



## INSTRUMENTO PARA MEDIR LA FUNCIONALIDAD Y USABILIDAD DE UN SOFTWARE COMERCIAL

### INSTRUMENT FOR MEASURING THE FUNCTIONALITY AND USABILITY OF A COMMERCIAL SOFTWARE



Carlos Enrique Castillo Diestra



Jorge Luis Gutierrez Gutierrez<sup>2</sup>

#### RESUMEN

El objetivo del presente estudio fue desarrollar un instrumento para medir la funcionalidad y usabilidad de un software comercial. Para ello se revisó bibliografía sobre funcionalidad y usabilidad del software y construcción de instrumentos de medición. Luego se diseñó el instrumento denominado IMFYU, siguiendo las recomendaciones dadas por la norma ISO/IEC 9126. El IMFYU evalúa la funcionalidad y usabilidad. Para probar la validez de contenido, el instrumento fue sometido a revisión por un panel de expertos en Ingeniería de Computación y Sistemas, con más de 5 años en el manejo del software, obteniéndose una validez de contenido con un coeficiente de correlación intra-clase de 0,9266 para un intervalo de confianza del 95%. También se sometió el instrumento a una prueba piloto para determinar su confiabilidad, obteniéndose un coeficiente de confiabilidad alfa de Cronbach promedio de 0,8533. Las pruebas de validación y de confiabilidad se hicieron con el paquete estadístico SPSS 9.0.

Palabras clave: Tema: Calidad, Especialidad: Software, Objetivo: Desarrollar

Método: Inductivo

#### ABSTRACT

The objective of the present work was to develop an instrument to measure the business software functionality and usability. For it, bibliography on functionality and usability and on construction of measuring instruments were reviewed. Then, the instrument, which was named IMFYU, was developed following the recommendations given by ISO/IEC 9126 norm. The IMFYU measure the functionality and usability. The instrument was put under revision by a panel of experts in Computing and Systems Engineering, with more than 5 years in the handling of software, obtaining itself a content validity with an intra-class correlation coefficient of 0.9266 for a confidence interval of 95%. Also the instrument was put under a pilot test to determine its reliability, obtaining itself a Cronbach alfa coefficient average of 0.8533. The validation and reliability analysis were done with SPSS 9.0 statistical package.

Keywords: Software, Quality.

#### I. INTRODUCCIÓN

El software de computadora juega un rol importante en diversos ámbitos de nuestra vida como lo son: la toma de decisiones comerciales, la investigación científica, la resolución de problemas de ingeniería, etc. De aquí que el software se encuentra inmerso en sistemas de todo tipo por lo que es importante y necesario que en cada una de estas aplicaciones se posea un software de calidad; es decir, que tenga concordancia con los requisitos funcionales y de rendimiento explícitamente

establecidos, los estándares de desarrollo explícitamente documentados y las características implícitas que se esperan de todo software desarrollado profesionalmente (7). Es así que empleamos las medidas para valorar la calidad de los productos de ingeniería o de los sistemas que construimos.

Lo primero que se debe considerar al abordar la calidad de software que éste constituye un producto con características muy peculiares (8) razón por la cual este tema ha constituido durante años un debate entre los miembros de la comunidad de software. Asimismo en el transcurrir del tiempo se han emprendido varios trabajos de investigación abordando este interesante tema; pero quienes desarrollamos, compramos o estudiamos algún tipo de software, nos encontramos con un problema: la carencia de instrumentos que brinden a los desarrolladores e incluso a los usuarios criterios de evaluación de la calidad de algún producto software. Por ello presentamos nuestra investigación hacia el desarrollo de un instrumento que permita determinar si un determinado software es de calidad, considerando los atributos de funcionalidad y usabilidad.

