

Un Análisis Experimental sobre el Efecto del Acoplamiento en la Comprensibilidad y Facilidad de Modificación de Expresiones OCL

L. Reynoso, M. Genero, M. Piattini y E. Manso

Resumen— Para solucionar el problema de la limitada capacidad expresiva de la notación de diagramas de UML, a la hora de expresar ciertas restricciones, se definió el lenguaje de restricciones OCL.

La relevancia que ha adquirido el uso de OCL nos motivó hace unos años a investigar aquellos aspectos que hacen que una expresión OCL sea más fácil de comprender y de modificar. Para cuantificar tales aspectos, definimos de manera metodológica un conjunto de medidas.

El principal objetivo de este artículo es la descripción de una familia de experimentos que hemos llevado a cabo con estudiantes de diversas universidades, para corroborar si existe una relación entre el acoplamiento (cuantificado con las medidas definidas) y la comprensibilidad y la facilidad de modificación de expresiones OCL. A partir de los resultados obtenidos tras la experimentación, se obtuvo evidencia empírica de que tal relación existe. Sin embargo, somos conscientes de que es necesario realizar experimentos adicionales con profesionales en entornos reales y también realizar casos de estudios para obtener un cuerpo de conocimiento sólido sobre el efecto del acoplamiento a la hora de comprender y modificar expresiones OCL.

Índice de Términos — Métricas de Software, Acoplamiento, Calidad de Software, OCL, UML, Modelado, Experimentos.

I. INTRODUCCIÓN

LA importancia del uso de modelos en el proceso de desarrollo de software Orientado a Objetos (OO) está siendo cada vez más reconocida. Este hecho se corrobora con las recientes iniciativas de desarrollo de software como

Este trabajo de investigación es parte del proyecto ENIGMAS (PCC-03-003-1) financiado por la “Consejería de Ciencia y Tecnología de la Junta de Comunidades de Castilla-La Mancha”, del proyecto CALIPO financiado por la “Dirección General de Investigación del Ministerio de Ciencia y Tecnología (España)” (TIC2003-07804-C05-03), y el proyecto MEC-FEDER (TIN2004-03145).

L. Reynoso pertenece al Departamento de Ciencias de la Computación, Universidad Nacional del Comahue, Argentina (e-mail: lreynoso@uncoma.edu.ar).

M. Genero y M. Piattini pertenecen al Grupo de Investigación Alarcos, Departamento de Tecnologías y Sistemas de Información, Universidad de Castilla La Mancha, España (e-mail: Marcela.Genero@uclm.es y Mario.Piattini@uclm.es).

M^a.E. Manso pertenece al grupo de Investigación GIRO, Departamento de Ciencias de Computación, Universidad de Valladolid, España (email: manso@infor.uva.es).

son el Modelado Dirigido por Modelos (Model-Driven Development, MDD) [1] y la Arquitectura Dirigida por Modelos (Model-Driven Architecture, MDA) [15], las cuales sostienen que el esfuerzo de desarrollo de software debe centrarse en el desarrollo de modelos [19].

Actualmente, el Lenguaje Unificado de Modelado (Unified Modeling Language, UML) [14] es el lenguaje estándar para el modelado OO. La expresividad de las técnicas de modelado utilizadas (por ejemplo la notación, etc.) afecta a la comprensibilidad de los modelos [19]. Los diagramas UML sólo proveen una vista adecuada de la arquitectura de software [10] pero son imprecisos debido a que la notación basada en diagramas no es suficientemente expresiva para expresar ciertas restricciones [8]. Para subsanar esta carencia de UML surgió el Lenguaje de Restricciones de Objetos (Object Constraint Language, OCL) [13]. Y así surgieron lo que actualmente se conoce como modelos UML/OCL combinados.

Debido a la importancia de OCL en el modelado OO algunos investigadores han corroborado empíricamente que OCL permite mejorar significativamente la comprensibilidad y la facilidad de mantenimiento de los modelos basados en UML [5]. Continuando con ese interés despertado sobre el uso de OCL hemos definido en [17] un conjunto de medidas con el objetivo de obtener indicadores de la comprensibilidad y facilidad de modificación de las expresiones OCL (ver tabla 1). Dichas medidas fueron definidas de manera rigurosa y metodológica aplicando un proceso de definición de medidas que consta de tres pasos principales [6]: definición, validación teórica y empírica de las medidas. Si bien las tres etapas son relevantes en este artículo nos centraremos en la validación empírica de las medidas propuestas. Como muchos autores han mencionado [2], [9], [11], [18] la validación empírica de medidas, a partir de experimentos es fundamental para asegurar que las medidas definidas sean realmente significativas y útiles en la práctica.

