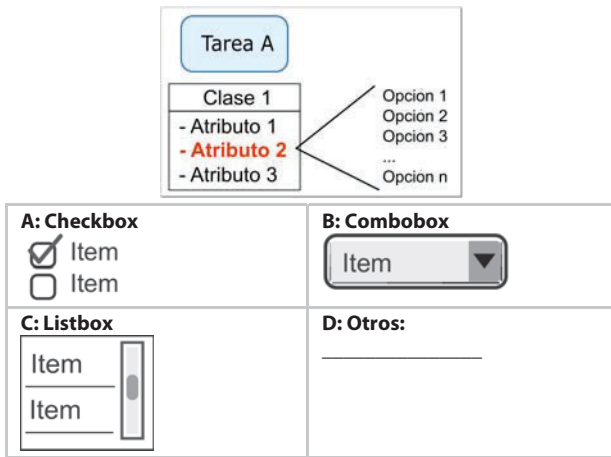


Figura 5: Pregunta 5 y alternativas de diseño.

Pregunta 6 (Contexto de uso 7): El valor del Atributo 2 de la Tarea A de tipo usuario debe mostrarse en una lista de valores visibles en un componente gráfico, y contar con menos o igual a 8 opciones.

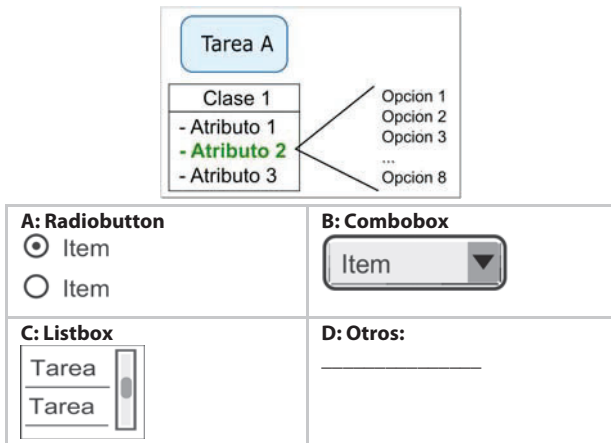


Figura 6: Pregunta 6 y alternativas de diseño.

Pregunta 7 (Contexto de uso 23): Si el evento tiene un Tiempo máximo de ejecución, ese tiempo máximo lo representamos en un:

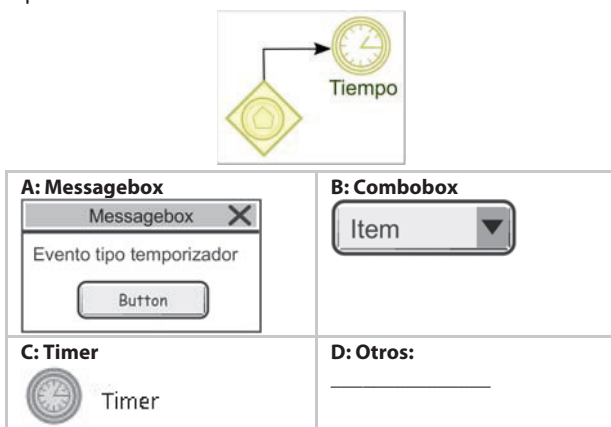


Figura 7: Pregunta 7 y alternativas de diseño

Pregunta 8 (Contexto de uso 16): Si las Tareas A y B de tipo usuario tienen dependencia y deben mostrarse en páginas clasificadas por separado.

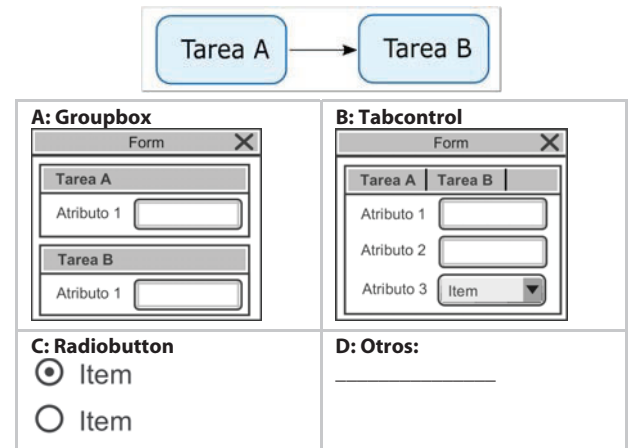


Figura 8: Pregunta 8 y alternativas de diseño

Pregunta 9 (Contexto de uso 18): Si cada Tarea de tipo usuario cuenta con menos de 8 atributos y todas las tareas deben mostrarse en un solo formulario.

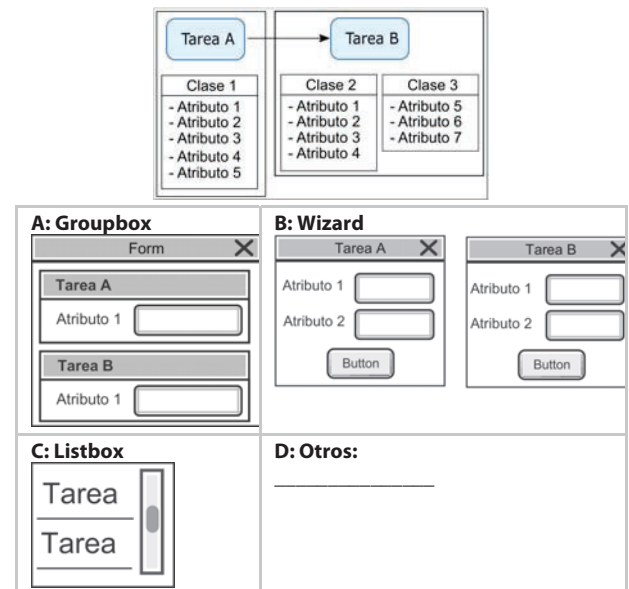


Figura 9: Pregunta 9 y alternativas de diseño.

Pregunta 10 (Contexto de uso 22): Cuando el usuario necesita una navegación entre interfaces con el evento tipo simple.

