

# Una Propuesta para Especificar Cambios de Replicaciones de Experimentos en Ingeniería del Software

Margarita Cruz, Beatriz Bernárdez, Amador Durán

Dept. Lenguajes y Sistemas Informáticos, University of Seville,  
Avda. Reina Mercedes s/n, 41012, Seville, Spain.

{cruz,beat,amador}@us.es

**Resumen** *Contexto:* La replicación de estudios empíricos en Ingeniería del Software es necesaria para consolidar el conocimiento adquirido. No obstante, para incrementar el conocimiento que se genera mediante la replicación, es necesario que la información se publique de forma que permita una comprensión profunda del estudio. *Objetivo:* Al diseñar una replicación, habitualmente surge la necesidad de introducir cambios. El objetivo de este trabajo es facilitar la especificación de dichos cambios proponiendo una plantilla que permita definirlos sistemáticamente y documentarlos de forma homogénea. *Método:* Se ha definido el metamodelo para formalizar la información sobre replicaciones y cambios que son relevantes para la plantilla propuesta. Posteriormente, se ha detallado la plantilla y se ha aplicado a una replicación concreta. *Resultados y Conclusiones:* La aplicación de la plantilla a una replicación concreta ha sido satisfactoria pero se debe aplicar a otras replicaciones o familias de experimentos para su validación y mejora a través de la retroalimentación obtenida.

**Keywords:** replicación, plantilla, patrones lingüísticos, estudio empírico, ingeniería del software empírica

## 1. Introducción

La Ingeniería del Software (IS) empírica permite la evaluación de nuevos métodos, técnicas y herramientas para conocer la conveniencia de usarlos en el proceso de desarrollo. Una vez evaluado empíricamente un artefacto por primera vez, el estudio necesita ser replicado en diferentes contextos y condiciones. No solo para consolidar el conocimiento adquirido, sino también para saber si sus resultados son generalizables [3].

A pesar de la importancia de las replicaciones y aunque en los últimos años ha aumentado su práctica [10], el número de replicaciones en IS sigue siendo bajo [27]. A la hora de identificar las causas de esta situación, son varios los problemas encontrados. Por un lado, la falta de unicidad de criterios para reportar las replicaciones [9]. Por otro lado, el conocimiento tácito, ya que el investigador tiene conocimiento que no explicita en la publicación [24]. Además, se da una

carencia o incompletitud de los llamados *laboratory packages* que son necesarios para poder realizar replications [27]. A todo esto se une el esfuerzo y los recursos necesarios para llevar a cabo un experimento [10].

En otras áreas de conocimiento distintas a la IS como Psicología, esta situación se conoce como *crisis de replicabilidad* [7]. Incluye además problemas con el uso de la estadística, gran cantidad de falsos positivos publicados y la falta de rigor debido a los sesgos individuales en la recogida de evidencias. Para atenuar esta crisis, en [7] se sugiere ser rigurosos en los aspectos metodológicos y una adecuada transparencia en la transmisión de los resultados de la investigación.

Según [14], es recomendable que las primeras replications del experimento original sean realizadas por los mismos experimentadores (replications internas), para evitar la influencia de variables que puedan intervenir al cambiar de sitio o de experimentadores. Posteriormente, es valioso realizar replications externas ya que permiten conocer el rango de condiciones bajo las cuales se mantienen los resultados y demuestran que éstos son independientes de las condiciones del estudio original [25,8].

Tanto si la replicación es interna como externa, es posible la necesidad de incorporar cambios en el experimento original por varios motivos, por ejemplo, ampliar los resultados del estudio inicial [25] o adaptar un experimento a otro entorno distinto del original [3]. Dentro de la documentación a publicar sobre las replications, las guías de Carver [9] ponen de manifiesto la necesidad de describir dichos cambios mediante una sección concreta de cambios al experimento original. Para cada cambio con respecto al experimento original es necesario registrar en qué consiste, la situación que ha causado dicho cambio y determinar su impacto en las amenazas a la validez de la replicación. Además, conocer a fondo los cambios facilita ver si se cumple el principio fundamental de independencia base del análisis agregado de datos según Kitchenham [18], que promueve la introducción de cambios en las replications para evitar propagar problemas del experimento original.