En [16] presentamos una familia de experimentos (un experimento y sus dos réplicas) para evaluar si existe correlación entre la profundidad de navegación (medida por DN) y la cantidad de diferentes objetos acoplados (NNC) de una expresión OCL y su comprensibilidad y facilidad de modificación. A partir de dicho estudio empírico obtuvimos que la comprensibilidad y facilidad de

modificación de expresiones OCL es dependiente de cuán lejos de la instancia contextual están los objetos a los cuales ésta se acopla (expresado a través de la medida DN), y de cuántos objetos distintos están acoplados a la instancia contextual (expresado a través de la NNC). Los resultados obtenidos, aunque preliminares, nos motivaron a investigar el efecto del acoplamiento sobre la comprensibilidad y facilidad de modificación de las expresiones OCL. El acoplamiento ha sido históricamente un atributo muy crítico de los productos software OO [4], que es recomendable minimizar. Pero hasta ahora se ha investigado el acoplamiento a nivel de código. Creemos que la disponibilidad de información sobre el acoplamiento en las etapas iniciales, a nivel de modelos UML/OCL combinados, sería útil por ejemplo para detectar a qué clases se les debería prestar mayor atención y verificación. Consideramos que un modelo UML/OCL revela mayor información de acoplamiento que un modelo especificado solamente con UML, debido al hecho de que la navegación OCL define un acoplamiento entre los objetos involucrados [21]. Por consiguiente, el objetivo de este artículo es describir una segunda familia de experimentos que hemos llevado a cabo con alumnos de diferentes universidades para corroborar si existe correlación entre el acoplamiento y la comprensibilidad y facilidad de modificación de expresiones OCL. De ser así, las medidas definidas serían empíricamente válidas y además habremos obtenido indicadores de la comprensibilidad y facilidad de modificación de las expresiones OCL útiles para los diseñadores.

Este artículo comienza con la descripción de una familia de experimentos, en la sección 2. Luego, en la sección 3 se presenta el análisis e interpretación de los datos empíricos. Finalmente, la última sección presenta las conclusiones y algunas líneas de investigación que surgen del presente trabajo.

TABLA I
MEDIDAS PARA EXPRESIONES OCL

Medida	Nombre
NNR	Número de Relaciones Navegadas (Number of Navigated Relationships)
NAN	Número de Atributos referidos a través de Navegaciones (Number of Attributes referred through Navigations)
NNC	Número de Clases Navegadas (Number of Navigated Classes)
WNCO	Número Ponderado de Operaciones de Colección (Weighted Number of Collection Operations)
DN	Profundidad de las Navegaciones (Depth of Navigations)
WNN	Número Ponderado de Navegaciones (Weighted Number of Navigations)
NEI	Número de Variables Iteradoras Explícitas (Number of Explicit Iterator Variables)
NKW	Número de Palabras Reservadas de OCL (Number of OCL KeyWords)
NES	Número de <i>Self</i> Explícitos (Number of Explicit <i>Self</i>)
NCO	Número de Operadores de Comparación (Number of Comparison Operators)

II. FAMILIA DE EXPERIMENTOS

Estudios empíricos aislados difícilmente proveen respuestas definitivas [12] [3]. Mediante la ingeniería de software empírica se obtienen resultados más sólidos y relevantes a partir de familias de experimentos en lugar de ejecutar un experimento individual. Por ello, con el propósito de evaluar el efecto del acoplamiento sobre la comprensibilidad y facilidad de modificación de expresiones OCL hemos llevado a cabo una familia de experimentos. Dicha familia consistió en tres experimentos realizados en distintas universidades de diferentes países. Aunque hemos seguido el proceso experimental sugerido por Ciolkowski et al. [7] y por Wohlin et al. [22] por razones de brevedad sólo mencionaremos sus principales características.

En el primer experimento (Abril de 2004) invitamos a los estudiantes de 3^{er} año de Informática de la Universidad de Alicante (UA, España) a participar de un seminario sobre OCL de 5 horas de duración y como parte de dicho seminario les solicitamos realizar el experimento. Participaron 60 estudiantes, quienes estaban motivados a participar en el experimento debido a que podían obtener un punto extra en el resultado final del examen final de la asignatura Ingeniería de Software. Denominamos a los datos obtenidos en este experimento “UAE”.