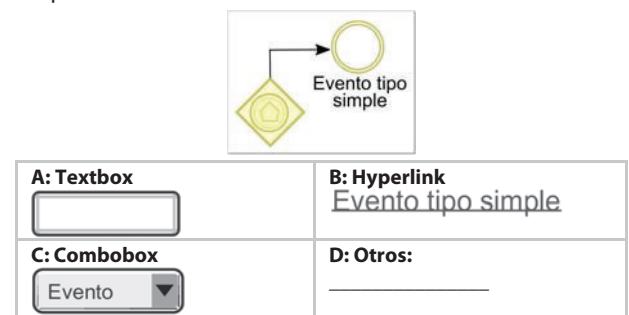


Figura 10: Pregunta 10 y alternativas de diseño.

Pregunta 11 (Contexto de uso 1): El Atributo 2 de la Tarea tipo usuario necesita introducir una cadena o un número corto en una sola línea.

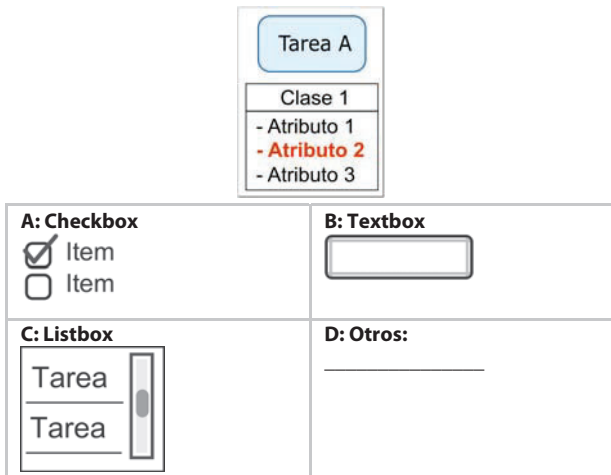


Figura 11: Pregunta 11 y alternativas de diseño.

Pregunta 12 (cOntexto de uso 3): Si el Atributo 3 de la Tarea tipo usuario puede ser cualquier valor que el sistema no pueda predecir.

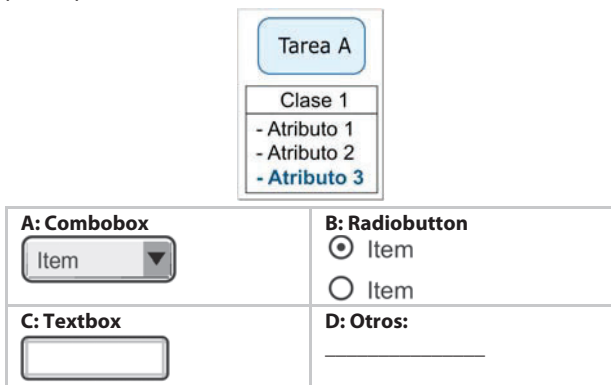


Figura 12: Pregunta 12 y alternativas de diseño.

Pregunta 13 (cOntexto de uso 9): Los valores de los atributos de la Tarea tipo usuario, se deben mostrar en una lista de opciones. De esta lista de opciones, el usuario debe poder elegir cuantos quiera (desde ninguno a todos ellos).

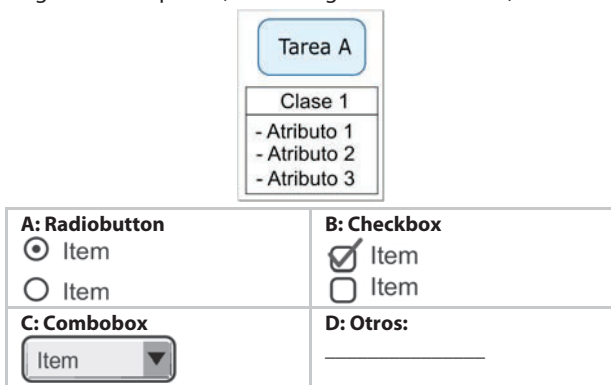


Figura 13: Pregunta 13 y alternativas de diseño.

Pregunta 14 (cOntexto de uso 21): Si la compuerta de decisión exclusiva tiene dos opciones "Si" y "No", sólo se puede elegir una opción.

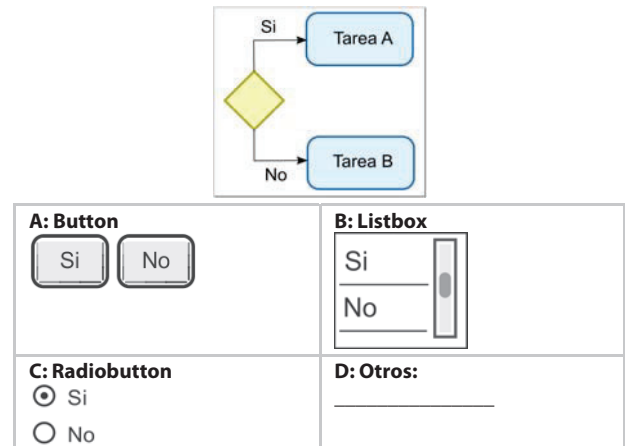


Figura 14: Pregunta 14 y alternativas de diseño.

4.4 Procedimiento

El procedimiento de cada réplica se aplicó en una sesión de 40 minutos en cada una de ellas (Curso 2018/2019 y 2019/2020). A continuación, se detallan los pasos del procedimiento en orden cronológico:

Introducción del Cuestionario, antes de asistir al experimento, los sujetos deben revisar un tutorial de BPMN y del diagrama de clases. El tutorial fue entregado una semana antes del experimento para que tuvieran tiempo de trabajarlo. El día del experimento, un experimentador dedicó 10 minutos de una breve introducción a BPMN y al diagrama de clases.

Completar un cuestionario demográfico, los sujetos completaron un cuestionario demográfico con los siguientes campos (correo electrónico, nombre, apellido, sexo, edad, conocimiento de BPMN, diagrama de clase UML e IGU) antes de ejecutar el experimento para identificar sus antecedentes. Cada sujeto firmó un formulario de consentimiento.