Esta investigación busca medir la funcionalidad y usabilidad de un software comercial de manera confiable y válida. Entendiéndose a la **funcionalidad como el grado en que el software satisface las necesidades indicadas por el usuario; y usabilidad al grado en que el software es fácil de usar por el usuario.**

El diseño del instrumento para medir la funcionalidad y usabilidad de un software comercial se basa en los principales modelos de calidad del software existentes, por una parte, el concepto de calidad es demasiado complejo como para poder ser evaluada o medida mediante una única medida. La norma UNE-EN ISO 8402 (9) define el aseguramiento de la calidad (AC) como "el conjunto de acciones planificadas y sistemáticas implantadas dentro del sistema de calidad, y demostrables para proporcionar la confianza adecuada de que una entidad cumplirá los requisitos para la calidad". Estos requisitos deben reflejar totalmente las necesidades y expectativas del usuario. En definitiva, el AC debe recoger el conjunto de acciones necesarias para asegurar que el cliente recibe el producto software acordado y, por tanto, queda satisfecho. Cada vez más se está asociando el concepto de calidad a la satisfacción del usuario y este hecho supone una mayor complejidad y ambigüedad en la obtención de mediciones reales y fiables de la calidad del software. En cualquier caso, se reconoce que la idea de calidad varía de un cliente a otro, de un proyecto a otro. De hecho, uno de los campos en los que más se ha trabajado es en la utilización de modelos de evaluación de calidad de software que tratan de aportar un medio para definir y descomponer el concepto de calidad de software con características más sencillas de evaluar y medir. Así podemos encontrar modelos de evaluación como el FCM (Factores/Criterios/Métricas) de McCall, métodos para crear modelos propios en cada proyecto como COQUAMO o el de Gilb e, incluso, estándares que abordan esta cuestión como ISO 9126 o IEEE 1061.

Por otra parte, se han propuesto multitud de mediciones, principalmente de productos, que proclaman su utilidad para evaluar la calidad del software, o alguna de sus facetas. En este caso, el trabajo se centró en analizar cada una de las propuestas (tanto teóricamente como en la práctica), cómo contribuyen o se pueden usar para el aseguramiento de la calidad en un proyecto y qué característica miden realmente y si pueden encajar dentro de los anteriormente mencionados modelos de evaluación de calidad de software y dentro de las actividades de aseguramiento de calidad (10).

Los modelos de calidad por lo general se definen con un criterio jerárquico, o sea la calidad se produce como consecuencia de la evaluación de un conjunto de indicadores o métricas en diferentes etapas:

En el nivel más alto de jerarquía se encuentran los factores de calidad definidos a partir de la visión del usuario del software, y conocidos también como atributos de calidad externos. Cada uno de los factores se descompone en un conjunto de criterios de calidad, o sea aquellos atributos que cuando están presentes contribuyen a obtener un software de la calidad. Se trata de una visión de la calidad técnica, desde el punto de vista del producto software y se les denomina también atributos de calidad internos.

Finalmente para cada uno de los criterios de calidad se definen un conjunto de métricas o medidas cuantitativas de ciertas características del producto que indican el grado en que dicho producto posee un determinado atributo de calidad.

De esta manera, a través de un modelo de calidad se concretan los aspectos relacionados con ella de tal manera que se puede definir, medir y planificar. Además el empleo de un modelo de calidad permite comprender las relaciones que existen entre diferentes características de un producto software.