En la primera réplica (Octubre de 2004) participaron 26 estudiantes de un curso de OCL realizado en la Escuela Internacional de Ciencias de Computación (organizado por la Universidad de La Matanza, Argentina). La duración del curso fue de 20 horas y durante las últimas dos horas ejecutamos la réplica del experimento. Los sujetos fueron estudiantes de grado de diferentes cursos y docentes de distintas universidades. Los datos obtenidos en esta réplica se denominan “ULME”.

La segunda réplica (Noviembre de 2004), fue realizada con 29 estudiantes de la asignatura de Ingeniería de Software de la Universidad Austral de Chile, quienes participaron en un curso de 20 horas sobre OCL. Se aplicó un incentivo similar al utilizado en la Universidad de Alicante para motivar la participación de los estudiantes. Les informamos a los estudiantes que si ellos completaban un examen, que consistía en la realización de un experimento, obtendrían una puntuación extra para el examen final de la asignatura de Ingeniería de Software. Los datos obtenidos fueron denominados “UACHe”.

Los tres cursos de OCL y sus correspondientes experimentos estuvieron a cargo del mismo profesor, concretamente el profesor Luis Reynoso, coautor de este artículo. A continuación describiremos las principales características del proceso experimental seguido en los tres experimentos:

Variables Independientes y Dependientes: La variable independiente (VI) es el acoplamiento de expresiones OCL. Las variables dependientes (VDs) son la comprensibilidad y facilidad de modificación de las expresiones OCL.

Material Experimental: Los objetos experimentales

fueron nueve diagramas de clases, cada uno de los cuales contenía una expresión OCL. Los diagramas y sus expresiones OCL fueron diseñados para cubrir un amplio rango de valores de las medidas que nos interesan. En realidad, es imposible cubrir todas las combinaciones posibles de los valores de medidas. Inicialmente diseñamos quince diagramas, pero varios de ellos eran similares y el hecho de tener diagramas con el mismo nivel de acoplamiento podía sesgar los resultados. Por esta razón, llevamos a cabo una clasificación jerárquica de los 15 diagramas agrupándolos en tres grupos (de acuerdo a los valores de las medidas) indicando tres niveles de acoplamiento: Bajo, Medio y Alto (identificamos cada nivel por sus siglas AB, AM, AA respectivamente). Finalmente, extrajimos tres modelos de cada grupo.

Instrumentación: Las tareas realizadas por los sujetos fueron:

1. Tareas de Comprensión: Los sujetos debían contestar un cuestionario que contenía cuatro preguntas y cuyo propósito era comprobar si los mismos habían comprendido la expresión OCL asociada a cada diagrama de clases.
2. Tareas de Modificación: Los sujetos tenían que modificar la expresión OCL de acuerdo a nuevos requisitos expresados en lenguaje natural.
3. Tareas de Valoración: Luego de finalizar cada tipo de tarea, ya sea de comprensión o de modificación, los sujetos debían valorar utilizando una escala de cinco etiquetas lingüísticas, la complejidad o dificultad de las tareas realizadas (por ejemplo, para valorar las tareas de comprensión se utilizaron las etiquetas “Muy fácil de comprender”, “Algo fácil de comprender”, “Normal”, “Algo difícil de comprender”, “Muy difícil de comprender”). La valoración de las tareas nos daba una idea de la percepción de los sujetos acerca de cuán difícil resultó para ellos comprender y modificar las respectivas expresiones OCL.

Le asignamos aleatoriamente tres diagramas a cada sujeto, cada uno contenía una expresión OCL de distinto nivel de acoplamiento (AB, AM y AA). En este artículo identificamos con C_1 a la colección de tests realizados en primer lugar por todos los sujetos, con C_2 a los realizados en segundo lugar y con C_3 a los realizados en tercer lugar. Los nueve diagramas fueron examinados por el mismo número de sujetos en cada C_i , obteniendo así un diseño balanceado.