Evaluación mediante un test previo, los sujetos fueron evaluados con un test sobre el modelo BPMN y el diagrama de clases antes de participar en el experimento. Este test asegura que los sujetos reclutados tienen los conocimientos mínimos como para poder entender los diagramas. El test consiste en 10 preguntas, cada pregunta tenía 4 alternativas y solo una posible respuesta correcta. La respuesta correcta se calculó como un punto, por lo que los posibles puntos estaban entre cero (sin respuestas correctas) y 10 (todas las respuestas correctas). Consideramos que los sujetos que obtuvieron más de 5 puntos fueron capaces de participar en el experimento. La Tabla 3 muestra los resultados, todos los sujetos obtuvieron una calificación alta en el test, por lo tanto todos los sujetos participan en el experimento.

Tabla 3: Resultados del test

	Curso 2018/2019				Curso 2019/2020				
Puntaje	7	8	9	10	6	7	8	9	10
Sujetos	2	6	16	19	3	2	4	20	18

Rellenar cuestionario, cada sujeto rellena el cuestionario de las 14 preguntas en línea, además no se impuso ninguna restricción. Se instruyó a los sujetos para que eligieran la opción que ellos prefieran, ya que a partir de este cuestionario se extraen los valores para atractividad y correctitud. El proceso de rellenar el cuestionario duró 20 minutos.

5. Amenazas a la Validez

Esta sección discute las amenazas a la validez de la familia de experimentos. Describimos las amenazas de acuerdo a la clasificación de Wohlin (Wohlin et al., 2012). Para cada grupo de amenazas, distinguimos entre las amenazas que no hemos podido abordar, las amenazas cuyo efecto hemos minimizado y las amenazas resueltas. Clasificamos las amenazas en 4 grupos: (1) **Validez de la conclusión**: este tipo de amenazas se refiere a la capacidad de sacar la conclusión correcta sobre las relaciones del tratamiento y el resultado. El experimento puede sufrir amenazas de este tipo: *Sujetos de heterogeneidad aleatoria*, lo que significa que siempre hay heterogeneidad en un grupo de estudio. Existe el riesgo de que la variación debida a diferencias individuales sea mayor que debido al tratamiento. Para minimizar esta amenaza, reclutamos sujetos con perfiles similares (todos son estudiantes del mismo grado), además utilizamos un cuestionario demográfico que aseguró que todos tenían un perfil similar. Otra amenaza que aparece es la *Pesca*, que significa que los experimentadores están buscando resultados específicos. Para minimizar esta amenaza, los sujetos fueron evaluados sin conocer el objetivo del experimento. Otra amenaza que aparece es la *Fiabilidad de la medida*, que significa que la validez de un experimento depende en gran medida de la fiabilidad de las medidas. Para minimizar esta amenaza, las métricas fueron calculadas por el experimentador, la Atractividad y la Correctitud se calculó en una hoja de cálculo, lo que reduce los posibles errores. (2) **Validez interna**: este tipo de amenazas analiza las influencias que pueden afectar el factor de causalidad. El experimento puede sufrir las siguientes amenazas: *Historial*, que significa que pueden surgir diferencias cuando los tratamientos se aplican en diferentes momentos. Para minimizar esta amenaza, realizamos el experimento en una sesión para cada réplica. Otra amenaza que puede aparecer es la *Instrumentación*, que significa que los artefactos utilizados en el experimento podrían afectar en los resultados. Para minimizar esta amenaza hemos seguido a (Kitchenham & Pfleeger, 2002) para validar nuestro cuestionario a través de 4 tipos de validaciones: (i) validación aparente: El cuestionario

fue validado por 3 investigadores que revisaron las preguntas del cuestionario, estos investigadores participaron en el desarrollo del artículo. Podemos asegurar que el cuestionario cumple con los objetivos de la investigación de capturar las preferencias de los sujetos con respecto a las IGUs. (ii) validez de contenido: cada pregunta es un contexto de recomendaciones de usabilidad (ver Tabla 1) que se basan en una recopilación de recomendaciones ya existentes en la literatura adaptadas a nuestro enfoque. (iii) validez de criterio: Es una posible amenaza porque no hemos podido comparar con otros instrumentos existentes, porque existe poca literatura y cuestionarios relacionado a nuestro enfoque. (iv) Construcción válida: El cuestionario fue utilizado en dos réplicas diferentes. Consideramos que se realizarán más evaluaciones para recopilar datos y verificar si producen resultados similares. (3) **Validez del constructo**: este tipo de amenaza se refiere a generalizar el resultado del experimento al concepto o teoría detrás del experimento. El experimento puede sufrir las siguientes amenazas: *Aprehensión de Evaluación* que significa que algunas personas temen ser evaluadas. Para minimizar esta amenaza, comunicamos a los sujetos que los contenidos del experimento estaban relacionados con los objetivos de la asignatura. (4) **Validez externa**: este tipo de amenaza se refiere a condiciones que limitan nuestra capacidad de generalizar los resultados de nuestros experimentos a la práctica industrial. El experimento puede sufrir las siguientes amenazas: *Interacción de selección y tratamiento*, que significa el efecto de tener una población de sujetos no representativa de la población que queremos generalizar. No podemos garantizar que los resultados del experimento sean válidos para sujetos con un perfil diferente a los reclutados en este experimento. Otra amenaza que puede aparecer es la *Interacción del entorno y el tratamiento*, que significa el efecto de no tener el entorno experimental o material representativo de la práctica industrial. El contexto de nuestro experimento es solo un entorno académico y los resultados solo pueden generalizarse en dicho contexto.

6. Resultados

Esta sección muestra los resultados de la familia de experimentos. En el análisis se agregó los datos de las 2 réplicas ya que el diseño de ambas réplicas es idéntico.

