

Re-evaluando la Intención de Uso de un Procedimiento de Medición basado en COSMIC-FFP

Nelly Condori-Fernández y Oscar Pastor
Departamento de Sistemas Informáticos y Computación
Universidad Politécnica de Valencia, Camino de Vera s/n , 46022, Valencia
nelly@dsic.upv.es, opastor@dsic.upv.es

Abstract

This paper describes the replication of an empirical study that was designed to evaluate the adoption of RmFFP in practice. RmFFP is a measurement procedure designed to measure the functional size of object-oriented systems from requirements specifications obtained in the context of the OO-Method approach. This procedure has been designed in accordance with the COSMIC-FFP standard method. The evaluation is based on the Method Adoption Model (MAM), where the intention to use a method is determined by the users' perceptions. The results show that an intention to use RmFFP exists, and that it is more influenced by usefulness than ease of use.

Keywords: COSMIC-FFP, functional size, RmFFP procedure, Method Adoption Model, Intention of use.

Resumen

Este artículo describe la réplica de un estudio empírico que fue diseñado para evaluar la adopción de RmFFP en la práctica. RmFFP es un procedimiento de medición diseñado para medir el tamaño funcional de sistemas orientado a objetos a partir de especificaciones de requisitos obtenidos en el contexto de OO-Method. Este procedimiento ha sido diseñado de acuerdo con el método estándar COSMIC-FFP. La evaluación fue llevada a cabo en base al Modelo de Adopción de Métodos (MAM), donde la intención a usar un método es determinado por las percepciones de los usuarios. Los resultados muestran que existe una intención de usar RmFFP y que esta intención de uso es determinado más por su utilidad percibida que por su facilidad de uso.

Palabras clave: COSMIC-FFP, tamaño funcional, procedimiento RmFFP, Modelo de Adopción de Métodos, intención de uso.

1. Introducción

Los métodos de medición de tamaño funcional (FSM) juegan actualmente un rol crucial en la administración de proyectos de software, siendo IFPUG FPA el método más popular y con mayor aceptación en la industria. Sin embargo, una evolución rápida de los paradigmas de desarrollo ha ocasionado un nuevo método de medición de tamaño funcional, COSMIC-Full Function Point (FFP) [1], el cual es compatible con los conceptos modernos de ingeniería de software y aplicable a varios dominios del software. Además, COSMIC-FFP ha sido aprobado por el estándar ISO/IEC19761 [2] y es considerado como el primer método de segunda-generación.

En los últimos años, se ha estado trabajando sobre un método basado en la transformación de modelos llamado OO-Method [3], el mismo que es soportado por la herramienta de generación de código automático Oliva Nova [4]. Esta herramienta incluye un módulo que permite medir en puntos de función el tamaño funcional de los modelos conceptuales (modelo de objetos, modelo dinámico, y modelo funcional) [5]. Sin embargo, este proceso de medición es llevado a cabo durante la fase de análisis del proceso de desarrollo de OO-Method, lo que da lugar a querer enriquecer la herramienta OLIVANOVA con la idea de estimar el tamaño funcional a partir de una etapa más temprana mediante la utilización de especificaciones de requisitos.

Para esto, se diseñó un procedimiento de medición de tamaño funcional llamado RmFFP, basado en el método estándar COSMIC-FFP con el fin de estimar el tamaño funcional de los sistemas orientado a objetos generados con OO-Method [6]. Además este procedimiento ha sido aplicado a varios casos de estudio (un sistema de gestión de alquiler de coches, un sistema de gestión de golf y un sistema de gestión de servicios de mantenimiento de un hospital) con el fin de evaluar su reproducibilidad y productividad [7]. Sin embargo, existe también la necesidad de valorar la respuesta de los usuarios hacia el nuevo procedimiento y evaluar su intención de ser usado en el futuro. Para esto, hemos diseñado un estudio empírico que ha sido llevado a cabo dos veces. En la primera evaluación, se utilizó como sujetos experimentales a estudiantes de informática de último año. Sin embargo, en la segunda evaluación se utilizó a estudiantes de doctorado, siendo esta última evaluación la constituye el ámbito del presente artículo.

