

Comparación de Técnicas de Lectura de Documentos de Requisitos de Software: Diseño de un Experimento

Comparison of Software Requirement Reading Techniques: Design of a Experiment

María Inés Lund, Esp., Myriam Herrera, MSc., Laura Aballay, Prog., Sergio Zapata, Esp.
Instituto de Informática, Facultad de Ciencias Exactas Físicas y Naturales
Universidad Nacional de San Juan, Argentina
{mlund, mherrera, laballay, szapata}@iinfo.unsj.edu.ar

Recibido para revisión 26 de Marzo de 2007, Aceptado 03 de Diciembre, Versión final 10 de Diciembre de 2007

Resumen—Dentro de la Ingeniería de Software, el proceso de especificación de requisitos es de gran relevancia. Alcanzar la satisfacción del cliente/usuario del software depende, en buena medida, de la calidad del documento de requisitos de software. Fundándose en esta premisa, las prácticas de evaluación de documentos de requisitos, específicamente las inspecciones de software, son herramientas valiosas para detectar defectos y mejorar su calidad. En este trabajo se presenta una propuesta de diseño experimental que compara dos técnicas de lectura de documentos de requisitos. El diseño experimental propuesto está basado en los principios de la Ingeniería de Software Experimental, los cuales promueven la aceptación de conclusiones sólo una vez que se verifiquen una serie de pruebas experimentales que confirmen su veracidad. El experimento se realizó en el ámbito de la Universidad Nacional de San Juan (Argentina) con la participación de alumnos avanzados y graduados recientes en carreras de informática.

Palabras Clave—Técnicas de Lectura, Inspecciones de Software, Requisitos de Software, Ingeniería de Software Experimental.

Abstract—For a software tool, the target of reaching customer/user satisfaction depends to a large extent on the quality of the software-requirement document. On this premise, the evaluation of software requirement documents - specifically software inspection - is a valuable tool for detecting defects and improve the software quality. The current work presents an experimental design proposal that compares two techniques for reading software requirement documents. The proposed experimental design is based on the Experimental Software Engineering principles, which only promote the acceptance of conclusions once a series of experimental tests has been done. The experiment was made in academic settings of the Universidad Nacional de San Juan, Argentina, with the participation of undergraduate and graduate students of Computer Science careers.

Keywords—Reading Techniques, Software Inspections, Software Requirements, Experimental Software Engineering.

I. INTRODUCCIÓN

A ctualmente se hacen grandes esfuerzos, tanto en la academia como en la industria, para mejorar la calidad del software, atendiendo a una necesidad concreta reclamada por la sociedad. La falta de calidad se puede reflejar, entre otros aspectos, en productos de software que no cumplen con lo especificado por el interesado del mismo.

En un documento de especificación de requisitos (DER) se modela lo que el interesado desea que efectivamente realice el software. Alcanzar la satisfacción del interesado de software depende, en buena medida, de la calidad del documento de requisitos de software.

Fundándose en esta premisa, las prácticas de evaluación de documentos de requisitos, específicamente las inspecciones de software, son herramientas valiosas para detectar defectos y mejorar su calidad [15], [16], [17].

Las inspecciones de software [2] son técnicas, cada vez más extendidas y definidas, que ayudan a mejorar la calidad aplicándose a cualquier subproducto de software que resulte durante el proceso de desarrollo. Consecuentemente, realizar inspecciones de los documentos de especificación de requisitos es una acción que favorece a la calidad anhelada.

Si no se detectan tempranamente errores cometidos en la etapa de definición de requisitos, el alto costo de remoción de defectos en etapas posteriores impactará fuerte y negativamente en la ejecución del proyecto de software, tanto en el cumplimiento de cronogramas y presupuestos como en la satisfacción del interesado [2].

La técnica de inspección de software, propuesta por Fagan

[7], consiste en un proceso de seis pasos y roles de inspectores bien definidos:

1. Planificación: Cuando el ingeniero de software completa un producto de trabajo, se conforma un grupo de inspección y se designa un moderador. El moderador controla que el producto de trabajo satisfaga los criterios de entrada para la inspección, asigna diferentes roles a las personas que integran el grupo de inspección, define los recursos necesarios para la inspección y hace una planificación de actividades a realizar para llevar a cabo el proceso de inspección en su totalidad.

2. Overview: Si los inspectores no están familiarizados con el producto de trabajo a inspeccionar, el autor del mismo presenta el producto en una reunión introductoria. Este paso es opcional.