6.1 Atractividad de las alternativas de diseño de las IGUs

La atractividad de las alternativas de diseño de IGUs se mide como el porcentaje de sujetos que han votado por cada una de ellas (el mayor porcentaje significa mayor atractividad). La Figura 15(a) muestra un gráfico circular de la atractividad de las alternativas de diseño de IGUs para la **Pregunta 1**. La alternativa para diseñar un Tabcontrol a partir de la primitiva

BPMN de 3 Tareas tipo usuario, tiene un valor de atractividad de 48%, en Groupbox de 29%, en wizard de 22%, y en otros elementos gráficos de 1%. Por tanto, no hay ninguna alternativa de diseño que obtenga mayor del 50% de votos de los sujetos, lo que indica que no hay consenso para una alternativa preferida. La Figura 15(b) muestra resultados de la **Pregunta 2**, la alternativa para diseñar un Report a partir de una Tarea tipo servicio que debe mostrar información para imprimir, tiene un valor de atractividad de 81%, esto significa que la mayoría de sujetos consideran que esta alternativa de diseño es la preferida. La Figura 15(c) muestra los resultados de la **Pregunta 3**, la alternativa para diseñar un Tabcontrol a partir de tres Tareas de tipo usuario que tienen dependencia tiene un valor de atractividad del 51%, por lo que se considera la preferida con alto consenso.

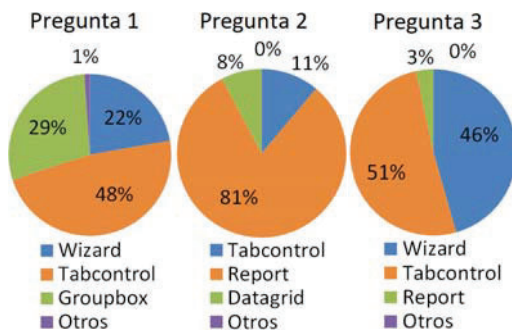


Figura 15: Atractividad de las alternativas de diseño de IGUs por los sujetos en: (a) Pregunta 1. (b) Pregunta 2. (c) Pregunta 3.

La Figura 16(a) muestra los resultados de la **Pregunta 4**, la alternativa para diseñar un Datagrid a partir de una Tarea de tipo servicio que debe mostrar información en tablas, tiene un valor de atractividad de 87%, por lo tanto se considerada como la preferida por los sujetos. La Figura 16(b) muestra los resultados de la **Pregunta 5**, la alternativa para diseñar un Combobox a partir de una Tarea tipo usuario complementado con un atributo de una clase que debe ser seleccionado de una lista de más de 13 opciones, tiene un valor de atractividad de 52%, por lo que se considera como la preferida. La Figura 16(c) muestra los resultados de la **Pregunta 6**, la alternativa para diseñar un Combobox a partir de una Tarea tipo usuario complementado con un atributo de una clase que debe mostrar una lista de valores visibles con menos de 8 opciones, tiene un valor de atractividad de 49%, en Listbox de 38%, y en Radiobutton de 13%. Por tanto, no hay ninguna alternativa con un alto valor y además hay bastante dispersión entre Radiobutton, Combobox y Listbox.

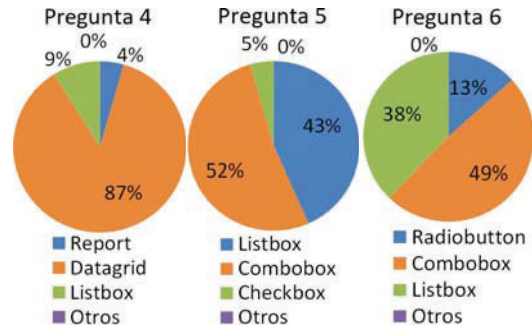


Figura 16: Atractividad de las alternativas de diseño de IGUs por los sujetos en: (a) Pregunta 4. (b) Pregunta 5. (c) Pregunta 6.

La Figura 17(a) muestra los resultados de la **Pregunta 7**. La alternativa para diseñar un Timer a partir de un Evento temporizador, tiene un valor de atractividad del 78%, podemos decir que esta alternativa de diseño es la preferida. La Figura 17(b) muestra los resultados de la **Pregunta 8**, la alternativa para diseñar un Groupbox a partir de 2 Tareas tipo usuario, tiene un valor de atractividad de 49%, en Tabcontrol de 34%, y en Radiobutton de 17%. Por tanto, podemos decir que no hay ninguna alternativa de diseño preferida y hay bastante dispersión entre las 3 alternativas. La Figura 17(c) muestra los resultados de la **Pregunta 9**, la alternativa para diseñar un Wizard a partir de 2 Tareas tipo de usuario complementado con clases, tiene un valor de atractividad de 42%, en Groupbox de 38%, y en Listbox de 20%. Por tanto, podemos concluir que no hay ninguna alternativa de diseño preferida y los votos están muy dispersos.

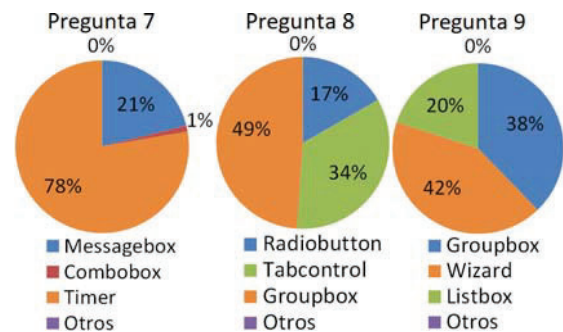


Figura 17: Atractividad de las alternativas de diseño de IGUs por los sujetos en: (a) Pregunta 7. (b) Pregunta 8. (c) Pregunta 9.

La Figura 18(a) muestra los resultados de la **Pregunta 10**, la alternativa para diseñar un Hyperlink a partir de un Evento tipo simple, tiene un valor de atractividad de 69%, podemos decir que esta alternativa de diseño es la preferida. La Figura 18(b) muestra los resultados de la **Pregunta 11**, la alternativa para diseñar un Textbox a partir de una Tarea tipo usuario complementado con un atributo del diagrama de clases que debe ingresar un texto corto, tiene un valor de atractividad de 87%, por lo que podemos decir que es la preferida. La Figura 18(c) muestra los resultados de la **Pregunta 12**, la alternativa para diseñar un Textbox a partir de una Tarea tipo de usuario complementado con un atributo del diagrama de clases que

puede ser un valor abierto, tiene un valor de atraktividad de 78%, lo que indica que es la alternativa preferida.