## II. MATERIAL Y MÉTODOS

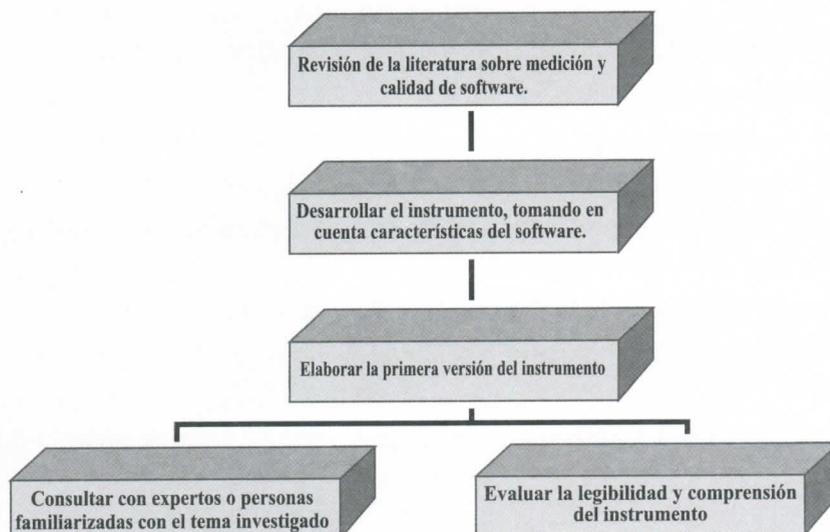
### *Muestra*

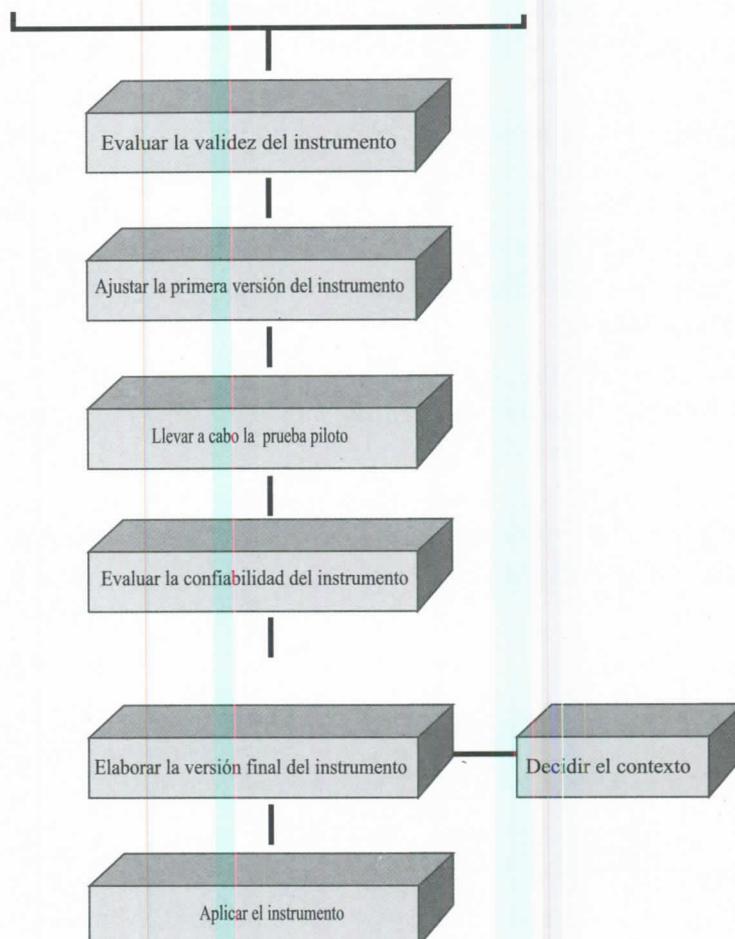
Para evaluar la validez y confiabilidad del instrumento se consideró una muestra no probabilística, pues nuestro estudio, requirió no tanto de una representatividad de elementos de una población, sino de una cuidadosa y controlada elección de sujetos con ciertas características específicas. (13).

Para evaluar la validez de contenido del instrumento la muestra fue de 7 expertos en Computación, todos con más de 5 años de experiencia en software comercial. Para evaluar la legibilidad y comprensión del instrumento se tomó una muestra de 10 personas relacionadas a la Computación y para evaluar la confiabilidad del instrumento se tomó una muestra de 12 usuarios de software de 4 empresas peruanas.

### *Procedimiento de la Investigación*

El proceso del diseño del instrumento se muestra en la Figura N° 01.





**Figura N° 01:** Proceso para construir un instrumento de medición

### ***Instrumentos Utilizados***

Para evaluar la validez y confiabilidad del instrumento diseñado se utilizaron los instrumentos respectivos. (Anexo I, Anexo II y Anexo III del informe completo).

### ***Procedimiento de recolección de información***

La recolección de información para la elaboración del instrumento se realizó mediante investigación bibliográfica y búsquedas de información a través de internet.