3. Preparación Individual: Los inspectores deben estudiar los productos de trabajo y el material relacionado, antes de la reunión de inspección. Es aconsejable la utilización de listas de chequeos para ayudar a encontrar defectos comunes. El tiempo de ejecución de esta etapa dependerá fuertemente de la experiencia que tenga el inspector.

4. Reunión de Inspección: Es la etapa en donde se evalúa el producto. El moderador debe asegurarse que todos los inspectores están suficientemente preparados. La persona designada en el rol de lector presenta el producto de trabajo, mientras que cada inspector identifica explícitamente posibles defectos. Al terminar con la reunión, el grupo determina si el producto es aceptado o debe ser retrabajado para una posterior inspección.

5. Retrabajo: El autor corrige todos los defectos encontrados en el producto por los inspectores.

6. Seguimiento: El moderador chequea las correcciones realizadas del autor. Si el moderador está satisfecho, la inspección se completa formalmente.

Como se observa en el proceso de inspección planteado, una etapa importante es la denominada *preparación individual*. Esta etapa consiste en el análisis individual, por parte de técnicos que cumplen el rol de inspectores, del documento a inspeccionar con el fin de encontrar posibles defectos. En esta etapa es aconsejable aplicar alguna técnica de lectura [11] que ayude a encontrar defectos en los documentos inspeccionados, siguiendo un conjunto de pautas preestablecidas.

Es justamente en esta etapa del proceso de inspección en donde el actual trabajo pretende realizar un aporte, al presentar un diseño experimental [12] para comparar dos técnicas de lecturas aplicables a documentos de requisitos de software.

El diseño experimental propuesto está basado en los principios de la Ingeniería de Software Experimental [14], los cuales promueven la aceptación de conclusiones sólo una vez que se verifiquen una serie de pruebas experimentales que confirmen su veracidad. Se realizó el diseño experimental en el ámbito de la Universidad Nacional de San Juan (Argentina) con la participación de alumnos avanzados y

graduados recientes en carreras de informática.

Este artículo esta organizado de la siguiente manera: en la sección 2 se brinda una descripción de las Técnicas de Lectura, exponiendo su objetivo y área de aplicación. Especialmente de las técnicas de Lectura Ad Hoc y de Lectura Basada en Perspectivas (PBR), que son las usadas en el experimento. En la Sección 3 se realiza una descripción del Proceso Experimental, en términos del modelo teórico de Wohlin *et al.* [12]. También en esta sección se presentan los resultados del experimento. Para finalizar en la sección 4 se presentan las conclusiones del trabajo.

II. TÉCNICAS DE LECTURA

Las técnicas de lectura son pautas o guías que tienen como objetivo encontrar defectos en los artefactos de software que están en proceso de inspección. Estas guías, generalmente basadas en un mecanismo de checklist, apuntan a obtener un conocimiento acabado del artefacto de software que permita al inspector encontrar hasta los defectos más sutiles [9].

En la literatura especializada se proponen varias técnicas de lectura, cada una con sus características, generalmente en función del artefacto de software al cual se aplicarán. Estas técnicas de lectura, surgen para adaptarse mejor a las particularidades de los artefactos de software (productos de trabajo) [4].

Para llevar a cabo la evaluación de calidad de los requisitos se han propuesto varias técnicas de lectura que consisten en realizar una lectura detenida de la especificación de requisitos con el objeto de encontrar irregularidades que sea necesario eliminar [4].

Se presenta en este trabajo un diseño experimental para comparar dos técnicas de lectura de documentos de requisitos de software, las técnicas de lectura *ad-hoc* y la técnica basada en perspectivas PBR (Perspective-Based Reading). La primera se la eligió por ser la más utilizada, mientras que a la segunda se la escogió por ser una de las técnicas mejor definidas y con mayores referencias bibliográficas en la actualidad.

A Técnica de Lectura Ad hoc

La técnica de Lectura Ad hoc no cuenta con procedimientos formalizados, es la que comúnmente utilizan los grupos de desarrollo y se basa simplemente en la experiencia de los inspectores, no necesitando conocer de antemano ningún tipo de proceso o secuencia de pasos para la detección de defectos.

Bajo esta técnica, el artefacto a inspeccionar se entrega a los inspectores sin ninguna indicación o guía sobre cómo proceder ni sobre qué buscar. Por eso se la conoce como técnica de Lectura sin Checklists.