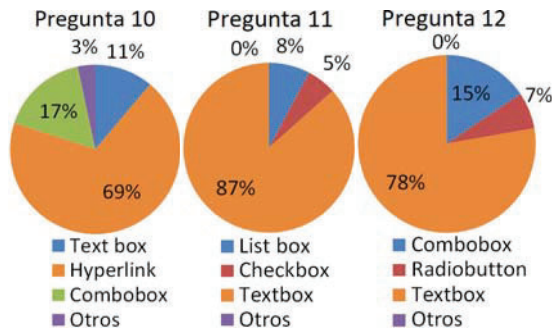


Figura 18: Atractividad de las alternativas de diseño de IGUs por los sujetos en: (a) Pregunta 10. (b) Pregunta 11. (c) Pregunta 12.

La Figura 19(a) muestra un gráfico de los resultados de la **Pregunta 13**. La alternativa para diseñar un Checkbox a partir de una Tarea tipo usuario complementado con un atributo del diagrama de clases que debe mostrar una lista de opciones en donde se puede elegir cuantos quiera o ninguna opción, tiene un valor de atraktividad de 67%, lo que indica que es la preferida. La Figura 19(b) muestra los resultados de la **Pregunta 14**. La alternativa para diseñar Buttons a partir de una Compuerta de decisión exclusiva con dos opciones que solo se tiene que elegir una opción, tiene un valor de atraktividad de 61%, lo que indica que es la preferida.

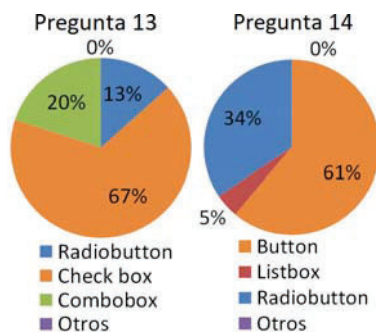


Figura 19: Atractividad de las alternativas de diseño de IGUs por los sujetos en: (a) Pregunta 13. (b) Pregunta 14.

La Tabla 4 muestra un resumen de las alternativas de diseño de IGUs que obtuvieron más votos por los sujetos (valores superiores al 50%), lo cual indica que son las alternativas preferidas.

Tabla 4: Alternativas de diseño con alta Atractividad

Primitivas BPMN/ diagrama de clases	Alternativa de diseño	% Atractividad
Tarea tipo servicio que debe mostrar información para imprimir.	Report	81%
Tareas tipo usuario que tienen dependencia.	Tabcontrol	51%
Tarea tipo servicio que debe mostrar información en tablas.	Datagrid	87%
Tarea tipo usuario complementado con un atributo que debe ser	Combobox	52%

seleccionado de una lista de más de trece opciones.		
Evento temporizador.	Timer	78%
Evento tipo simple.	Hyperlink	69%
Tarea tipo usuario complementado con un atributo que se puede ingresar un texto corto.	Textbox	87%
Tarea tipo usuario complementado con un atributo que es un valor abierto.	Textbox	78%
Tarea tipo usuario complementado con atributos que deben mostrar una lista de opciones, y que se puede elegir cuantos quiera o ninguna.	Checkbox	67%
Compuerta decisión exclusiva con dos opciones y solo se debe elegir una.	Buttons	61%

Concluimos que 10 alternativas de diseño de IGUs (de un total de 14 analizadas) superaron más del 50% indicando que **H1** se satisface mayormente.

6.2 Correctitud de las recomendaciones de usabilidad

La Figura 20 muestra un diagrama de frecuencia con el porcentaje de correctitud de las recomendaciones de usabilidad. Las preguntas 2, 4, 5, 7, 10, 11, 12 y 13 reciben los mejores valores de correctitud (superaron el 50% de los votos de los sujetos), mientras que las preguntas 1, 3, 6, 8, 9 y 14 reciben los peores valores de correctitud, lo que sugiere que la mayoría de los sujetos tienden a preferir IGUs alternativas a nuestras recomendaciones de usabilidad. En estos casos donde no se ha obtenido una correspondencia entre nuestra propuesta y las preferencias de los sujetos, las alternativas de diseño más votados han sido:

- Pregunta 1 y 3: un tabcontrol (48% y 51% votos de sujetos respectivamente) a partir de Tareas tipo usuario. La recomendación del wizard es más baja (22% y 46% de correctitud respectivamente).
- Pregunta 6: un combobox (49% de votos) a partir de una Tarea tipo usuario complementado con un atributo del diagrama de clases que debe mostrar las opciones visibles y tener menos de 8 opciones. La recomendación del listbox es más baja (38% correctitud).
- Pregunta 8: un groupbox (49% votos) a partir de 2 Tareas tipo usuario que tienen dependencia y deben mostrarse en el mismo formulario. La recomendación del tabcontrol es baja (34% de correctitud).
- Pregunta 9: un wizard (42% de votos) a partir de Tareas tipo usuario complementadas con atributos y que deben mostrarse en un formulario. La recomendación del groupbox es más baja (38% de correctitud).
- Pregunta 14: button (61% de votos) a partir de Compuerta de decisión exclusiva que tiene 2

opciones y solo se tiene que elegir una. La recomendación del radio button es más baja (34% correctitud).

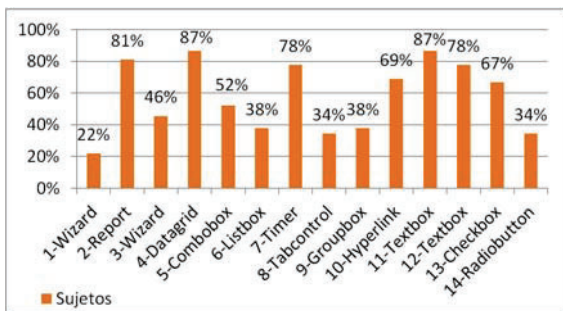


Figura 20: Porcentaje de Correctitud de las recomendaciones de usabilidad

La Tabla 5 muestra un resumen de las alternativas de diseño que tuvieron un alto porcentaje de correctitud respecto a las recomendaciones de usabilidad. Debemos tener en cuenta que aquellas alternativas de diseño con alta atractividad coinciden con nuestras alternativas recomendadas. Concluimos que 8 alternativas de diseño que recomendamos (de un total de 14 que evaluamos) superaron el 50% de los votos de los sujetos, indicando que **H2** se satisface mayormente.