En este trabajo se propone un enfoque para la especificación de los cambios en las replications basado en: i) un metamodelo para formalizar la información, ii) una plantilla que facilita la reusabilidad, la representación visual y evita redundancias y iii) patrones lingüísticos que facilitan la redacción. El uso de la plantilla propuesta persigue una doble finalidad, por una lado, invita al investigador a hacer explícito el cambio y sus detalles (reduciendo el conocimiento tácito) y, por otro, ayuda al lector a comprender mejor la replicación y a seguir la traza de cómo evoluciona el experimento original en la sucesión de experimentos de una familia.

El resto del artículo está organizado de la siguiente forma: en la sección 2 se presenta el metamodelo en el que se sustenta la plantilla; en la sección 3 se propone la plantilla para la especificación de los cambios de las replications; los trabajos relacionados se discuten en la sección 4 y finalmente, en la sección 5 resumimos nuestras conclusiones y las perspectivas para trabajos futuros.

## 2. Metamodelo sobre replications y cambios

En esta sección se presenta el metamodelo en el que se basa la plantilla (Fig. 1) usando la notación de diagramas de clases UML.

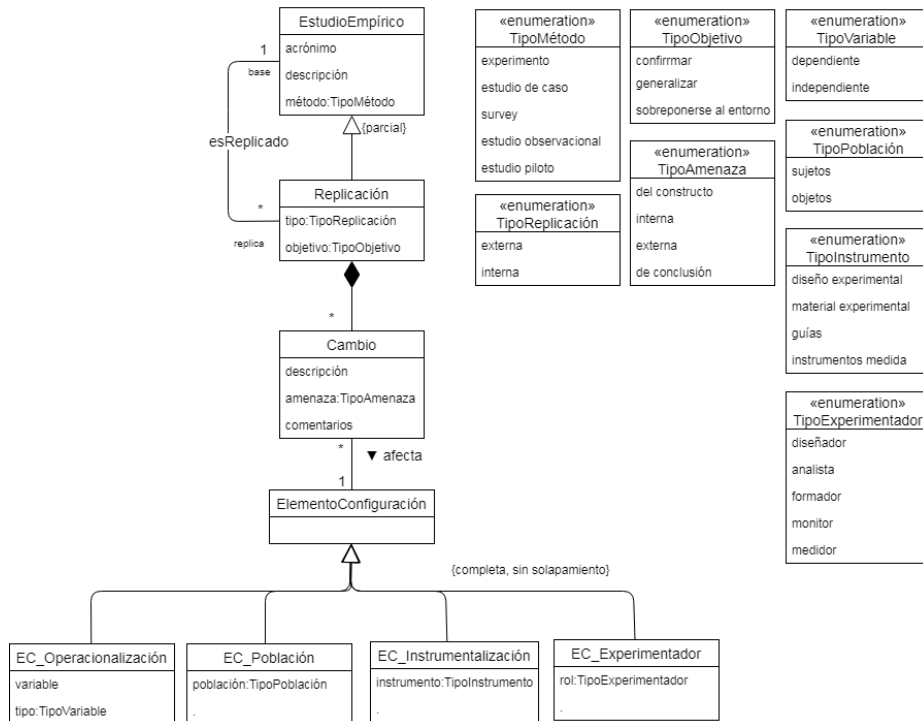


Figura 1. Diagrama de clases UML

En el modelo, una replicación es un tipo de estudio empírico, cuya base es otro estudio empírico (bien el original o bien otra replicación). La herencia es parcial ya que los objetos que son instancias del “experimento original” se consideran en el supertipo. Entre los atributos de la clase Replicación, aparecen:

- **Acrónimo.** Se utilizará un código ó acrónimo relativo al experimento de referencia.
- **Método empírico.** Los tres principales métodos empíricos son *experimentos controlados*, *surveys* y *estudios de caso*. El método empírico *cuasi experimentos*, en el que la asignación de sujetos a tratamientos no es aleatoria, se incluyen en los *experimentos controlados* [30]. Aunque menos frecuentes, consideramos además, otros dos métodos empíricos: *estudios piloto* y *estudios observacionales*.