La variable independiente fue medida a partir de las medidas mostradas en la Tabla 1. Concretamente utilizamos las medidas NNR, NAN, NNC, WNCO, DN, WNN y NES debido a que algún aspecto del concepto de navegación es capturado en la intensidad de sus definiciones [17]. También utilizamos la medida NEI la cual está relacionada con variables iteradoras de operaciones de colección, y permite definir el contexto dentro de las operaciones de colección. El resto de las medidas NKW

(Número de Palabras Reservadas) y NCO (Número de Operadores de Comparación) se utilizaron necesariamente para definir una expresión OCL simple pero no estaban relacionadas con el acoplamiento. Debido a que no estamos interesados en estudiar estas dos últimas medidas tratamos de conservar sus valores lo más constante posible. Por ejemplo todas las expresiones OCL utilizadas en los objetos experimentales fueron definidas con tres palabras reservadas de OCL.

Consideramos que el tiempo que tardaba cada sujeto al realizar las tareas requeridas (i.e., Tiempo de Comprensión (Tiempo COM) y Tiempo de Modificación (Tiempo MOD) no era la medición más precisa para las VDs. Por ello decidimos utilizar la Eficiencia de la Comprensión (Ef-COM) y la Eficiencia de la Modificación (Ef-MOD):

$$Ef - COM = \frac{respuestas_correctas}{Tiempo_COM} \quad (1)$$

$$Ef - MOD = \frac{modificaciones_correctas}{Tiempo_MOD} \quad (2)$$

A través de las tareas de valoración obtuvimos dos medidas denominadas Complejidad Subjetiva de la Comprensión (CompSub COM) y Complejidad Subjetiva de la Modificación (CompSub MOD) respectivamente. Estas medidas, a nuestro criterio, pueden reflejar estimar la carga cognitiva de los sujetos a la hora de manipular modelos UML/OCL combinados.

Formulación de Hipótesis: Formulamos diferentes hipótesis de acuerdo a nuestras suposiciones o creencias:

- Suposición 1: El nivel de acoplamiento en las expresiones OCL tiene influencia en el grado de corrección de las tareas de comprensión y modificación realizadas por tiempo, i.e. influyen en la eficiencia de los sujetos (Ef-COM o Ef-MOD). La hipótesis es:

Hipótesis 1: $H_{0,1}$ No hay correlación entre las medidas de acoplamiento definidas para expresiones OCL (ver tabla 1) y Ef-COM (ó Ef-MOD). $H_{1,1} = \neg H_{0,1}$

- Suposición 2: El nivel de acoplamiento de las expresiones OCL tiene influencia en la percepción de los sujetos sobre la dificultad o complejidad de las tareas de comprensión y modificación. La hipótesis es: **Hipótesis 2:** $H_{0,2}$ No existe correlación significativa entre las medidas para expresiones OCL relacionadas con el acoplamiento (ver tabla 1) y CompSub. $H_{1,2} = \neg H_{0,2}$

- Suposición 3: El criterio subjetivo de los sujetos al valorar las tareas de comprensión y modificación está influenciado por el tiempo que tardan en realizar dichas tareas. Por ejemplo, esperamos que los sujetos valoren aquellas tareas que más tiempo les requirieron como tareas “Algo Dificiles de Entender” ó “Muy Dificiles de Entender”. La hipótesis es:

Hipótesis 3: $H_{0,3}$ CompSub COM (ó MOD) no está correlacionada con Tiempo COM (ó MOD). $H_{1,3}: \neg H_{0,3}$

- Suposición 4: Creemos que el grado de corrección de las tareas realizadas por tiempo, i.e. Ef-COM (ó Ef-MOD), puede ser un indicador de la valoración subjetiva dada por los sujetos acerca de la complejidad o dificultad de las tareas de comprensión y modificación. La hipótesis:

Hipótesis 4: $H_{0,4}$ CompSub COM (ó MOD) no está correlacionada con Ef-COM (ó Ef-MOD). $H_{1,4}: \neg H_{0,4}$.

III. ANÁLISIS DE DATOS E INTERPRETACIÓN

En esta sección resumiremos los principales aspectos del análisis de los datos empíricos llevados a cabo con SPSS [20]. El análisis de los datos fue realizado de la siguiente forma:

- Primero llevamos a cabo un estudio exploratorio y descriptivo.
- Luego, verificamos las hipótesis formuladas. Como todas en todas las hipótesis se basan el grado de dependencia entre variables, realizamos un análisis de correlación. Los coeficientes de correlación como Spearman o Tau de Kendall se utilizan con pares de observaciones (X_i, Y_j) , sobre n -objetos (en nuestro caso 9 objetos), pero las observaciones deben ser independientes. Esto significa por ejemplo, que si estudiamos una variable dependiente, supongamos Ef-COM, del sujeto “j” en el i ésimo-diagrama no debemos considerar ninguna otra observación del mismo sujeto “j”. De este modo, las correlaciones de las hipótesis formuladas serán verificadas para cada C_i (el cual representa el conjunto de los i ésimos-diagramas utilizados por todos los sujetos experimentales).