Tabla 5: IGUs con alta correctitud de recomendaciones de usabilidad

Primitivas BPMN/ diagrama de clases	Alternativa recomendada	% Correctitud
Tarea tipo servicio debe mostrar información para imprimir.	Report	81%
Tarea tipo servicio que debe mostrar información en tablas.	Datagrid	87%
Tarea tipo usuario complementado con un atributo debe ser seleccionado de una lista de más de trece opciones.	Combobox	52%
Evento temporizador.	Timer	78%
Evento tipo simple.	Hyperlink	69%
Tarea tipo usuario complementado con un atributo que permite un texto.	Textbox	87%
Tarea tipo usuario complementado con un atributo de un valor abierto.	Textbox	78%
Tarea tipo usuario complementado con atributos mostrados en lista de opciones, puede elegir cuantos quiera.	Checkbox	67%

7. Discusión

Los resultados para Atractividad de las alternativas de diseño de IGUs muestran que 10 preguntas cuentan con más del 50% de votos de los sujetos. Las alternativas que presentan los porcentajes más altos de atractividad se puede deber al hecho de que representan componentes gráficos de la IGUs que son ampliamente conocidos por los sujetos. A continuación, analizamos las alternativas con mejor atractividad. La pregunta 2, tiene un alto valor de atractividad Report (81%) a

partir de Tarea tipo servicio y atributos, se puede deber a que los sujetos prefieren mostrar información a través de informes antes que se muestren en tablas (datagrid) o un conjunto de tabs. La Pregunta 3, tiene un valor de atractividad del 51% para generar un Tabcontrol a partir de 03 Tareas tipo usuario que se deben usar para una navegación secuencial. Esto se puede deber que los sujetos prefieren visualizar tareas tipo usuario en páginas clasificadas por separado a diferencia de un wizard que tiene un conjunto de formularios. La Pregunta 4 tiene un valor de atractividad del 87% al generar un Datagrid a partir de Tarea tipo servicio, se puede deber a que los sujetos prefieren ver información a través de tablas organizadas, en vez de usar listbox o un report. La Pregunta 5, tiene un valor de atractividad del 52% al generar un Combobox a partir de una Tarea tipo usuario y un atributo. Una posible justificación para esta referencia puede ser que los sujetos cuando quieren tener una lista de más de 13 ítems prefieren mostrar todas las opciones en una única lista cuyo contenido solo se ve al desplegarla. La Pregunta 7 tiene un valor de atractividad del 78% para un Timer generado a partir de un Evento tipo temporizador. Esto se puede deber a que los sujetos prefieren usar un control no visible para manejar un tiempo máximo, en vez de messagebox que es más visual y puede provocar más interrupciones en la interacción. La Pregunta 10, tiene un valor de atractividad del 69% para Hyperlink generado a partir de un Evento tipo simple. Este resultado se puede deber a que los sujetos prefieren usar este elemento para navegar a otro lugar del sistema en vez de combobox o textbox. La Pregunta 11 y Pregunta 12 tienen un alto valor de atractividad al generar un Textbox (87% y 78% respectivamente). En estas preguntas se genera a partir de una Tarea tipo usuario y un atributo. Esto se puede deber a que los sujetos consideran que el Textbox es común para ingresar datos al sistema, en vez de listbox o combobox, que contienen opciones ya definidas. La Pregunta 13 tiene un valor de atractividad del 67% para generar un Checkbox a partir de Tarea tipo usuario y un atributo, se puede deber que los sujetos consideran ver atributos como lista de opciones en checkbox donde puedan elegir cuantos quieran o ninguno. La Pregunta 14 tiene un valor de atractividad del 61% para generar Buttons a partir de una compuerta decisión exclusiva. Esto se puede deber, que los sujetos prefieren ver opciones de Si/No de la compuerta exclusiva mediante botones, en vez de mostrarse en radiobutton (34%) o listbox (5%), que normalmente tienen función distinta a la pregunta Sí/No.

Se debe tener en cuenta que la mayoría de sujetos tienen un conocimiento medio y alto en el diseño de IGUs, esto puede haber favorecido el hecho de altos valores de atractividad para una misma alternativa en la mayoría de las preguntas. Las alternativas de diseño que no superaron el 50% de votos de los sujetos, se encuentran en las siguientes preguntas: 1, 6, 8,