- **Tipo de replicación.** Atendiendo a quién hace la replicación, éstas se clasifican en *internas* (son realizadas por los experimentadores originales) y *externas* (son llevadas a cabo por experimentadores independientes y por tanto su poder de confirmación es mayor que en las internas) [8]. Existen otras taxonomías atendiendo a diversos criterios, por ejemplo una posible clasificación es según el grado en que se sigue el procedimiento del experimento original. En este aspecto, hay autores que hablan de replications *exact* y *conceptual* [25], *closed* y *differentiated* [1] [19] [17]. Basili et al. [5] se refieren a replications *strict* y presentan una clasificación en tres grupos principales según se varíe o no la hipótesis de investigación o se amplíe la teoría. En [15] se propone una clasificación de las replications en *literal*, *operational* y *conceptual* dependiendo de los cambios llevados a cabo y del propósito de la replicación. Debido a la falta de acuerdo en la terminología [15] y los matices diferenciadores, hemos optado por describir los cambios catalogando las replications como *internas* ó *externas* sin entrar en otras categorías más confusas.
- **Objetivo.** Para el objetivo de la replicación, se ha definido un tipo enumerado que puede tomar tres posibles valores: i) *confirmar resultados*, se trata de verificar los resultados obtenidos, ii) *generalizar los resultados* incluye, entre otros, cambios en los experimentadores para analizar si los resultados son consecuencia de la intervención de dichos experimentadores, cambios en la población y métricas para verificar si se siguen cumpliendo los resultados o que la hipótesis sigue siendo válida, iii) *sobreponerse a ciertas limitaciones del entorno del experimento original*, recoge los cambios en las replications que se producen como consecuencia de las limitaciones impuestas y a las que es necesario adaptar el experimento pero que deben de quedar registrados.

Para reflejar la relación que hay entre la clase replicación y la clase que describe sus cambios se ha utilizado una composición ya que los cambios son parte intrínseca de la replicación. Sus atributos son:

- **Descripción.** Permitirá describir en qué consiste el cambio.
- **Amenaza a la validez abordada.** Es un tipo enumerado con los valores: *incrementa la validez del constructo*, *externa*, *interna* y de *conclusión*. Este atributo sirve para indicar cómo el cambio introducido mitiga la amenaza a la validez de un determinado tipo.
- **Comentarios.** Información adicional sobre el cambio.

Cada cambio se lleva a cabo sobre un elemento de la configuración experimental. En [15] se identifican los elementos de la configuración experimental que pueden ser cambiados en una replicación. Dichos elementos se agrupan en dimensiones. En el diagrama de clases, se ha utilizado una jerarquía para clasificar los elementos configuracionales en las cuatro dimensiones. La clasificación es *completa* y *sin solapamiento*.

- **Dimensión y elementos.** Las dimensiones identificadas en [15] y que en el metamodelo se representan mediante los subtipos son:

- *Operacionalización*: engloba a los cambios relacionados con las *variables dependientes* (cambios en métricas y procedimientos de medición) e *independientes* (cambios en la forma de aplicar el tratamiento). Se corresponde con el subtipo EC\_Operacionalización con los atributos: variable (nombre de la variable a modificar) y tipo que puede tomar los valores *dependiente* o *independiente*.
- *Población*: incluye los cambios en las propiedades de los *sujetos* (ej. distinta experiencia, edad, etc.) y de los *objetos experimentales* (ej. distinto lenguaje de programación, nivel de dificultad, etc). Se corresponde con el subtipo EC\_Población con el atributo población que puede tomar los valores *sujetos* u *objetos*.
- *Instrumentalización*: incluye los cambios en el *diseño experimental*, en el *material experimental* (al utilizar distintas instancias del mismo tipo de objetos experimentales, ej. cambiar el enunciado de un problema por otro de igual dificultad), en las *guías* (ej. cambiar las instrucciones proporcionadas a los sujetos) y en los *instrumentos de medición* (ej. cambiar el cuestionario para recogida de datos). Se corresponde con el subtipo EC\_Instrumentalización con el atributo instrumento que puede tomar los valores *diseño experimental*, *material experimental*, *guías* o *instrumentos medida*.
- *Experimentadores*: referido a los cambios en los roles de los experimentadores. Se corresponde con el subtipo EC\_Experimentador con el atributo rol que puede tomar los valores *diseñador*, *analista*, *formador*, *monitor* o *medidor*.

### 3. Plantilla para cambios en replicaciones

En esta sección se propone una plantilla para facilitar la especificación de los cambios de las replicaciones basada en el metamodelo descrito en la sección anterior. El uso de plantillas ayuda a estructurar la información en una forma fija, reduce la ambigüedad, promueve la reutilización y también sirve de guía para evitar que falte información relevante [12]. Para algunos campos de las plantillas se han identificado frases que son habituales y que se han parametrizado. Estas frases se denominadas patrones lingüísticos, o abreviadamente patrones-L [28].

En la notación usada para describir los patrones-L, las palabras o frases entre  $\langle \text{ y } \rangle$  deben ser convenientemente reemplazadas, las palabras o frases que se encuentren entre  $\{ \text{ y } \}$  y separadas por  $|$  representan opciones de las que se debe escoger una y las palabras entre  $[ \text{ y } ]$  son opcionales, es decir, pueden aparecer o no cuando está instanciada la plantilla.

En el cuadro 1 se muestra la plantilla propuesta. El significado de los campos que la componen es el siguiente:

- **Acrónimo de la replicación**: Con objeto de conseguir una rápida identificación de los distintos experimentos de una familia, es conveniente que la descripción del conjunto de replicaciones del mismo experimento comience

**Cuadro 1.** Plantilla para especificar cambios en una replicación

< <i>Acrónimo</i> > #< <i>n</i> >	Replicación del experimento < <i>Acrónimo experimento</i> >#< <i>m</i> >
Método empírico	{Experimento controlado   Survey   Estudio de caso   Estudio observacional   Estudio piloto }
Tipo de replicación	{Interna   Externa }
Objetivo	{Confirmar resultados   Generalizar los resultados   Sobreponerse a ciertas limitaciones del entorno del experimento original }
Cambio #< <i>i.j</i> >	Originalmente, < <i>descripción de la situación en el experimento original.</i> > En la replicación < <i>descripción de la situación en la replicación</i> > {con el fin de   debido a } < <i>causa del cambio.</i> >
[Dimensión modificada]	{Operacionalización [(, en concreto, de la variable {dependiente   independiente} < <i>nombre de la variable</i> >)]   Población[(, en concreto, los {sujetos   objetos} experimentales)]   Instrumentalización [(, en concreto, {diseño experimental   material experimental   guías   instrumentos de medida} )]   Experimentadores [(, en concreto, {diseñador   analista   formador   monitor   medidor})] }
Amenaza abordada	El cambio incrementa la validez {del constructo   externa   interna   de conclusión }
[Comentarios]	< <i>Comentarios</i> >

por un código ó acrónimo relativo al experimento de referencia y seguido de un número secuencial.