A. Estudio exploratorio y descriptivo

Primero analizamos la normalidad de los datos a través del test de Shapiro-Wilk, descubriendo que no eran normales. Removimos los valores extremos (outliers) de todas las DVs. Como describimos previamente, los conjuntos C_i representan el orden de las tareas realizadas, el cual nos permite mostrar como varía el tiempo insumido en las tareas asociadas a cada diagrama. Luego de analizar las variables Tiempo COM y MOD a medida que el tiempo pasaba, comprobamos que el tiempo disminuía durante la ejecución del experimento. En el caso de Ef-COM y Ef-MOD, esperábamos que la eficiencia de los sujetos mejorara significativamente a medida que pasaba el tiempo, pero no fue así, excepto en el experimento UA para Ef-COM. Sin embargo, comprobamos que al organizar los datos de acuerdo a su nivel de acoplamiento (AB, AM y AA) tanto Tiempo COM como Ef-COM mejoraban a medida que el nivel de acoplamiento disminuía. Este no es el caso de Tiempo MOD y Ef-MOD debido a que las tareas de modificación correspondientes a los diagramas de acoplamiento medio (AM) resultaron más difíciles de modificar que las tareas correspondientes al acoplamiento alto (AC). Esta situación ocurrió en los tres experimentos.

La principal diferencia entre los diagramas AM y AA es que el acoplamiento de nivel AM está basado principalmente en navegaciones complejas, mientras que en los diagramas clasificados como AA el acoplamiento está basado en operaciones de colección entrelazadas unas con otras. Suponemos que para los sujetos fue más difícil - durante la modificación de expresiones- identificar que relaciones deberían ser utilizadas (su nombre de rol, su nombre de atributo, etc.) que identificar qué operaciones de colección debían ser utilizadas para realizar la modificación. Creemos que complementar los diagramas de clase UML con una descripción en lenguaje natural de la intención de las relaciones hubiera facilitado a los sujetos el darse cuenta de cual relación debían utilizar.

Las medias de Tiempo COM y Tiempo MOD presentaron valores más altos en UACHe comparados con los de ULME y UAE, y entre estos dos últimos, la media más pequeña corresponde a UAE. Los estudiantes chilenos tenían menor experiencia en UML, y esto les requirió mayor tiempo extra para completar la tarea. Por otro lado, aunque UAE presenta una media de Tiempo COM mayor que ULME los valores de Ef-COM fueron similares si comparamos sus C_i .

B. Contraste de las hipótesis 1 y 2

Para contrastar las primeras dos hipótesis, realizamos un análisis de correlación utilizando el coeficiente de Spearman's con un nivel de significación $\alpha = 0,05$, lo cual indica que el nivel de confianza es 95% (i.e. la probabilidad de aceptar H_0 cuando H_0 es verdadero es 0,95). No incluimos las tablas con los coeficientes correspondientes por problemas de espacio.

Las conclusiones a las que llegamos son:

- Hipótesis 1: Todas las medidas presentan un coeficiente de correlación negativo, excepto diversas medidas como NAN y NCO en Ef-MOD y NES y NCO en Ef-COM en algunas observaciones de los sujetos (i.e. para algún C_i). Un coeficiente negativo significa que a medida que suben los valores de las medidas los sujetos eran menos eficientes, lo cual es coherente, de acuerdo a nuestras suposiciones. Las medidas NNC, WNCO y NEI presentan diversas correlaciones con Ef-COM en los experimentos UAE y UACHe. Es lógico que el número de clases involucradas en las expresiones OCL (NNC), el número de operadores de comparación (NCO) y el número de variables iteradoras en operaciones de colección (NEI) influyan en la eficiencia de los sujetos. Esta influencia parece ser independiente del orden de las tareas realizadas para UAE debido a que encontramos correlación para la mayoría de los C_i . La profundidad de la navegación (DN) presenta correlaciones con Ef-MOD en los tres experimentos. NNN, NAN, NES y WNN también tienen diversas correlaciones con Ef-MOD, pero no para los tres experimentos. NAN, NES y NCO presentan un

coeficiente de correlación positivo, i.e. los sujetos eran más eficientes cuando los valores de las medidas eran más altos, lo que parece sospechoso.