y 9. La Pregunta 1, fueron muy dispersos los valores de atractividad de los sujetos cuando hay 3 tareas tipo usuario y atributos: tabcontrol (48%), groupbox (29%), wizard (22%), y otras IGUs (1%). Esto se puede deber, que las 3 alternativas son similares para agrupar tareas y contener componentes, además que se puede aplicar navegación. La Pregunta 6, los valores de atractividad son dispersos cuando se genera una IGU a partir de una Tarea tipo usuario y un atributo que debe mostrar una lista menos de 8 opciones visibles: combobox (49%), listbox (38%), y radiobutton (13%). Esto se puede deber a que estas IGUs muestran una lista de opciones donde el usuario solo puede elegir una opción. La Pregunta 8 tiene valores para atractividad muy dispersos cuando se genera una IGU a partir 2 tareas tipo usuario: groupbox (49%), tabcontrol (34%), y radiobutton (17%). En groupbox y tabcontrol, su elección tiene sentido, ambos pueden usar grupos y conjunto de pestañas para cada tarea. Esta forma de interactuar parece más natural que usar un radiobutton que muestra opciones y donde eliges una. La Pregunta 9 presenta valores similares para la atractividad cuando se genera una IGU a partir de 2 Tareas tipo usuario complementada con menos de 8 atributos en cada tarea: wizard (42%) y groupbox (38%). Esto se puede deber a que estas alternativas agrupan componentes y son similares a diferencia del listbox (20%) que contiene ítems. Los resultados de correctitud de las recomendaciones de usabilidad muestran que la mayoría de sujetos han preferido nuestras recomendaciones. De las 14 recomendaciones de alternativas evaluadas, 8 han sido preferidas por los sujetos. A continuación pensamos posibles justificaciones para las alternativas recomendadas que no han sido elegidas por los sujetos (correctitud inferior al 50%). La Pregunta 1 muestra que Tabcontrol tiene 48% de votos de los sujetos, se genera a partir de 3 tareas tipo usuario complementado con más de 8 atributos en cada tarea, mientras que nuestra recomendación es un Wizard con 22% de votos. Esto puede ser porque los sujetos prefieren ver cada tarea con atributos en cada página clasificada, y además una parte de los sujetos prefieren ver cada tarea en Groupbox con 29% de los votos. La pregunta 3 muestra que Tabcontrol tiene 51% de votos de los sujetos, que se genera a partir de 3 Tareas tipos usuario que tienen dependencia y debe usarse para navegación secuencial. Mientras que nuestra recomendación es Wizard con 46% de votos. Podemos ver que los votos son similares, aunque la mayoría de sujetos prefieren ver las tareas en tabs. La pregunta 6 muestra que Combobox tiene 49% de votos de los sujetos que se genera a partir de una Tarea tipo usuario y un atributo que debe mostrar una lista de valores visibles y contar con menos e igual a 8 opciones. Mientras que nuestra recomendación es un Listbox con 38% de votos. Esto puede ser porque los usuarios consideran que combobox es más común y usable en este contexto, teniendo en cuenta que esta IGU oculta las de más opciones a diferencia del Listbox que

recomendamos. La pregunta 8 muestra que Groupbox tiene 49% de votos de los sujetos que se genera a partir de 2 Tareas tipo usuario que tienen dependencia y deben mostrar en páginas clasificadas, mientras que nuestra recomendación es un Tabcontrol con 34% de votos. Esto puede ser porque los sujetos prefieren ver tareas agrupadas. La pregunta 9 muestra que un Wizard tiene 42% de votos de los sujetos, que se genera a partir de 2 tareas tipo usuario con menos de 8 atributos y deben mostrarse en un formulario, mientras que nuestra recomendación es un Groupbox con 38% de votos. Puede ser que un grupo de sujetos se confundieran con el wizard ya que este elemento mostraría cada tarea en un formulario de navegación secuencial. La pregunta 14 muestra que los botones tienen 61% de votos de los sujetos, que se genera a partir de una compuerta decisión exclusiva, mientras que nuestra recomendación es Radiobutton con 34% de votos. Esto puede ser porque los botones son ampliamente conocidos por los sujetos y permite realizar una acción a diferencia del radiobutton que permite elegir una opción.

8. Conclusiones y Trabajos Futuros

El artículo propone una lista de recomendaciones para optimizar la usabilidad en el proceso de generar IGUs a partir de modelos BPMN. Estas recomendaciones se han evaluado en una familia de experimentos con 2 réplicas de 43 y 47 sujetos respectivamente. Las variables respuesta analizadas fueron Atractividad de las alternativas de diseño de IGUs (porcentaje de sujetos que han votado por cada alternativa de diseño de IGU) y Correctitud (porcentaje de las preferencias de los sujetos que concuerdan con nuestras recomendaciones de usabilidad). Ambas variables se midieron a través de un cuestionario, cada pregunta mostraba 4 alternativas de elementos gráficos para una primitiva BPMN (solo 1 de las 4 alternativas era la recomendada). Los resultados de estas medidas son: (1) Atractividad indica que hay consenso entre los sujetos en 10 alternativas (de 14): (1) De una Tarea tipo servicio que debe mostrar información para imprimir se genera un Report. (2) De 3 Tareas tipo usuario que existe dependencia y se deben usar para una navegación secuencial se genera un Wizard. (3) De una Tarea tipo servicio debe mostrar información en tablas organizadas se genera un Datagrid. (4) De una Tarea tipo usuario y un atributo que debe ser seleccionado de una lista de más de 13 opciones se genera un Combobox. (5) De un Evento tipo temporizador se genera un Timer. (6) De un Evento tipo simple se genera un Hyperlink. (7) De una Tarea tipo usuario y un atributo que debe ingresar una cadena o números se genera un Textbox. (8) De una Tarea tipo usuario y un atributo que puede ser un valor que el sistema no pueda predecir se genera un Textbox. (9) De una Tarea tipo usuario y atributos que deben ser mostradas en lista de opciones, donde el usuario debe poder elegir cuantos quiera o ninguno se genera un Checkbox. (10) De una

compuerta decisión exclusiva se genera Radiobutton. (II) Correctitud de recomendaciones de usabilidad indica que de 14 recomendaciones evaluadas, 8 coinciden con preferencias del usuario, para no repetir referenciamos los números que se resumieron en los resultados de Atractividad: 1, 3, 4, 5, 6, 7, 8 y 10, que son las mismas primitivas BPMN y alternativas de diseño para una alta correctitud.

Los elementos gráficos de las IGUs con altos votos de sujetos son consideradas más usables: Report 81%, Datagrid 87%, Timer 78%, Hyperlink 69%, Textbox 87% y 78%, Checkbox 67% y Buttons 61%. Estos componentes gráficos de las IGUs concuerdan con nuestras recomendaciones. Los sujetos que usen estas recomendaciones de usabilidad en la práctica podrán elaborar sistemas interactivos más usable sin tener un conocimiento avanzando o experiencia en usabilidad.

El experimento sufre de las siguientes limitaciones: la mayoría de sujetos no tienen un conocimiento amplio de modelos BPMN y diagrama de clases UML, pero nosotros intentamos cubrir esta limitación con un tutorial. Además, la mayoría de

sujetos tienen un alto conocimiento de diseño de IGUs. Los resultados deben ser interpretados dentro del contexto en el que se ha ejecutado el experimento: (1) los sujetos son estudiantes del Grado de Ingeniería Informática. (2) el cuestionario tiene 14 preguntas e involucra a 4 alternativas de componentes gráficos por pregunta.