Para evaluar la validez del instrumento se recolectó la información, mediante la aplicación del instrumento del Anexo I, a 7 expertos en Computación con más de 5 años en el manejo de software. La aplicación se realizó a través de internet (correo electrónico).

Para evaluar la legibilidad y comprensión del instrumento se recolectó la información mediante la aplicación del instrumento del Anexo II a 10 bachilleres en Informática de la Universidad Nacional de Trujillo. La aplicación del instrumento se hizo en un aula durante 30 minutos.

Finalmente, para evaluar la confiabilidad del instrumento la recolección de la información se realizó mediante la autoadministración del instrumento del Anexo III a una muestra piloto de 12 personas, quienes evaluaron 4 software de 4 empresas diferentes. La distribución de los evaluadores fue como se observa en la Tabla N° 01.

**Tabla N° 01: Distribución de los evaluadores del instrumento**

Tipo de evaluador	N° Ítems	Software I	Software II	Software III	Software IV
Usuario	1-30	4	3	3	2

Por reserva de la información no se menciona el nombre del software y de la empresa donde fue tomada, sino se reemplaza por el de software I, software II, software III y software IV.

### III. RESULTADOS

Se evaluó la validez de contenido, para lo cual se siguió dos pasos. El primer paso fue elaborar el instrumento inicial en base al estándar de calidad ISO/IEC 9126. El segundo paso fue aplicar el instrumento al panel de expertos, conformado por 7 profesionales en Computación, todos con más de cinco años de experiencia con software en diversas organizaciones del Perú, quienes evaluaron el grado de relación de los ítems con cada una de las subcaracterísticas de calidad dada por el estándar de calidad del software ISO/IEC 9126 (12). La evaluación del instrumento fue en forma independiente y para determinar el grado de relación de los ítems con la característica de calidad se usó una escala bipolar (13) como se aprecia en la Figura N° 2.

Sin relación				Bien relacionada		
1	2	3	4	5	6	7

**Figura N° 2: Escala utilizada por los expertos para determinar el grado de relación de los ítems con la característica de calidad del software**

Para determinar la confiabilidad del instrumento para medir la calidad de un software de gestión se aplicó la segunda versión del instrumento conformada por 30 ítems, a una muestra conformada por 12 usuarios de software. El instrumento que se aplicó modificado a partir de la primera versión como consecuencia de los resultados de la validación de contenido por los expertos y de los resultados de la prueba de legibilidad y comprensión.

El instrumento de medición propuesto se basó en el estándar de calidad ISO/IEC 9126 (12). El instrumento diseñado se compone de un total de 30 ítems o preguntas, para ser administrado a los usuarios del software. Para la valoración de cada ítem se utilizó un escalamiento tipo Likert (13), como puede apreciarse en la Tabla N° 02.

**Tabla N° 02: Escala de valoración de los ítems del IMFYU**

<i>Valor</i>	<i>Clasificación</i>
4	Muy de acuerdo
3	De acuerdo
2	Ni de acuerdo, ni en desacuerdo
1	En desacuerdo
0	Muy en desacuerdo

Los resultados obtenidos se pueden interpretar según la Tabla N° 03.

**Tabla N° 03: Interpretación de las valoraciones**

<i>Valor</i>	<i>% Calidad</i>	<i>Clasificación</i>	<i>Criterio de Evaluación</i>	<i>Conclusión</i>
4	100%	Excelente	No se requieren hacer modificaciones en el software	Aceptado
3	75%	Buena	Se necesitan hacer pequeñas modificaciones en el software	Aceptado
2	50%	Regular	Se deben hacer grandes modificaciones en el software	Rechazado
1	25%	Mala	El software debe ser elaborado nuevamente	Rechazado
0	0%	Pésima	El software debe ser elaborado nuevamente	Rechazado

#### IV. DISCUSIÓN

La medición de la Funcionalidad y Usabilidad de un software comercial es una tarea difícil debido a las múltiples características que la componen y a la dificultad que entraña su medición directa de una manera objetiva.

El instrumento de autoadministración desarrollado en el presente estudio, al cual llamamos Instrumento para Medir la Funcionalidad y Usabilidad de un Software Comercial (IMFYU), se compone de 30 ítems.