El término "*ad-hoc*" sólo se refiere al hecho de no proporcionar a los inspectores de un procedimiento común para aplicar. En este caso la detección de los defectos depende completamente de las habilidades, conocimientos y experiencia

del inspector.

Normalmente, el inspector busca secuencialmente los defectos típicos del artefacto que está leyendo. Por ejemplo, si se está inspeccionando requisitos, la verificación de requisitos se adecua a las características propias de la especificación concreta y no sigue un patrón establecido a priori [4], el inspector, buscará sistemática y secuencialmente defectos de corrección, de completitud, de ambigüedad, etc.

B. Técnica de Lectura PBR

La técnica de lectura PBR (Perspective-Based Reading) [3] cuyo origen fue la Técnica de Lectura denominada Scenario-Based Reading, provee una guía de procedimientos, adaptada a los requisitos de software, expresados en lenguaje natural.

La técnica PBR destaca la heterogeneidad existente respecto a los intereses de los diferentes usuarios en cuanto a la información que brinda el documento de especificación de requisitos. Es necesario por lo tanto identificar los distintos usuarios de un artefacto de software específico. Este proceso es similar a la construcción de un modelo de casos de uso, el cual requiere identificar quién usará el sistema y de qué manera.

Basándose en la existencia de los distintos usuarios del documento de requisitos, se han desarrollado diferentes perspectivas, las cuales pueden variar de acuerdo a las necesidades de la organización o proyecto; las más frecuentes, debido a su constante uso y aplicación, son las siguientes [10]:

- Perspectiva del usuario: el revisor debe verificar si el documento de requisitos describe correctamente el conjunto de funcionalidades y restricciones que el sistema final debe satisfacer.
- Perspectiva del diseñador: en este escenario el revisor debe verificar si el documento de requisitos conforma una base correcta para el diseño del sistema.
- Perspectiva del tester: el revisor debe verificar que el documento de requisitos conforma una herramienta adecuada de comparación con el sistema final a implementar.

III. PROCESO EXPERIMENTAL

El proceso experimental propuesto por Wohlin *et al.* [12], como se observa en la Fig.1, tiene las siguientes etapas principales:

- *La definición de los objetivos del experimento*, el porqué del mismo.
- *La planificación del experimento*, que tiene como salida el diseño del experimento, y en donde se determinará el tipo de análisis estadístico que se deberá realizar.
- *La operación*, es la fase de ejecución del experimento, normalmente se lleva a cabo un entrenamiento piloto previo y luego el experimento propiamente dicho.

- *Análisis e interpretación de los resultados*, el tipo de análisis estará condicionado por la hipótesis del experimento, por el tipo de variables que se haya elegido y el propio diseño del experimento.

- A partir de las conclusiones del experimento se pueden poner en marcha nuevos estudios para confirmar los hallazgos, extender su validez o descubrir nuevo conocimiento.

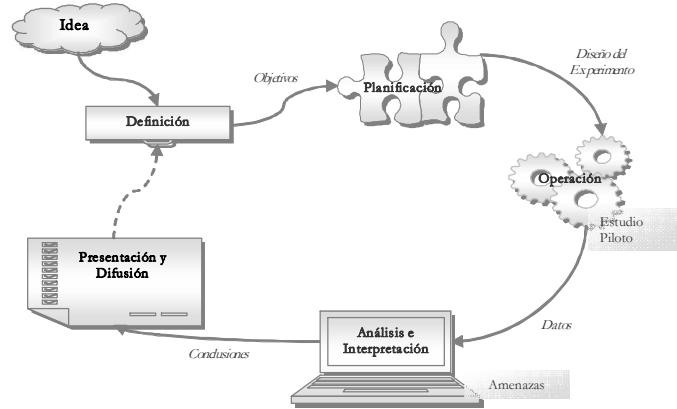


Figura. 1. Vista general del Proceso Experimental

En este trabajo se presenta el experimento completo. Siguiendo las pautas del proceso experimental de Wohlin *et al.* [12].

A. Definición

En esta etapa se determina el fundamento del experimento. Se parte de una idea experimental y se obtiene como resultado la definición del experimento.

Tomando como pauta de trabajo la plantilla *GQM* (Goal-Question-Metrics) [8] se puede resumir la definición del experimento de la siguiente forma:

Analizar: la Técnica de Lectura Ad hoc y Técnica de Lectura Basada en Perspectivas PBR,

Con el propósito de: evaluar,

Con respecto a: la cantidad de defectos detectados, tiempo de lectura, eficiencia, eficacia, facilidad de uso, preferencia e intención de uso,

Desde el punto de vista del: inspector,

En el contexto de: estudiantes de grado y recién recibidos del programa de grado de Licenciatura en Informática de la Universidad Nacional de San Juan, Argentina.