El estudio empírico diseñado está basado sobre un modelo teórico llamado Modelo de la Adopción de Métodos (MAM) [8]. Este modelo incluye los mismos constructores primarios del Modelo de Aceptación de Tecnologías [9] que han sido adaptados para explicar y predecir la adopción de métodos. Estos constructores son:

- **Facilidad de Uso Percibida:** Grado de esfuerzo que una persona percibe al usar un método en particular.
- **Utilidad Percibida:** Grado de eficacia de un método en particular que es percibido por una persona para el logro de sus objetivos propuestos.
- **Intención de Uso:** Grado de intencionalidad que tiene una persona para usar un método particular.

MAM reconoce que las percepciones de eficiencia (facilidad de uso) y efectividad (utilidad) de un método juegan un rol importante para que dicho método sea adoptado en la práctica, este modelo fue aplicado también por Poels [10] y Abrahao [11].

Este artículo es organizado de la siguiente manera: La sección 2 presenta una descripción general del proceso de evaluación utilizado. La sección 3 describe la evaluación de la intención de usar RmFFP en el futuro. En la sección 4 se discute el análisis e interpretación de estos resultados. Finalmente, en la sección 5 presentamos nuestras conclusiones y trabajos futuros.

2. Descripción general

El proceso de evaluación llevado a cabo se inició con la selección de participantes, quienes formaron parte de un proceso de entrenamiento. Este entrenamiento tiene como propósito desarrollar cierto nivel de pericia requerido en los sujetos para medir especificaciones de requisitos usando el procedimiento de medición RmFFP. Tal como se observa en la Figura 1, un conjunto de materiales de entrenamiento fue utilizado, el cual fue previamente diseñado y elaborado. Este material consta de casos de estudio especificados con el Modelo de Requisitos de OO-Method [77], ejemplos del uso de RmFFP y guía de medición.

Finalizado el proceso de entrenamiento, los participantes inician la demostración de lo aprendido, mediante la medición de un caso de estudio que fue previamente seleccionado. Si esta demostración no es satisfactoria, se reforzará nuevamente con la sesión de entrenamiento. En caso contrario, se procederá a la captura de percepciones e intenciones

que tienen los participantes sobre el uso de RmFFP mediante la aplicación de un cuestionario, el cual es descrito con mayor detalle en la sección 3. Luego, cada una de las respuestas de los participantes son registradas, validadas y analizadas para finalmente ser interpretadas y presentadas en un informe.

Este proceso de evaluación descrito fue llevado a cabo dos veces. Los primeros resultados fueron reportados en [13]. Sin embargo, una replicación del estudio empírico fue llevada a cabo con el propósito de mejorar la fiabilidad de los resultados, cuya planificación experimental se describe en la siguiente sección.

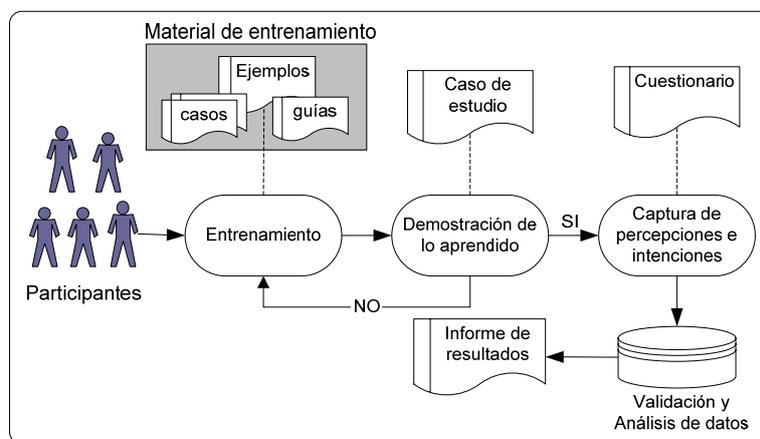


Figura 1. Proceso de evaluación de las percepciones e intención de usar RmFFP

3. Replicando la evaluación de la intención de uso de RmFFP

Usando la plantilla de GQM (Goal/Question/Metric) [14], la meta del presente estudio empírico fue de analizar las respuestas de los usuarios para valorar RmFFP con respecto a su intención de uso desde el punto de vista de los investigadores bajo el rol de estudiantes que miden especificaciones de requisitos de OO-Method.

3.1. Planificación

Los *sujetos experimentales* utilizados en el presente estudio fueron once estudiantes de doctorado de la Universidad Politécnica de Valencia, matriculados en el curso de "Tecnologías Software para Ambientes Web" durante el período de febrero a junio de 2006. Todos los sujetos estaban familiarizados con técnicas de modelado y medición.