El instrumento propuesto pretende cubrir la necesidad de contar con un instrumento específico que permita a los profesionales de Computación medir la Funcionalidad y Usabilidad de un software comercial de una manera rápida, válida y confiable.

**La rapidez** se pone de manifiesto en el tiempo necesario para la aplicación del IMFYU, que es corto, y en su versión final fue aproximadamente de 25 minutos, lo que está dentro de la recomendación planteada por Hernández y otros (13), quienes afirman que la aplicación de un instrumento que dura más de 35 minutos resulta tedioso.

**La validez** del IMFYU se realizó mediante la técnica de validez de contenido, para lo cual se determinó los ítems que serían incluidos en el instrumento mediante dos criterios: la desviación estándar y la media de las valoraciones dadas por los expertos (Tabla N° 03).

Además, utilizando el paquete estadístico SPSS 9.0 para Windows y los datos de la Tabla N° 03, el grado de acuerdo entre los 7 panelistas, en relación con las preguntas dirigidas a establecer la

validez de contenido, tuvieron un coeficiente de correlación intraclase de 0,9266 para un intervalo de confianza del 95% (IC 95% 0,8598-0,9685), con lo cual se verifica la validez de contenido; por lo tanto no se consideró necesario repetir la prueba de validez de contenido.

La confiabilidad del IMFYU se realizó mediante el coeficiente alfa de Cronbach, a partir de los datos de muestras de 4 software diferentes (Tabla N° 04) y la ayuda del paquete estadístico SPSS 9.0 para Windows. El alfa de Cronbach del instrumento obtuvo un valor promedio de 0,8533, lo que refleja un buen nivel de consistencia interna, pues se encuentran por arriba de 0,70 y no rebasan el 0,90, punto de corte que, de ser rebasado, podría reflejar redundancia de ítems.

## V. CONCLUSIONES

1. Se logró diseñar un instrumento para medir la funcionalidad y usabilidad de un software comercial, al que denominamos IMFYU, que es fácil de contestar por los usuarios y fácil de calificar por cualquier persona con un mínimo de conocimiento matemático.
2. **Utilizando el paquete estadístico SPSS 9.0 para Windows, se obtuvo una validez de contenido con un coeficiente de correlación** intraclase de 0,9052 para un intervalo de confianza del 95% (IC 95% 0,8565-0,9417) y una confiabilidad con un coeficiente alfa de Cronbach promedio de 0,8616. **Por lo tanto, la validez y la confiabilidad demostrada indican que es posible diseñar un instrumento que permita medir la funcionalidad y usabilidad de un software comercial, como el IMFYU.**
3. La evaluación de la calidad de un software está basada en modelos y métricas de calidad reconocidas (a partir de características y atributos) pero las decisiones están libradas a la intuición y al juicio surgido de la experticia humana, antes que a modelos y métodos cuantitativos de puntaje.
4. La mayoría de las técnicas y prácticas actuales para la evaluación de la calidad del software son ad-hoc o están emergiendo. No existe hasta el momento una metodología amplia y reconocida, cuantitativa y objetiva, para evaluar la calidad de un software.

## VI. REFERENCIAS BIBLIOGRÁFICAS

1. Garzotto F, Mainetti L, Paolini P. Hypermedia Design, Analysis, and Evaluation Issues. US: Ed. Addison-Wesley; 1995.
2. Garzotto F, Schwabe D, Paolini P. HDM, a model based approach to Hypermedia Application Design. ACM Transaction on Information System, (USA), 11 (1): 1-26; Enero 1993.
3. Garzotto F, Paolini P, Schwabe D. HDM, a model for a design of Hypertext Application, Proceed of Hypertext. US: Ed. ACM Press; 1991.
4. Kirakowski J, Corbett M. SUMI: the Software Usability Measurement Inventory. Brit J Ed Technology, 23; 1994.
5. Kirakowski J, Cierlik B. Measuring the Usability of Web Sites: Human Factors and Ergonomics Society Annual Conference. Chicago, US; 1998.
6. Buglione Luigi, Abran Alain. A Quality Factor for Software [en línea]. Italia, [fecha de consulta: 12 de junio 2002].  
Disponible en: <<http://space.tin.it/scienza/luigibug/qf/df.htm>>
7. Pressman Roger. Ingeniería de Software: un enfoque práctico. 4ª ed. Madrid: Ed. McGraw-Hill; 1998. p. 581.