B. Planificación

En la planificación se determina cómo se ejecutará el experimento. Como muestra la Figura 2, la etapa de planificación se puede dividir en varios pasos [12], que pueden ejecutarse en forma iterativa hasta alcanzar el resultado final de esta etapa, el diseño del experimento.

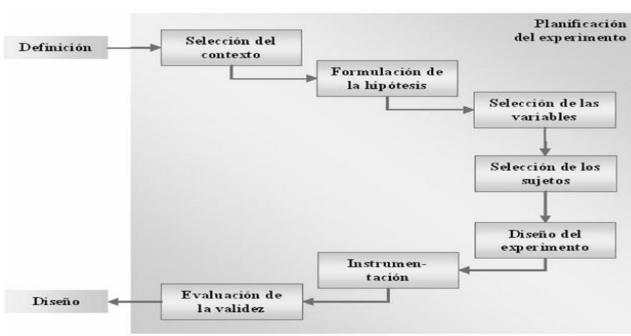


Figura. 2. Pasos de Planificación

1) Selección del contexto.

Para alcanzar resultados más generales el experimento debería ser realizado con proyectos reales y con profesionales experimentados, sin embargo esto es muy costoso en tiempo, dinero y esfuerzo; por lo tanto se torna difícil de llevar a cabo. Para reducir costos se ejecutan proyectos reales con estudiantes avanzados y egresados recientemente recibidos, en un entorno experimental controlado [12].

Desde el punto del contexto se deben tener en cuenta cuatro dimensiones que definen la generalidad del mismo [12], ellas son:

- El contexto del experimento se llevará cabo “off-line”, es decir, se hará en un entorno de laboratorio.
- Los sujetos serán estudiantes o recién egresados de carreras informáticas.
- El experimento abordará un problema real.
- El experimento será válido para un contexto específico.

2) Formulación de las hipótesis.

Este paso consiste en transformar la idea concebida para el experimento en una frase formal, es decir en hipótesis. Para formular las hipótesis debemos tener en cuenta que debemos definir la hipótesis nula y la hipótesis alternativa

A continuación se enuncian las diferentes hipótesis a validar con el experimento, definiendo TL como acrónimo de Técnica de Lectura.

a) Hipótesis de la cantidad de defectos detectados:

H0, a: Con la TL Ad hoc se detecta mayor o igual cantidad de defectos que con la TL PBR.

H1, a: Con la TL Ad hoc se detecta menor cantidad de defectos que con la TL PBR.

b) Hipótesis del tiempo de lectura:

H0, b: Con la TL Ad hoc se tarda mayor o el mismo tiempo que con la TL PBR.

H1, b: Con la TL Ad hoc se tarda menos tiempo que con la TL PBR.

c) Hipótesis de la eficiencia:

H0, c: La TL Ad hoc logra la misma eficiencia que la TL PBR.

H1, c: La TL Ad hoc no logra la misma eficiencia que la TL PBR

d) Hipótesis de la eficacia:

H0, d: La TL Ad hoc logra mayor eficacia que la TL PBR.

H1, d: La TL Ad hoc no logra mayor eficacia que la TL PBR.

e) Hipótesis de la facilidad de uso:

H0, e: La TL Ad hoc no es más fácil de usar por los sujetos que la TL PBR.

H1, e: La TL Ad hoc es más fácil de usar por los sujetos que la TL PBR.

f) Hipótesis de la preferencia de uso de los sujetos:

H0, f: La TL Ad hoc no alcanza una mayor preferencia por parte de los sujetos que la TL PBR.

H1, f: La TL Ad hoc alcanza una mayor preferencia por parte de los sujetos que la TL PBR.

g) Hipótesis de la intención de uso de los sujetos:

H0, g: La TL Ad hoc no alcanza una mayor intención de uso por parte de los sujetos que la TL PBR.

H1, g: La TL Ad hoc alcanza una mayor intención de uso por parte de los sujetos que la TL PBR.

3) Selección de las variables.