- **Acrónimo del experimento:** El mismo que el de la replicación. Cuando el experimento de referencia sea el experimento original, es conveniente que el número secuencial que le sigue sea el 1; de esta forma, se identificará fácilmente cuando el experimento de referencia sea ya una replicación.
- **Método empírico:** Se elegirá uno de los posibles valores: *experimento controlado*, *survey*, *estudio de caso*, *estudio piloto* o *estudio observacional*.
- **Tipo de replicación:** Se elegirá uno de los posibles valores: *interna* o *externa*.
- **Objetivo:** Permite seleccionar entre tres posibles valores: *confirmar resultados*, *generalizar los resultados* o *sobreponerse a ciertas limitaciones del entorno del experimento original*.
- **Descripción del cambio:** Los cambios se numeran secuencialmente y su descripción contiene un patrón-L que se debe completar y en el que se recoge la situación en el experimento original, el cambio en sí y la causa o consecuencia. En esta última parte del patrón, para describir la causa, se elige entre “con el fin de” o “causado por” según suene mejor en español al completar la frase.
- **Dimensión modificada y elementos:** El patrón-L se completa eligiendo una de las dimensiones propuestas; dentro de la dimensión elegida, se puede

concretar el elemento modificado eligiendo alguna de las opciones presentadas. Tanto la dimensión como el elemento afectado son opcionales. Es decir, por simplicidad, se permite describir un cambio sin tener la obligatoriedad de identificar la dimensión y el elemento al que afecta el cambio

- **Amenaza a la validez abordada:** Se elije alguna de las amenazas a la validez: *del constructo, externa, interna* y de *conclusión*. [30].
- **Comentarios.** Texto libre.

### 3.1. Aplicación de la plantilla

En el cuadro 2 se muestra el resultado de aplicar la plantilla y patrones-L para definir los cambios realizados en la replicación descrita en [6]. Se trata de una replicación interna llevada a cabo por parte de los autores del presente artículo que versa sobre el efecto de la práctica de *mindfulness* en el rendimiento de los alumnos al desarrollar modelos conceptuales.

**Cuadro 2.** Ejemplo de aplicación de la plantilla a la replicación [6]

<i>Mind #2</i>	Replicación del experimento <i>Mind #1</i>
Método empírico	Experimento controlado
Tipo de replicación	Interna
Objetivo	El objetivo de la replicación es confirmar resultados
Cambio #1	Originalmente, <i>durante 4 semanas se practicó mindfulness 4 días a la semana en sesiones de 10 minutos.</i> En la replicación <i>las sesiones fueron de 12 minutos y durante 6 semanas</i> con el fin de <i>hacer más evidentes los beneficios de Mindfulness.</i>
Dimensión modificada	Operacionalización, en concreto, la variable independiente <i>Training Workshop</i>
Amenaza abordada	El cambio incrementa la validez del constructo
Cambio #2	Originalmente, <i>la asignación de sujetos a tratamiento no era aleatoria.</i> En la replicación <i>se hace aleatoria</i> con el fin de <i>subsanan las amenazas a la validez interna de los quasi experimentos.</i>
Dimensión modificada	Instrumentalización, en concreto, diseño experimental
Amenaza abordada	El cambio incrementa la validez interna
Cambio #3	Originalmente, <i>se impartió un taller de oratoria al grupo de control a modo de placebo.</i> En la replicación <i>el taller de oratoria se llevó a cabo después del experimento</i> con el fin de <i>evitar un posible efecto de dicho taller en las mediciones de las variables dependientes.</i>
Dimensión modificada	Operacionalización, en concreto, la variable independiente <i>Training Workshop</i>
Amenaza abordada	El cambio incrementa la validez del constructo

Aunque es evidente que la validación de la plantilla necesitaría aplicarse a otras replications, con su estructura actual hemos podido hacer explícitos los cambios de la replicación citada.

#### 4. Trabajos relacionados

En esta sección, se analizará trabajos relacionados ya sea por i) presentar directrices sobre como reportar cambios en replications en IS; ii) presentar dichos cambios de forma que sean fácilmente interpretados, por ej. mediante tablas; y iii) definir o utilizar plantillas o patrones.

Carver [9] presenta una propuesta inicial de directrices sobre el contenido de las publicaciones que reportan replications. En dicho artículo, se indica el contenido que debería ser incluido sobre el estudio original, sobre la replicación y la comparación de ambos. Entre la información sobre la replicación se incluye un apartado de cambios al experimento original pero sin concretar cómo reportar dichos cambios.