Como trabajo futuro, planeamos replicar este experimento cambiando algunos elementos del diseño: Primero considerar más preguntas con un mayor número de componentes gráficos (más alternativas). Segundo, reclutar sujetos que estén en la industria del diseño de software.

AGRADECIMIENTOS

El primer autor agradece al Ministerio de Educación del Perú, PRONABEC – Beca Presidente de la República. Este proyecto además está con el soporte del Ministerio de Ciencia e Innovación Español a través del proyecto DataME (ref: TIN2016-80811-P).

Referencias

- Azuma, M. (2008). *Iso/iec cd 25010: Software engineering–software product quality requirements and evaluation (square)–software and quality in use models*. ISO/IEC JTC1/SC7, CANADA.
- BPMN. (2013). *Business Process Modeling Notation*. 2020, from <http://www.bpmn.org>
- Brambilla, M., Fraternali, P., & Vaca, C. (2011). *BPMN and design patterns for engineering social BPM solutions*. Paper presented at the International Conference on Business Process Management.
- Dahl, Y., & Svendsen, R.-M. (2011). *End-user composition interfaces for smart environments: A preliminary study of usability factors*. Paper presented at the International Conference of Design, User Experience, and Usability.
- Design, M. (2018). *Design*. from <https://material.io/design/>
- Design, U. S. W. (2018). *UI Components*. from <https://designsystem.digital.gov/components/>
- Díaz, E., Panach, J. I., Rueda, S., & Pastor, O. (2018a). *Generación de Interfaces de Usuario a partir de Modelos BPMN con Estereotipos*. Paper presented at the Jornada de la Sociedad de Ingeniería de Software y Tecnologías de Desarrollo de Software (SISTEDES).
- Díaz, E., Panach, J. I., Rueda, S., & Pastor, O. (2018b, 29-31 May 2018). *Towards a method to generate GUI prototypes from BPMN*. Paper presented at the 2018 12th International Conference on Research Challenges in Information Science (RCIS).
- Fiori, S. (2018). *Fiori Design Guidelines*. from <https://experience.sap.com/fiori-design-web/>
- Galitz, W. O. (2007). *The essential guide to user interface design: an introduction to GUI design principles and techniques*: John Wiley & Sons.
- Han, L., Zhao, W., & Yang, J. (2016). *An approach towards user interface derivation from business process model*. *Communications in Computer and Information Science*, 602, 19-28. doi: 10.1007/978-981-10-1019-4_2
- Hartmann, J., Sutcliffe, A., & De Angeli, A. (2007). *Investigating attractiveness in web user interfaces*. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
- Hassenzahl, M., Burmester, M., & Koller, F. (2003). *AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität Mensch & computer 2003* (pp. 187-196): Springer.
- IEEE. (2010). *Systems and software engineering -- Vocabulary*.
- Iso/iec. (2005). *ISO/IEC 25000 - Software engineering - Software product Quality Requirements and Evaluation (SQuARE) - Guide to SQuARE*. doi: citeulike-article-id:10949862
- Johnson, J. (2008). *Common user interface design don'ts and Dos*: Morgan Kaufmann.
- Kitchenham, B., & Pfleeger, S. L. (2002). *Principles of survey research part 4: questionnaire evaluation*. *ACM SIGSOFT Software Engineering Notes*, 27(3), 20-23.

- Kous, K., Pušnik, M., Heričko, M., & Polančič, G. (2020). Usability evaluation of a library website with different end user groups. *Journal of Librarianship and Information Science*, 52(1), 75-90.
- Lee, Z. C., & Yurchisin, J. (2011). The impact of website attractiveness, consumer-website identification, and website trustworthiness on purchase intention. *International Journal of Electronic Customer Relationship Management*, 5(3-4), 272-287.
- Microsoft, M. (2018). Common UI Controls and Text Guidelines
- Molich, R., Jeffries, R., & Dumas, J. S. (2007). Making usability recommendations useful and usable. *Journal of Usability Studies*, 2(4), 162-179.
- Osada, K., Muke, P. Z., Piwowarczyk, M., Telec, Z., & Trawiński, B. (2020). Comparative Usability Analysis of Selected Data Entry Methods for Web Systems. *Cybernetics and Systems*, 1-22.
- Park, K., Jeong, M., & Kim, K. (2018). Usability evaluation of menu interfaces for smartwatches. *Journal of Computer Information Systems*, 1-10.
- Payne, J. W., Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*: Cambridge university press.
- Schrepp, M., Held, T., & Laugwitz, B. (2006). The influence of hedonic quality on the attractiveness of user interfaces of business management software. *Interacting with Computers*, 18(5), 1055-1069.
- Sousa, K., Mendonça, H., & Vanderdonck, J. (2007). User Interface Derivation from Business Processes: A Model-Driven Approach for Organizational Engineering. Paper presented at the TAMODIA, Toulouse (France).
- Sreedhar, G. (2016). Optimizing Website Content to Improve Correctness of the Website Design Problem Solving and Uncertainty Modeling through Optimization and Soft Computing Applications (pp. 329-338): IGI Global.
- Stavnycha, M., Yin, H., & Römer, T. (2015). A large-scale survey on the effects of selected development practices on software correctness. Paper presented at the Proceedings of the 2015 International Conference on Software and System Process.
- Welie, M. v., & Traetteberg, H. (2000). Interaction Patterns in User Interfaces. Paper presented at the 7th. Pattern Languages of Programs Conference, Illinois, USA.
- Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). *Experimentation in Software Engineering: An Introduction*: Springer.
- Yongchareon, S., Liu, C., Zhao, X., Yu, J., Ngamakeur, K., & Xu, J. (2018). Deriving user interface flow models for artifact-centric business processes. *Computers in Industry*, 96, 66-85. doi: <https://doi.org/10.1016/j.compind.2017.11.001>