La *variable independiente* en nuestro estudio se corresponde al procedimiento de medición de tamaño funcional RmFFP. Las *variables dependientes* se corresponden a tres

variables de percepción basados en el MAM: facilidad de uso percibida (FUP), la utilidad percibida (UP), y la intención de uso (IU).

Las hipótesis identificadas son las siguientes:

- H1: Hay intención de usar RmFFP.
- H2: La intención de uso es determinada por la facilidad de uso percibida y la utilidad percibida.

De la Hipótesis H2 podemos derivar dos hipótesis sencillas, que son las siguientes:

- H3: La intención de uso es determinada por la facilidad de uso percibida.
- H4: La intención de uso es determinada por la utilidad percibida.

Los instrumentos utilizados en este experimento fueron: el *objeto experimental* conformado por la especificación de requisitos de OO-Method para los casos de estudio: alquiler de coches, gestión de servicios de mantenimiento en un hospital, y administración de Golf. Los *materiales de instrucción* utilizados fueron los siguientes: material sobre el Modelo de Requisitos de OO-Method y el procedimiento de medición RmFFP; un ejemplo ilustrativo sobre la aplicación de RmFFP y Guía de medición.

El cuestionario original utilizado en el primer estudio fue ajustado para replicar esta evaluación inicial. La tabla 1 resume los cambios principales que fueron llevados a cabo.

Cuestionario original, adaptado de [5]	Cuestionario ajustado
<ul style="list-style-type: none"> • La FUP incluyó 5 preguntas: Q1, Q3, Q4, Q6 y Q9. • La UP incluyó 5 preguntas: Q2, Q5, Q8, Q10, y Q11. • La IU incluyó 3 preguntas: Q7, Q12, e Q13. • Las preguntas Q2 y Q11 fueron ubicados para evaluar la FUP, constituyéndose como preguntas Q2 y Q12 respectivamente. • La pregunta Q12 fue ubicada para evaluar la FUP constituyéndose en la pregunta Q14. 	<ul style="list-style-type: none"> • La FUP incluyó 8 preguntas: Q1, Q2, Q3, Q4, Q6, Q9, Q12 y Q14. • La UP incluyó 3 preguntas: Q5, Q8, y Q11. • La IU incluyó 4 preguntas: Q7, Q10', Q13', y Q15. • Las preguntas Q10' y Q13' fueron adicionados para evaluar la IU. • Las preguntas Q11 y Q15 se corresponden con las preguntas Q10 y Q13 (del cuestionario original) respectivamente.

Tabla 1. Las diferencias entre la inspección original y la inspección ajustada.

Por lo tanto, este cuestionario ajustado incluyó quince preguntas cerradas, utilizándose la escala likert de 5 puntos. La Figura 2 muestra la distribución de preguntas para cada una de las variables dependientes.

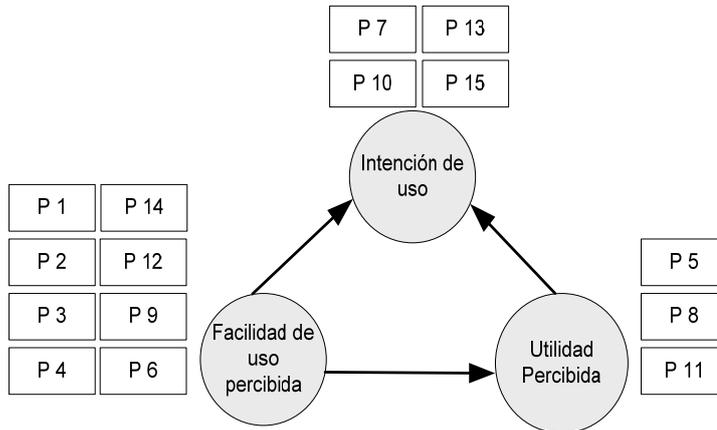


Figura 2. Organización del Modelo de Adopción de Métodos

3.2. Operación del experimento

Para que el diseño del estudio empírico sea ejecutado adecuadamente, la interacción entre los sujetos fue controlada para evitar que las respuestas sean influenciadas por opinión de otros sujetos. Sin embargo, ellos podían consultar al profesor del curso cualquier duda con respecto al cuestionario. No hubo límite en el tiempo para la ejecución de la encuesta.