8. Piattini Mario, Calvo-Manzano José, Cerera Joaquín, Fernández Luis. Análisis y diseño detallado de aplicaciones informáticas de gestión. México: Ed. Alfa Omega-Rama; 2000.
9. AENOR, UNE-EN-ISO 8402 Gestión de la calidad y aseguramiento de calidad. Vocabulario (ISO 8402:1994), AENOR; 1995.
10. Fernández Sanz Luis, Alarcón Rodríguez Miren. Necesidades de medición en la gestión y el aseguramiento de calidad del software [en línea]. [fecha de consulta: 1 de julio 2003]. Disponible en: <<http://www.sc.ehu.es/jiwdocoj/remis/docs/aseguracal.htm>>
11. IEEE Std 1061. IEEE Standard for a Software Quality Metrics Methodology. US: Ed. IEEE Computer Society Press; 1992.
12. ISO/IEC 9126 International Standard. Information technology – Software product evaluation – Quality characteristics and guidelines for their use. 1991.
13. Hernández Sampieri Roberto, Fernández Collado Carlos, Baptista Lucio Pilar. Metodología de la Investigación. 2ª ed. México: Ed. McGraw-Hill; 1999.

## VII. ANEXO

### Instrumento para medir la Funcionalidad y Usabilidad de un Software Comercial

Pregunta	Valoración				
	4	3	2	1	0
1. El software cumple con todos los requerimientos	4	3	2	1	0
2. El software siempre hizo lo que yo estaba esperando	4	3	2	1	0
3. Tiene consistencia interna (uniformidad de las pantallas, menús, reportes, mensajes, opciones, etc.)	4	3	2	1	0
4. Los resultados producidos por el software son correctos.	4	3	2	1	0
5. Se conecta y opera fácilmente con otros sistemas	4	3	2	1	0
6. Los términos usados están estandarizados (por ejm. Adicionar, Insertar, Agregar debe usarse un solo término.	4	3	2	1	0
7. Los símbolos (iconos) usados están estandarizados.	4	3	2	1	0
8. Previene el acceso no autorizado (sea accidental o deliberado) al software y datos.	4	3	2	1	0
9. Existen controles de acceso adecuado al software, subsistemas, funciones, archivos de datos, etc.	4	3	2	1	0
10. El uso del software es fácil de aprender	4	3	2	1	0
11. Los comandos del software se aprenden rápidamente	4	3	2	1	0
12. La información del software se presenta de manera clara y entendible.	4	3	2	1	0
13. La terminología usada se relaciona bien con el trabajo que hago	4	3	2	1	0

14.	Los mensajes que aparecen en la pantalla son claros	4	3	2	1	0
15.	La organización de los menús o listas de información parece bastante lógica	4	3	2	1	0
16.	Las interfaces o pantallas representan los objetos (por ejm, íconos) con análogos en mi ambiente de trabajo.	4	3	2	1	0
17.	La presentación del software es muy atractiva	4	3	2	1	0
18.	Los mensajes de entradas de datos son entendibles	4	3	2	1	0
19.	Los mensajes de prevención de error son adecuados	4	3	2	1	0
20.	El tipo, tamaño y formato de texto es correcto	4	3	2	1	0
21.	Es fácil moverse de una parte de una tarea a otra	4	3	2	1	0
22.	Los errores de ingreso de datos los puedo corregir fácilmente	4	3	2	1	0
23.	Hay tareas en el software que requieren de demasiados pasos	4	3	2	1	0
24.	Tengo que buscar ayuda la mayoría de veces cuando uso este software	4	3	2	1	0
25.	La información de ayuda dada por este software es muy útil (utilidad de ayudas)	4	3	2	1	0