- Hipótesis 2: Todas las medidas presentan un coeficiente de correlación positivo excepto diversos valores de NAN y NCC en CompSub MOD y NES y NCO en CompSub COM.

Encontramos pocas correlaciones entre las medidas y CompSub COM. Del conjunto de medidas que presentan correlación, sólo una de ellas está correlacionada dos veces. Los niveles de significación están entre 0,002 y 0,038.

DN, WNN y NNR están correlacionadas con CompSub MOD en los tres experimentos. Los niveles de significación están entre 0,000 y 0,041. DN tiene la correlación más fuerte en UAE, independientemente del orden de las tareas. Sin embargo en este experimento, la correlación de NNR y WNN es más alta a medida que el tiempo pasa.

C. Contraste de las hipótesis 3 y 4

Con el objeto de verificar las tercera y cuarta hipótesis, estudiamos la correlación utilizando tests no paramétricos para medidas ordinales. En primer lugar transformamos las variables CompSub COM y CompSub MOD, asignando números a las etiquetas lingüísticas: en un rango de 1 (asignado a "Muy Fácil de Comprender/Modificar") a 5 (correspondiendo a "Muy Difícil de Comprender/Modificar"). Utilizamos el coeficiente Tau de Kendall (aconsejable para medidas ordinales) para contrastar $H_{0,3}$ y $H_{0,4}$. Analizando los coeficientes de correlación llegamos a las siguientes conclusiones:

- Relación entre CompSub COM y Tiempo COM: En UAE y UChE parece que existe una correlación significativa entre la variable CompSub COM y Tiempo COM. Sin embargo en ULME sólo encontramos correlaciones en una observación (C_2).
- Relación entre CompSub MOD y Tiempo MOD: se obtuvieron los mismos resultados que en el caso previo.
- Relación entre CompSub COM (o MOD) y Ef-COM (o MOD): parece que existe una correlación significativa entre CompSub COM y Ef-COM, y entre CompSub MOD y Ef-MOD, en el caso de UAE y UChE. En ULME encontramos que CompSub MOD está correlacionada con Ef-MOD.

IV. CONCLUSIONES

En este artículo hemos presentado una familia de tres experimentos realizada con el objetivo de evaluar si realmente el acoplamiento influye a la hora de comprender y modificar las expresiones OCL. El acoplamiento lo cuantificamos a través de un conjunto de medidas que definimos previamente (ver tabla 1). Luego de realizar un exhaustivo análisis estadístico de los datos recogidos en los tres experimentos encontramos que parece que existe una

correlación estadísticamente significativa entre algunas de las medidas de acoplamiento y la eficiencia de las tareas de comprensión y modificación. Por otro lado, parece que el acoplamiento afecta de diferente forma a la comprensión y a la modificación de las expresiones OCL. Con respecto a la Ef-COM y Ef-MOD: las operaciones de colección, sus iteradores y el número de clases parecen afectar Ef-COM mientras que la longitud de las navegaciones y el número de relaciones influencia Ef-MOD. CompSub MOD, i.e. la valoración subjetiva de los sujetos acerca de la dificultad de las tareas de modificación parece estar afectada por la profundidad de las navegaciones, el número de relaciones y como están combinadas las navegaciones en las operaciones de colección.

En los experimentos UAE y UChE la valoración subjetiva realizada por los sujetos estuvo influenciada por el tiempo que ellos tardaron en comprender o modificar las expresiones OCL. En estos dos experimentos Ef-COM o Ef-MOD estuvieron fuertemente correlacionadas con CompSub COM y CompSub MOD. La razón por la cual no se obtuvieron los mismos resultados en el experimento ULME pudo deberse a la heterogeneidad de los sujetos (estudiantes de diferentes cursos y docentes de diferentes universidades), hecho que no pudimos evitar.

El efecto de aprendizaje puede haber afectado a la validez interna de los resultados relacionados con las tareas de comprensión. En efecto, en las tareas de comprensión el tiempo fue más bajo a medida que transcurría el tiempo, y algunas medidas estuvieron sólo correlacionadas para algunas observaciones.