Aparte de la propuesta metodológica de Carver, se han encontrado varios estudios que documentan los cambios de las replications haciendo uso de tablas *ad-hoc*. Entre ellos destacaremos los siguientes:

Solari [26] se refiere a identificar los incidentes experimentales que ocurren durante la replicación. Ejemplos de incidentes son cambios imprevistos en cualquiera de las variables definidas para el experimento. En el estudio, se documentan cinco replications del mismo experimento realizadas por diferentes experimentadores en diferentes lugares. Las incidencias se clasifican por replicación y categoría de la incidencia. Con la información almacenada se presenta una tabla resumen que facilita el análisis y la comparación de las replications.

El trabajo de Lung et al. [20], es el más relacionado con nuestra propuesta. Se documenta una replicación de un experimento con sujetos humanos presentando una tabla resumen de diferencias y cambios realizados para adaptar el procedimiento a las circunstancias locales. Para cada cambio introducido se explica claramente la situación en el experimento original, la situación en la replicación y las razones de hacer dicho cambio.

En [1], Almqvist presenta una revisión bibliográfica de replications. Incluye una plantilla para recoger la configuración de las replications y en una de las filas describe de forma conjunta todos los cambios llevados a cabo en la replicación.

En [2], Apa et al. presentan una replicación sobre un experimento de detección de fallos. Los cambios se describen, en una sección de cambios en el experimento original, en forma textual y en una tabla donde para cada cambio se compara la situación en el experimento original y en la replicación.

En [13], Fucci et al. documentan una replicación externa de un experimento sobre el desarrollo basado en pruebas. Contiene una sección de cambios en la configuración original donde presenta una tabla con los ajustes hechos al experimento base.

En [16], Juristo et al. documentan ocho replications de un experimento para evaluar la efectividad de tres técnicas de verificación y validación. Se presenta



una tabla donde para cada cambio, se analiza su situación en cada replicación y la decisión de diseño adoptada.

En [22], Quesada et al. describen un experimento controlado sobre análisis de puntos de función y sus dos replicaciones. Las diferencias en la configuración experimental entre el experimento y replicaciones se explican mediante una tabla y de forma textual.

Por otra parte, están los estudios sobre definición de plantillas en otras áreas dentro de (IS). La idea de utilizar plantillas y patrones para definir los cambios en las replicaciones de experimentos, está basada en las plantillas propuestas por Durán et al. [12] para la elicitación de requisitos. Las plantillas y patrones-L se han aplicado con éxito en diversas áreas; Del Río et al. [11] las han utilizado para la definición de indicadores de rendimiento de procesos de negocio y Segura et al. [23] para descripción de relaciones metamórficas.

En la descripción de experimentos, destacar la plantilla Goal–Question–Metric (GQM) de Basili [4], recomendada por Wohlin [30], para la definición de objetivos de los experimentos. Por último, en la metodología Design Science Research (DSR), Wieringa [29] define una plantilla para la especificación de un problema que conducirá una investigación futura.

**Cuadro 3.** Comparación de trabajos relacionados

Trabajos relac.	Tema	(IS)Empírica	Uso plantilla	Uso patrón
Carver [9]	Guías	✓	✗	✗
Solari [26]	Documenta replicación	✓	✗	✗
Lung et al. [20]	Documenta replicación	✓	✗	✗
Almqvist [1]	SLR replicaciones	✓	✗	✗
Apa et al. [2]	Documenta replicación	✓	✗	✗
Fucci et al. [13]	Documenta replicación	✓	✗	✗
Juristo et al. [16]	Documenta replicación	✓	✗	✗
Quesada et al. [22]	Documenta replicación	✓	✗	✗
Durán et al. [12]	Elicitación requisitos	✗	✓	✓
Del Río et al. [11]	Indicadores rendimiento	✗	✓	✓
Segura et al. [23]	Relaciones metamórficas	✗	✗	✓
Basili [4]	GQM	✓	✗	✓
Wieringa [29]	DSR	✗	✗	✓
Este trabajo	Plantillas cambios	✓	✓	✓

En el cuadro 3 se analiza, para cada uno de los trabajos anteriormente citados: i) el tema sobre el que versan, ii) su pertenencia al área de IS Empírica, iii) la utilización de plantillas para facilitar la reutilización y presentación visual y iv) la utilización de patrones para facilitar la redacción.