3.3. Evaluación de validez

Para asegurar la validez de los resultados experimentales consideramos algunas amenazas para la validez de la construcción como la explicación inadecuada de los constructores. Esta amenaza significa que los constructores no son definidos adecuadamente y por lo tanto el experimento no pueden estar suficientemente claros. Para evitar esta amenaza, un análisis de correlación de inter-ítems fue llevado a cabo para evaluar la validez de los ítems que corresponden a los constructores del MAM (FUP, UP, IU).

Esta evaluación fue basada en un análisis de correlación inter-ítems. Empleamos dos criterios, la *Validez Convergente* (VC) y la *Validez del Discriminante* (VD), para cada ítem; si la validez convergente es más alta que la validez discriminante, el ítem es validado. Sin embargo, aplicando estos criterios se encontró que el valor de VC fue más bajo que el valor

de VD para los ítems Q2 y Q12 (vea Tabla A1 del apéndice); por esta razón estos dos ítems fueron extraídos del análisis.

Además, se realizó también un análisis de fiabilidad de los ítems validados utilizando la técnica de alfa-Chronbach, donde el valor alfa obtenido para cada constructor del MAM se muestra en la Tabla 2.

Constructor	Cronbach (α)	Sin Q2 y Q12
FUP	0.71	0.802
UP	0.818	0.818
IU	0.846	0.846

Tabla 2. Fiabilidad para los constructores de MAM.

Estos valores indican que los ítems del cuestionario son fiables ya que un alfa mayor o igual de 0,7 es aceptable (Nunally [15]). Además, el alfa de UP y la IU fue mejor que el alfa obtenido en la primera evaluación (UP = 0,5 y la IU = 0,5). Otros detalles puede consultarse en [13].

4. Análisis e Interpretación

Una vez que los datos fueron recolectados y validados, los puntajes de cada sujeto fueron promediados sobre las diferentes preguntas del cuestionario que son relevantes para cada constructor. De este modo obtuvimos tres valores promedios para cada sujeto. La tabla 3 muestra la estadística descriptiva para cada constructor del MAM; obsérvese que el valor promedio de la IU obtenido con once sujetos es mayor a 3 en una escala de 5 puntos.

Estadístico	FUP	UP	IU
Promedio	3.98	3.67	3.61
Desv. estándar	0.59	0.87	0.89
Mínimo	2.83	2.33	1.75
Máximo	5.00	5.00	5.00

Tabla 3. Valores descriptivos para los constructores del MAM.

Para evaluar la intención de uso de RmFFP, la hipótesis H1 fue probada formalmente, verificando si los puntajes de los sujetos asignados fueron apreciablemente mejores el puntaje medio en la escala de 5 puntos de Likert. Para esto, verificamos en primer lugar la

normalidad de estos puntajes utilizando la prueba de Shapiro-Wilk. Como la distribución de los datos fue normal, utilizamos la prueba de T-Student para comprobar la diferencia entre los valores promedio y el valor 3 (puntaje medio). La prueba estadística fue aplicada con un nivel de significación de 5 % ($\alpha = 0,05$). Los resultados obtenidos (Tabla 4) permiten el rechazo de la hipótesis nula con un nivel de significancia medio. Por tanto, corroboramos empíricamente que existe una intención de usar RmFFP en el futuro.

Estadístico	IU
Diferencia de medias	0.614
95% Intervalo de confianza	0.016 (inferior)
	1.212 (superior)
T	2.29
p-valor	0.022

Tabla 4. T-Student: Intención de Uso.

Para probar la *hipótesis H2*, la técnica de análisis de regresión fue aplicada. La ecuación resultante fue: $IU = -0.18 + 0.76 * UP + 0.25 * FUP$.

El modelo de regresión tuvo un nivel de significación medio ($P = 0,0133$), lo que significa que *H2* fue confirmado. El coeficiente de determinación ($R^2 = 0,66$) indica que el 66% de la variación total en la intención de uso puede ser explicada por la variación en la utilidad percibida y la facilidad de uso percibida.

Con respecto a la *Hipótesis H3*: Facilidad de Uso Percibida \rightarrow Intención de Uso. La ecuación de regresión resultante a partir del análisis es: $IU = 1,151 + 0,618 * FUP$. Esta ecuación obtenida tuvo un nivel de significación bajo ($p = 0,212$), lo que significa que *H3* no fue confirmado.