En este paso se deben seleccionar tanto las variables independientes como las variables dependientes del experimento, elegir la escala de medición y el rango de variables posibles. Las variables independientes son aquellas variables que se pueden controlar y cambiar dentro del experimento e influyen en las variables dependientes. Para el caso presentado existe solamente una variable independiente, la cual se corresponde con la técnica de lectura utilizada para realizar la lectura individual dentro del proceso de inspección de software, por lo tanto esa variable cuenta con dos niveles: TL Ad hoc y TL PBR. En el caso de las variables dependientes se distinguen dos tipos: aquellas que se basan en la performance (variables objetivas) y aquellas que se basan en la percepción (variables subjetivas).

Las variables basadas en la performance miden el resultado de cada uno de los sujetos en la ejecución o reproducción de las tareas experimentales. Las variables que se definen a continuación son consideradas como las variables objetivas más relevantes para este experimento:

- Cantidad de defectos detectados: definida como el número de defectos detectados que resulta de la aplicación de la TL al Documento de Especificación de Requisitos seleccionado.
- Tiempo de lectura: tiempo insumido en la aplicación de la TL al Documento de Especificación de Requisitos seleccionado.
- Eficiencia: esfuerzo, medido en tiempo, requerido para encontrar defectos con una TL. (Correlacionado con las variables cantidad de defectos detectados y tiempo de lectura).
- Eficacia: grado en que la TL logra el objetivo de detectar la totalidad de los defectos. (Correlacionado con la variable cantidad de defectos detectados y cantidad total de defectos)

Por otra parte, la medición de las variables basadas en la percepción, variables subjetivas, se realizará por medio de:

- Facilidad de uso: el grado en el que el sujeto percibe que la TL es simple de usar.
- Preferencia de uso: grado de predilección de los sujetos respecto de una TL.
- Intención de uso: el grado en el que el sujeto proyecta utilizar en el futuro una TL.

4) Selección de sujetos.

Teniendo en cuenta que la población se define como la totalidad, en este experimento se considera a todas aquellas personas que utilizan una TL en particular. Para generalizar los resultados a la población deseada, la selección debe ser representativa de ella. Para la selección o muestra de sujetos del experimento se usará la técnica de muestreo por conveniencia (no probabilística). La muestra estará compuesta por alrededor de 40 personas vinculadas a la carrera de Licenciatura en Informática, de la Facultad de Ciencias Exactas, Físicas y Naturales, dependiente de la Universidad Nacional de San Juan (Argentina) y que satisfagan alguna de las siguientes condiciones:

- Estudiantes cursantes del quinto año (último) de la carrera.
- Estudiantes avanzados que ya hayan realizado el curso de grado de Ingeniería de Software.
- egresados recientes (no mas de 2 años de egreso).

La participación de estudiantes en la experimentación no invalida los resultados, siendo éstos, útiles en muchos casos como pruebas piloto para experimentaciones que luego pueden trasladarse a la industria [18].

La posible diferencia de experiencia entre alumnos avanzados y graduados recientes es en realidad casi inexistente, en mercados laborales como el de San Juan (Argentina). Esto se debe a que, debido a la demanda de recursos humanos, los alumnos avanzados ya están insertos laboralmente en la industria con una experiencia similar a los graduados recientes. Adicionalmente, los alumnos avanzados y los graduados recientes, no utilizaban en su ámbito laboral las técnicas de lectura de documentos de requisitos de software comparadas en este trabajo.

5) Diseño del experimento.

Para obtener conclusiones significativas del experimento se aplicarán métodos de análisis estadísticos, tanto en la recolección como en la interpretación de los resultados. En el experimento el factor es la Técnica de Lectura. Este factor puede tomar dos valores o tratamientos: TL Ad hoc o TL PBR. Se opta por un diseño intra -sujetos, en el que a cada sujeto se le asigna ambos tratamientos.

De acuerdo a lo mencionado, se determina un diseño experimental completamente *aleatorizado unifactorial* con dos niveles.

TABLA 1. Ensayos del experimento

Objeto	Técnica de Lectura	
	Ad-hoc	PBR
Documento de Requisitos de Software A	Grupo Experimental A	Grupo Experimental B
Documento de Requisitos de Software B	Grupo Experimental B	Grupo Experimental A

Mediante un procedimiento aleatorio se formarán dos grupos con igual número de sujetos cada grupo, que serán los que intervendrán en ambos tratamientos. La Tabla 1 muestra la combinación de sujeto-tratamiento-objeto que tendrá este experimento.

6) Instrumentación.

Esta fase de planificación tiene que ver con los instrumentos que son necesarios para realizar el experimento y monitorearlo. Incluye los objetos experimentales, materiales varios, guías y formularios de pre-acondicionamiento y de control experimental y por último, las planillas de medición.