## 5. Conclusiones y trabajos futuros

En este trabajo se presenta un metamodelo y una plantilla junto con los patrones lingüísticos para facilitar la especificación de los cambios de las replications. La estructuración de la información en forma de plantilla y la propuesta de frases estándar, pensamos que puede facilitar la redacción de los cambios en las replications.

El desarrollo de la plantilla se ha ilustrado con su aplicación a una replicación concreta. A corto plazo, los trabajos futuros irán encaminados a desarrollar un formulario web o bien un asistente para que los experimentadores puedan definir los cambios amigablemente y con dicha información se genere, por ejemplo, el código `LATEX` para insertar en los artículos. Por otro lado, se validará la plantilla aplicándola a varias familias de experimentos e introduciendo las mejoras pertinentes. Más a largo plazo, nuestra intención es extender el repositorio de información experimental EXEMPLAR [21] con aspectos de replications basándonos en el metamodelo y plantilla propuestos.

## Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos de investigación BELi (TIN2015-70560-R) del Plan Nacional y COPAS (TIC-1867) del Plan Andaluz de Investigación.

## Referencias

1. Johan Per Fredrik Almqvist. Replication of controlled experiments in empirical software engineering - a survey, 2006. Master's thesis, Department of Computer Science, Faculty of Science, Lund University, Sweden, 2006.
2. Cecilia Apa, Oscar Dieste, et al. Effectiveness for detecting faults within and outside the scope of testing techniques: an independent replication. *Empirical Software Engineering*, 19(2):378–417, 2014.
3. Maria Teresa Baldassarre, Jeffrey Carver, Oscar Dieste, and Natalia Juristo. Replication types: Towards a shared taxonomy. In *Proceedings of EASE '14*, pages 18:1–18:4, 2014.
4. Victor R Basili, Gianluigi Caldiera, and H Dieter Rombach. Goal question metrics paradigm. *Encyclopedia of Software Engineering*, pages 528–53, 1994.
5. Victor R Basili, Forrest Shull, and Filippo Lanubile. Building knowledge through families of experiments. *IEEE Transactions on Software Engineering*, 25(4):456–473, 1999.
6. Beatriz Bernárdez, Amador Durán, José A. Parejo, and Antonio Ruiz–Cortés. An experimental replication on the effect of the practice of mindfulness in conceptual modeling performance. *Journal of Systems and Software*, 136:153–172, 2018.
7. Fernando Blanco, José César Perales López, and Miguel A Vadillo. Puede la psicología rescatarse a sí misma? incentivos, sesgos y replicabilidad, 2017.
8. Andrew Brooks, John Daly, James Miller, Marc Roper, and Murray Wood. Replication of experimental results in software engineering. *Technical Report ISERN-96-10, University of Strathclyde*, 1996.