Finalmente, con respecto a la *Hipótesis H4*: Utilidad Percibida \rightarrow Intención de Uso. La ecuación de regresión resultante a partir del análisis es: $IU = 0,621 + 0,816 * UP$. La regresión tuvo un nivel alto de significación ($p = 0,003$), lo que significa que *H4* fue confirmado. El coeficiente de la determinación ($r^2 = 0,635$) mostró que el 63,5% de la variación total en la intención de uso puede ser explicada por la variación en la utilidad percibida. La figura 3 representa la regresión lineal obtenida.

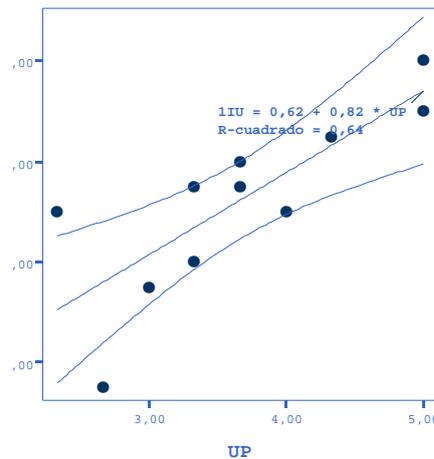


Figura 3. Modelo de regresión: IU vs. UP

5. Conclusiones y trabajos futuros

Este artículo describe la réplica de un estudio empírico que evalúa la intención de usar el procedimiento RmFFP, que fue diseñado para medir el tamaño de los sistemas orientado a objetos de acuerdo al método COSMIC-FFP. Los resultados indican que hay una intención de usar RmFFP para medir especificaciones de requisitos especificadas con OO-Method. Aunque RmFFP es percibido como fácil de usar, los resultados muestran que la utilidad percibida puede tener una mayor influencia sobre la intención a usar RmFFP que la facilidad de uso percibida. Esto significa que el usuario piensa utilizar RmFFP más por causa de su utilidad que por las otras ya mencionadas.

Por tanto, se puede concluir que la relación del MAM entre Facilidad de Uso Percibida e Intención de Uso no pudo ser verificada empíricamente en el dominio de la medición de tamaño funcional del software.

En un estudio futuro, planearemos la identificación y la evaluación de otras variables que pueden afectar a la intención de uso de un procedimiento de medición.

Agradecimientos

Los autores dan la gracias a Alain Abran para sus comentarios y sugerencias que contribuyeron a la mejora del presente artículo. Además, este trabajo ha sido subvencionado por el proyecto DESTINO con ref. TIN2004-03534 del Ministerio de Ciencia y Tecnología.

Referencias

- [1] Abran A., J. M. Desharnais, S. Oigny, D. St-Pierre, and C. Symons, *COSMIC-FFP Measurement Manual Version 2.2, The COSMIC Implementation Guide for ISO/IEC 19761:2003*, École of technologie supérieure- ETS, Montreal (Canada) 2003. Disponible en: www.gelog.etsmtl.ca/cosmic-ffp
- [2] ISO, *ISO/IEC 19761 Software Engineering-COSMIC-FFP-A Functional Size Measurement Method*, International Organization for Standardization_ISO, Geneva, 2003.
- [3] Pastor O., Gomez J., Insfran E., Pelechano V., 2001. "The OO-Method approach for information systems modelling: from object-oriented conceptual modelling to automated programming", *Journal Information Systems*, vol 26, pp. 507-534.
- [4] CARE S.A: <http://www.care-t.com/> , visitado el 15 September 2006
- [5] Abrahão S., *On the Functional Size Measurement of Object-Oriented Conceptual Schemas: Design and Evaluation Issues*, tesis doctoral, Department de Sistemas de Informáticos y Computación, Universidad Politécnica de Valencia, octubre, 2004.
- [6] N. Condori-Fernández, S. Abrahão, O. Pastor, "Towards a Functional Size Measure for Object-Oriented Systems from Requirements Specifications". *IEEE Quality Software Int. Conf.* 2004, Germany, pp. 94-101.
- [7] N. Condori-Fernández, O. Pastor, "Evaluating the Productivity and Reproducibility of a Measurement Procedure", *Springer LNCS*, ER Workshop on Quality of Information Systems, Minnesota, USA, November 2006.
- [8] Moody D. L., *"Dealing with Complexity: A Practical Method for Representing Large Entity Relationship Models"*, tesis doctoral, Department of Information Systems, University of Melbourne, Australia, 2001.
- [9] F. D. Davis, "Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology", *MIS Quarterly*, 1989, pp. 319-340.
- [10] Poels G., Maes A., Gailly F., Paemeleire R., "Measuring User Beliefs and Attitudes towards Conceptual Schemas: Tentative Factor and Structural Equation Model", Fourth Annual Workshop on HCI Research in MIS, December 2005.
- [11] Abrahao S., Poels G., and Pastor O. "A Functional Size Measurement Method for Object-Oriented Conceptual Schemas: Design and Evaluation Issues". *Journal Software & System Modelling*, Vol. 5 n°1, pp. 48-71, 2005.
- [12] Insfran E., Pastor O. and Wieringa R., "Requirements Engineering-Based Conceptual Modelling". *Journal Requirements Engineering*, , vol 7, n° 2, pp. 61-72, 2002.
- [13] N. Condori-Fernández, O. Pastor, "An Empirical Study on the Likelihood of Adoption in Practice of a Size Measurement Procedure for Requirements Specification", *IEEE Quality Software International Conference*, China, octubre, 2006.
- [14] Wohlin C., Runeson P., Höst M., M. C. Ohlsson, B. Regnell, and A. Wesslén, *Experimentation in Software Engineering: An Introduction*, Springer-Verlag, 2000.
- [15] Nunally J., *Psychometric Theory*, McGraw-Hill, 2nd ed., New York, NY1978.