(1) Objetos a inspeccionar

Se les entregará a los sujetos un documento impreso que contendrá una detallada descripción de los Documentos de Especificación de Requisitos a inspeccionar. Estos documentos están bajo el formato de la Norma PSS-05 (Estándar de Ingeniería de Software de la ESA) [6].

(2) Guías de pre-acondicionamiento

A efectos de que los sujetos del experimento adquieran cierta destreza y nivelación en el manejo de ambas técnicas, será necesario organizar una serie de sesiones de instrucción preliminar, para lo cual se les facilitará una guía impresa conteniendo una breve descripción de las TL Ad hoc y PBR, además se les impartirá clases teóricas de ambas técnicas.

(3) Guías de las prácticas de entrenamiento y experimentales

Mediante el uso de formularios se intentará que, de una manera sencilla y ordenada, cada uno de los sujetos avance en los objetivos del experimento.

7) Evaluación de la validez

Es importante considerar cuán válidos son los resultados del experimento. Es decir, deben tener una validez adecuada para poder generalizar los resultados a la población de interés. Existen factores que afectan la capacidad de arribar a la conclusión correcta sobre las relaciones entre el tratamiento y el resultado de un experimento.

Una pregunta fundamental con respecto a los resultados del experimento es justamente ¿cómo validarlos? El grado de credibilidad de cualquier experimento depende de cómo se establece la validez de los resultados, por lo tanto se debe considerar la cuestión de la validez en forma anticipada, dentro de la etapa de planificación para poder planear y alcanzar adecuadamente la validación de los resultados del experimento.

(1) Validez de las Conclusiones.

Se refiere a la relación entre el tratamiento y el resultado.

Se debe asegurar que es una relación estadística significativa.

Se tendrán en cuenta los siguientes puntos:

- Los datos resultantes de la ejecución se analizarán e interpretarán con test y herramientas estadísticas adecuadas a efectos de potenciar el tratamiento del experimento.

Las variables objetivas, ya sean paramétricas o no paramétricas, se analizarán por medio de test de hipótesis de acuerdo a las distribuciones de los datos y de acuerdo a la cantidad de datos con que se cuenten.

Para realizar el tratamiento a las variables subjetivas, que se obtendrán por medio de una encuesta, será necesario analizar previamente la fiabilidad de los datos por medio del Alpha de Cronbach [5]

- Para controlar las diferencias individuales entre sujetos, se seleccionará un grupo homogéneo de estudiantes y recién egresados.

- El nivel de significancia que se tomará será el de la práctica estadística habitual, es decir, valores de α y β pequeños, que representan la probabilidad de cometer un error Tipo I y Tipo II respectivamente.

- El riesgo de que un grupo interfiera en el proceso de aplicación de una TL de otro grupo, se cree que es mínima. Sin embargo, para la ejecución del experimento se separará físicamente a ambos grupos.

- Se tratará de minimizar las perturbaciones que puedan llegar a distraer a los sujetos, como sesiones fuera de un horario adecuado, ruidos en el ambiente, climatización deficiente, hambre y sed, entre otras.

Validez Interna. Entendiendo como amenazas a la validez interna a aquellas que pueden llegar a influenciar a la variable independiente con respecto a la causalidad, se tomarán los siguientes recaudos:

- Se estandarizó el diseño del material de recolección de datos para que no haya una diferencia significativa en el proceso de inspección, ni en el llenado de la encuesta post-tareas. Además, a ambos grupos experimentales se les proveerá del mismo Documento de Especificación de Requisitos escrito de acuerdo al formato de la ESA [6].

- No se descarta que los participantes puedan cometer errores en algún cálculo de tiempos o en defectos encontrados, en el caso de que ocurra, en esa muestra se intentará rescatar re-calcular los valores y/o preguntando al sujeto en cuestión.

- Como se llevará a cabo una nivelación de conocimientos para todos los sujetos y la asignación de los participantes a los dos grupos será totalmente al azar, se entiende que se reducirá notablemente la amenaza de diferencias entre sujetos

- El experimento se llevará a cabo por vez primera con los sujetos seleccionados, de manera que el efecto de la experiencia por persistencia es prácticamente nula.

- Se estima que la muestra es representativa de la población que se estudia, aunque se considera que replicaciones adicionales con muestras de similar cantidad, ayudarán a consolidar los resultados del experimento en cuestión.

- Se tratará de minimizar el efecto de fatiga adecuando la duración de las sesiones programadas.