9. Jeffrey C Carver. Towards reporting guidelines for experimental replications: a proposal. In *1st International Workshop on Replication in Empirical Software Engineering*, 2010.
10. Fabio QB Da Silva, Marcos Suassuna, A César C França, Alicia M Grubb, Tatiana B Gouveia, Cleiton VF Monteiro, and Igor Ebrahim dos Santos. Replication of empirical studies in software engineering research: a systematic mapping study. *Empirical Software Engineering*, 19(3):501–557, 2014.
11. Adela del Río-Ortega, Manuel Resinas, Amador Durán, and Ruiz-Cortés Antonio. Defining process performance indicators by using templates and patterns. In *Proceedings of BPM'12*, pages 223–228. Springer, 2012.
12. Amador Durán Toro, Beatriz Bernárdez Jiménez, Antonio Ruiz Cortés, and Miguel Toro Bonilla. A requirements elicitation approach based in templates and patterns. In *Proceedings of WER'99*, pages 17–29, 1999.
13. Davide Fucci, Giuseppe Scanniello, Simone Romano, Martin Shepperd, Boyce Sigweni, Fernando Uyaguari, Burak Turhan, Natalia Juristo, and Markku Oivo. An external replication on the effects of test-driven development using a multi-site blind analysis approach. In *Proceedings of EMSE'16*. ACM, 2016.
14. Omar S Gómez, Natalia Juristo, and Sira Vegas. Replications types in experimental disciplines. In *Proceedings of ESEM'10*. ACM, 2010.
15. Omar S Gómez, Natalia Juristo, and Sira Vegas. Understanding replication of experiments in software engineering: A classification. *Information and Software Technology*, 56(8):1033–1048, 2014.
16. N Juristo, Sira Vegas, Martín Solari, Silvia Abrahao, and Isabel Ramos. Comparing the effectiveness of equivalence partitioning, branch testing and code reading by stepwise abstraction applied by subjects. In *2012 IEEE Fifth International Conference on Software Testing, Verification and Validation*, pages 330–339. IEEE, 2012.
17. Natalia Juristo and Sira Vegas. The role of non-exact replications in software engineering experiments. *Empirical Software Engineering*, 16(3):295–324, 2011.
18. Barbara Kitchenham. The role of replications in empirical software engineering—a word of warning. *Empirical Software Engineering*, 13(2):219–221, 2008.
19. R Murray Lindsay and Andrew SC Ehrenberg. The design of replicated studies. *The American Statistician*, 47(3):217–228, 1993.
20. Jonathan Lung, Jorge Aranda, Steve Easterbrook, and Gregory Wilson. On the difficulty of replicating human subjects studies in software engineering. In *Proceedings of ICSE'08*, pages 191–200. IEEE, 2008.
21. José Antonio Parejo, S. Segura, P. Fernandez, and Antonio Ruiz-Cortés. Exemplar: An experimental information repository for software engineering research. In *Actas de JISBD'2014*, pages 155–159, 2014.
22. Christian Quesada-López, Denisse Madrigal, and Marcelo Jenkins. An empirical evaluation of automated function points. In *Ibero-American Conference on Software Engineering*, pages 151–165, 2016.
23. Sergio Segura, Amador Durán, Javier Troya, and Antonio Ruiz Cortés. A template-based approach to describing metamorphic relations. In *Metamorphic Testing (MET), 2017 IEEE/ACM 2nd International Workshop on*, pages 3–9. IEEE, 2017.
24. Forrest Shull, Victor Basili, Jeffrey Carver, José Carlos Maldonado, Guilherme Horta Travassos, Manoel Mendonça, and Sandra Fabbri. Replicating software engineering experiments: addressing the tacit knowledge problem. In *Proceedings of EMSE'2002*, pages 7–16. IEEE, 2002.

25. Forrest J Shull, Jeffrey C Carver, Sira Vegas, and Natalia Juristo. The role of replications in empirical software engineering. *Empirical Software Engineering*, 13(2):211–218, 2008.
26. Martín Solari. Identifying experimental incidents in software engineering replications. In *Proceedings of EMSE'2013*, pages 213–222. IEEE, 2013.
27. Martín Solari, Sira Vegas, and Natalia Juristo. Content and structure of laboratory packages for software engineering experiments. *Information and Software Technology*, 2017. In Press.
28. Amador Durán Toro and Beatriz Bernárdez Jiménez. Metodología para la elicitación de requisitos de sistemas software. *Informe Técnico LSI-2000-10. Facultad de Informática y Estadística Universidad de Sevilla*, 2000.
29. Roelf J. Wieringa. *Design science methodology for information systems and software engineering*. Springer, 2014.
30. Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. *Experimentation in software engineering: an introduction*. Springer Publishing Company, Incorporated, 2012.