Consideramos que debemos profundizar en el estudio de tareas de modificación debido a que su corrección y eficiencia no fue suficientemente buena como en el caso de tareas de comprensión. En las tareas de modificación fue más difícil para los sujetos identificar y navegar por las relaciones que identificar las operaciones de colección que debían ser utilizadas para realizar la modificación. Más allá de los importantes resultados obtenidos, creemos que se requiere validación adicional para fortalecer los resultados acerca del efecto de acoplamiento en la comprensibilidad y facilidad de modificación de las expresiones OCL. Por tal razón nuestro objetivo inmediato es continuar avanzando y reforzando estos resultados realizando experimentos con profesionales en entornos reales y también llevando a cabo casos de estudio.

V. AGRADECIMIENTOS

Los autores agradecen a la profesora Cristina Cachero de la Universidad de Alicante y del profesor Luis Álvarez de la Universidad Austral de Chile por permitir la realización de los cursos de OCL y de los experimentos con sus alumnos.

VI. REFERENCIAS

- [1] C. Atkinson, T. Kühne, "Model-Driven Development: A Metamodeling Foundation", IEEE Software, 20(5), 2003, pp. 36- 41.

- [2] V. R. Basili, H. D. Rombach, "The TAME project: towards improvement-oriented software environments", en *IEEE Trans. on Softw. Eng.* Vol.14 N° 6, 1998, pp. 758-773.
- [3] V. R. Basili, F. Shull, F. Lanubile, "Building knowledge through families of experiments", en *IEEE Trans. on Softw. Eng.*, Vol. 25 N° 4, 1999, pp. 456-473.
- [4] L. C. Briand, L. C. Bunse, J. W. Daly, "A Controlled Experiment for evaluating Quality Guidelines on the Maintainability of Object-Oriented Designs", en *IEEE Trans. on Soft. Eng.*, Vol. 27 N° 6, 2001, pp. 513-530.
- [5] S. L. C. Briand, Y. Labiche, H. D. Yan, M. Di Penta, "A controlled Experiment on the Impact of the OCL in UML-based Maintenance", en *IEEE Int. Conference on Softw. Maintenance*, 2004.
- [6] C. Calero, M. Piattini, M. Genero, "Method for Obtaining Correct Metrics", en *Proc. of the 3rd Int. Conference on Enterprise and Information Systems (ICEIS 2001)*, 2001, pp. 779-784.
- [7] M. Ciolkowski, F. Shull, S. Biffl, "A Family of Experiments to Investigate the Influence of Context on The Effect of Inspection Techniques" en *Proc. of the 6th Int. Conference on Empirical Assessment in Softw. Eng. (EASE)*, 2002, pp. 48-60.
- [8] S. Cook, A. Kleepe, R. Mitchell, B. Rumpe, J. Warmer, A. Wills, "The Amsterdam Manifiesto on OCL", en *Advances in Object Modelling with the OCL*, Springer, Berlin, LNCS 2263, 2001, pp. 115-149.
- [9] N. E. Fenton, S. L. Pfleeger, "Software Metrics: A Rigorous and Practical Approach". Chapman & Hall, London, 2nd Edition. International Thomson Publishing Inc. 1997.
- [10] M. Giese, R. Haldal, "From Informal to Formal Specification in UML". UML 2004, LNCS 3273, pp. 197-211, 2004.
- [11] B. Kitchenham, S. Pfleeger, N. Fenton, "Towards a Framework for Software Measurement Validation". *IEEE Trans. on Softw. Eng.*, Vol. 21 N° 12, 1995, pp. 929-944.
- [12] J. Miller, "Applying Meta-Analytical Procedures to Software Engineering Experiments", *Journal of Systems and Software*, 54, 2000, pp. 29-39.
- [13] Object Management Group. UML 2.0 OCL 2nd revised submission. OMG Document. Available at <http://www.omg.org>
- [14] Object Management Group. UML Specification Version 1.5, OMG Document. Available at <http://www.omg.org>
- [15] Object Management Group. MDA The OMG Model Driven Architecture. Available: <http://www.omg.org/mda/>, 2002.
- [16] L. Reynoso, M. Genero, M. Piattini, "Validating OCL metrics through a Family of Experiments", en *Jornadas de Ingeniería del Software y Base de Datos*. 2004.
- [17] L. Reynoso, M. Genero, M. Piattini, "Measuring OCL Expressions: An approach based on Cognitive Techniques". Imperial College Press, UK. 2005.
- [18] N. F. Schneidewind, "Methodology for Validating Software Metrics", en *IEEE Trans. on Softw. Eng.*, Vol. 18 N° 5, 1992, pp. 410-422.
- [19] B. Selic, "The Pragmatics of Model-Driven Development", en *IEEE Software*. Vol. 20, N°5, 2003, pp 19-25.
- [20] SPSS, 2002 SPSS 11.5. "Syntax Reference Guide". Chicago. SPSS Inc. 2002.
- [21] J. Warmer, A. Kleppe, "The Object Constraint Language". Second Edition. Getting Your Models Ready for MDA. Addison-Wesley, Massachusetts, 2003.
- [22] C. Wohlin, P. Runeson, M. Höst, M. Ohlson, B. Regnell, A. Wesslén, "Experimentation in Softw. Eng.: An Introduction", Kluwer Academic Publishers, 2000.