- Los participantes del experimento asistirán a un Taller de Técnicas de Lectura para Inspección de Software, con certificación del Instituto de Informática dependiente de la Universidad Nacional de San Juan. De esta manera se motivará a los sujetos y se intentará minimizar el efecto por mortalidad (abandono del experimento en cualquier punto del proceso).

- A los sujetos se les informará que no es posible hablar entre ellos durante la experimentación, además, el instructor estará presente en todo momento, para ayudar y verificar que no exista plagio ni influencias entre ellos.

(2) Validez de la Construcción.

Tiene que ver con la relación entre lo teórico y la observación. Si es posible generalizar los resultados del experimento a la teoría que lo sostiene. Las siguientes son amenazas a la validez de la construcción del experimento:

- Las variables dependientes objetivas elegidas son usadas siguiendo la propuesta de la ISO/IEC 14143 [13]. Las variables dependientes subjetivas son usadas y medidas con instrumentos de la literatura convencional. Se entiende, por ende, que la adopción de las variables mencionadas es la mejor manera de verificar las hipótesis planteadas.

- Están diferenciados los efectos que producen los tratamientos, en particular sobre el documento de requisitos de software (SRD) a inspeccionar, por lo que se cree es mínima la posibilidad de interacción entre tratamientos.

- Como no se dará a conocer inicialmente a los sujetos los objetivos buscados, las expectativas y el comportamiento no estarán basados en la suposición de ninguna hipótesis, ya que simplemente perseguirán el aprender las técnicas de lectura, obtener el certificado y participar del sorteo de un Pen Drive.

Validez Externa. Está relacionada con aquellas amenazas a la generalización de los resultados:

- Los estudiantes seleccionados, al estar en la última etapa de la carrera, poseen el perfil aproximado de los profesionales de la región, es decir, la interacción entre los sujetos y los tratamientos es alta. Existe mucha demanda en el medio para contratar estudiantes avanzados y recién egresados de la Universidad Nacional de San Juan.

- Al ser muy conveniente la experimentación con profesionales del medio, se invitará a participar del experimento, a los egresados de los dos últimos años de la carrera Licenciatura en Informática y de la carrera Programador con conocimientos en la temática abordada.

- No existe ninguna amenaza con respecto al momento en que se realizará el experimento.

- El documento de requisitos de software (SRD) a evaluar en el experimento, corresponde a un proyecto de desarrollo real. Sería muy bueno el realizar otros experimentos controlados con SRDs de mayor tamaño y complejidad, y casos de estudio en entornos diferentes a los de laboratorio.

C. Operación

Esta es la etapa en la que se ejecuta el experimento, es decir, donde se realiza la recolección de los datos que serán analizados.

La realización del experimento es un proceso que tiene como entrada el diseño del experimento y como salida principal los datos validados. La operación consta de tres pasos:

Tabla 2. Planificación de las tareas

Día	Tarea	Grupo Experimental A	Grupo Experimental B
1	1	Llenado de la Solicitud de Inscripción por parte de los aspirantes.	
	2	Procesamiento de las Solicitudes, selección adecuada de los sujetos y generación aleatoria de dos grupos experimentales	
	3	Disposición de todos los materiales (de acuerdo a la cantidad de participantes) que se utilizarán en las diferentes sesiones.	
2	4	Toma de asistencia, presentación del taller a los participantes y anuncio de los integrantes del Grupo A y del Grupo B.	
	5	Sesión de pre-acondicionamiento respecto del proceso de Inspección de Software y las diferentes Técnicas de Lectura con que cuentan los inspectores para la detección de defectos.	
3	6	Entrenamiento en la TL Ad hoc	Entrenamiento en la TL PBR
	7	Entrega del SRD para evaluación individual con la TL Ad hoc	Entrega del SRD para evaluación individual con la TL PBR
	8	Finalización del Proceso de Inspección y recolección de los datos de las variables objetivas.	
4	9	Llenado de la Encuesta sobre la TL Ad hoc	Llenado de la Encuesta sobre la TL PBR
	10	Examen Final para Aprobación del Taller	
5			
6			

2) Ejecución

El experimento se llevó a cabo en todas y cada una de las etapas planificadas en tiempo y forma.