Apéndice A

	Facilidad de Uso Percibida								Utilidad Percibida			Intención de uso				Promedio		Válidc
	P1	P2	P3	P4	P6	P9	P12	P14	P5	P8	P11	P7	P10	P13	P15	VC	VD	
FUP P1	1,00	-0,07	0,63	0,36	0,28	0,61	0,56	0,52	0,26	0,32	0,33	0,69	0,30	0,71	0,47	0,49	0,44	SI
P02	-0,07	1,00	-0,31	-0,44	-0,13	-0,26	0,27	0,09	0,38	0,15	0,09	0,31	0,13	0,24	0,08	0,02	0,20	NO
P03	0,63	-0,31	1,00	0,79	0,07	0,63	0,01	0,72	0,31	0,03	0,18	0,18	-0,11	0,29	0,39	0,44	0,18	SI
P04	0,36	-0,44	0,79	1,00	0,48	0,25	-0,12	0,49	0,28	-0,11	-0,12	-0,12	0,07	-0,07	0,27	0,35	0,03	SI
P06	0,28	-0,13	0,07	0,48	1,00	0,03	0,23	0,12	0,13	0,29	0,04	0,23	0,65	-0,08	0,20	0,26	0,21	SI
P09	0,61	-0,26	0,63	0,25	0,03	1,00	0,18	0,74	0,26	0,43	0,60	0,50	-0,11	0,44	0,37	0,40	0,35	SI
P12	0,56	0,27	0,01	-0,12	0,23	0,18	1,00	-0,02	0,47	0,79	0,70	0,82	0,72	0,83	0,68	0,27	0,71	NO
P14	0,52	0,09	0,72	0,49	0,12	0,74	-0,02	1,00	0,57	0,16	0,22	0,25	-0,21	0,33	0,20	0,46	0,22	SI
UP P05	0,26	0,38	0,31	0,28	0,13	0,26	0,47	0,57	1,00	0,49	0,42	0,38	0,31	0,56	0,52	0,64	0,37	SI
P08	0,32	0,15	0,03	-0,11	0,29	0,43	0,79	0,16	0,49	1,00	0,91	0,71	0,66	0,64	0,58	0,80	0,39	SI
P11	0,33	0,09	0,18	-0,12	0,04	0,60	0,70	0,22	0,42	0,91	1,00	0,74	0,43	0,62	0,70	0,78	0,38	SI
IU P07	0,69	0,31	0,18	-0,12	0,23	0,50	0,82	0,25	0,38	0,71	0,74	1,00	0,57	0,73	0,76	0,77	0,43	SI
P10	0,30	0,13	-0,11	0,07	0,65	-0,11	0,72	-0,21	0,31	0,66	0,43	0,57	1,00	0,38	0,53	0,62	0,26	SI
P13	0,71	0,24	0,29	-0,07	-0,08	0,44	0,83	0,33	0,56	0,64	0,62	0,73	0,38	1,00	0,52	0,66	0,41	SI
P15	0,47	0,08	0,39	0,27	0,20	0,37	0,68	0,20	0,52	0,58	0,70	0,76	0,53	0,52	1,00	0,70	0,41	SI

Tabla A1. Análisis de correlación inter-ítems.