VII. BIOGRAFÍAS



Luis Reynoso es Licenciado en Ciencias de la Computación por la Universidad Nacional del Sur, Argentina, desde 1993; Master en Ciencias de Computación por la misma universidad en 2003. Es profesor asociado en la Universidad Nacional del Comahue en Neuquén, Argentina. Estudiante de doctorado en la Universidad de Castilla La Mancha en Ciudad Real, España, desde el año 2002. Actualmente esta llevando a cabo su tesis doctoral.

Becario del Instituto Internacional de Tecnología de Software, de la

Universidad de Naciones Unidas, en Macao, durante 6 meses, en el año 2000. Sus líneas de investigación se centran en la utilización de OCL, medidas orientadas a objetos y la combinación de métodos formales e informales aplicados a la ingeniería del software. Actualmente goza de una beca otorgada por un acuerdo estratégico entre la Provincia de Neuquén (Argentina) y la empresa Repsol-YPF. Ha publicado artículos en diversas conferencias tanto internacionales como nacionales (METRICS, ISESE, JISBD, etc.).



Marcela Genero es Licenciada en Ciencias de la Computación por la Universidad Nacional del Sur, Argentina, desde 1989. Doctora en Informática por la Universidad de Castilla La Mancha, Ciudad Real, España desde 2002. Actualmente es profesora asociada en la Universidad de Castilla La Mancha. Sus principales líneas de investigación son: diseño avanzado de base de datos, medidas de software, calidad de modelos conceptuales, calidad de bases de datos. Ha publicado diversos artículos en prestigiosas conferencias y revistas (CaISE, E/R, OOIS, METRICS, ISESE, SEKE, Journal of Systems and Software, Data and Knowledge Engineering, International Journal of Software Engineering and Knowledge Engineering, Information and Software Technology, Software Quality Journal, etc.). Además es coeditora de los libros "Information and Database quality", 2002, Kluwer Academic Publisher, USA y "Metrics for Software Conceptual Models", 2005, Imperial College Press, Reino Unido.



Mario Piattini es Ingeniero y Doctor en Informática por la Universidad Politécnica de Madrid (en los años 1989 y 1994 respectivamente). Auditor Certificado de Sistemas de Información y Gerente Certificado de Seguridad de la Información por la ISACA (Information System Audit and Control Association). Catedrático en el Departamento Tecnologías y Sistemas de Información de la Universidad de Castilla La Mancha, en Ciudad Real, España desde el año 2002. Sus principales líneas de investigación son: diseño avanzado de base de datos, calidad de bases de datos, medidas de software, mantenimiento de software y seguridad en sistemas de información. Autor de diversos libros y artículos sobre bases de datos, Ingeniería de Software y Sistemas de Información. Dirige el grupo de investigación ALARCOS de la Universidad de Castilla La Mancha. Ha coeditado diversos libros: "Advances Databases: Technology and Design", 2000, Artech House, Reino Unido; "Auditing Information Systems", Idea Group Publishing, 200, USA; "Information and database quality", 2002, Kluwer Academic Publishers, USA y "Metrics for Software Conceptual Models", 2004, Imperial College Press, Reino Unido.



María Esperanza Manso está licenciada en Ciencias Matemáticas por la Universidad de Valladolid (España) desde 1977. Es profesora titular de escuela universitaria en la Universidad de Valladolid desde 1991. Su investigación se centra principalmente en la experimentación en reutilización, reingeniería y mantenimiento del software. Actualmente está desarrollando su tesis doctoral. Ha publicado diversos artículos en prestigiosas conferencias tanto nacionales como internacionales (JISBD, OOIS, CaISE, METRICS, ISESE, etc.) y como capítulos de libros.