La información obtenida mediante el formulario de inscripción permitió seleccionar sólo los que cumplían con los requisitos prefijados. Así, la cantidad de sujetos aptos para esta etapa fue de 28. La asistencia a las sesiones programadas era una condición indispensable, sólo asistieron a todas las sesiones 19 personas. Los grupos quedaron divididos con 10 y 9 sujetos cada uno. Tanto la asignación de grupos, como de roles en la técnica de lectura PBR fue hecha aleatoriamente, disminuyendo así cualquier tipo de amenaza a la validez de las conclusiones. Es necesario aclarar que si bien las cantidades de inspectores fueron las mismas para las TL *Ad-hoc* y PBR, los datos resultantes fueron en menor cantidad en la técnica de lectura PBR, debido a que para obtener una muestra se necesitaron tres inspectores, uno por cada perspectiva (Diseñador, Tester y Usuario). Por ello se contaron con 19 observaciones para TL *Ad-hoc* y 6 para PBR.

3) Validación de Datos

Se tuvo sumo cuidado en la recolección de los datos, cumplimentando las planillas y formularios experimentales por cada uno de los grupos de sujetos que participaron en

1) Preparación

En este paso, según Wohlin [12], se debe seleccionar e informar a los participantes y preparar todo el material que se va a utilizar.

En la Tabla 2 se muestra la planificación de tareas para el experimento.

dicho experimento. Concluida la etapa de recolección los datos, se procedió a la digitalización de los mismos y luego se auditó la carga de datos anterior con el fin de remover errores de carga y de otro tipo. Todas las inconsistencias fueron salvadas.

4) Interpretación de los resultados

Los datos sin procesar se encuentran alojados en www.idei.unsj-cuim.edu.ar/datos-exp1/index.htm, para realizar el análisis estadístico de los mismos se utilizó el software InfoStat versión 2.0. Primero se muestra un análisis descriptivo de las variables objetivas medidas bajo la técnica *Ad-Hoc* y PBR (Tabla 3 y Tabla 4). Luego se testean las hipótesis del trabajo y por último se muestran los resultados de las variables subjetivas.

Tabla 3. Aplicación de la técnica de lectura *Ad-Hoc*

Resumen	<i>Ad-hoc</i> Defectos	<i>Ad-hoc</i> Tiempo	<i>Ad-hoc</i> Eficiencia	<i>Ad-hoc</i> Eficacia
n	19,00	19,00	19,00	19,00
Media	11,11	59,47	20,67	22,21
D.E.	3,33	17,47	10,34	6,66
Var(n-1)	1,10	305,26	106,95	44,40
Mín	6,00	28,00	10,00	12,00
Máx	17,00	87,00	44,74	34,00

P(25)	9,00	45,00	13,75	18,00
P(50)	11,00	60,00	16,00	22,00
P(75)	13,00	79,00	26,67	26,00

Tabla 4. Aplicación de la técnica de lectura PBR:

Resumen	PBR Defectos	PBR Tiempo	PBR Eficiencia	PBR Eficacia
n	6,00	6,00	6,00	6,00
Media	33,33	259,50	13,38	66,67
D.E.	5,75	50,22	4,20	11,50
Var(n-1)	33,07	2521,90	17,61	132,27
Mín	27,00	167,00	10,04	54,00
Máx	43,00	310,00	21,56	86,00
P(25)	29,00	245,00	11,23	58,00
P(50)	31,00	269,00	11,72	62,00
P(75)	36,00	290,00	13,87	72,00

A continuación se procede a testear las hipótesis planteadas.

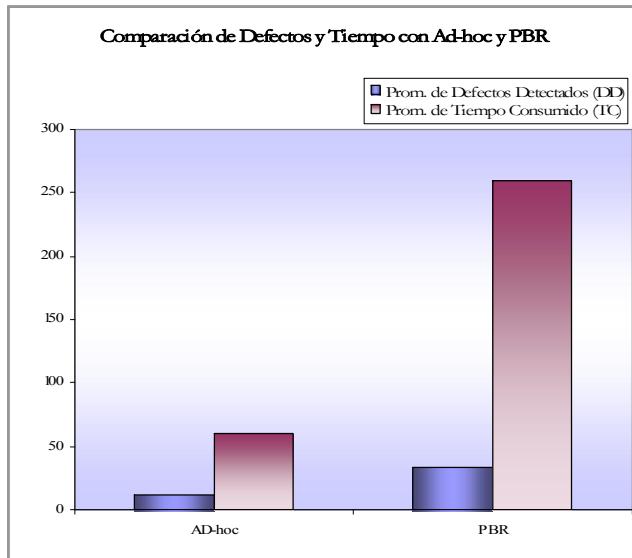
a) *Comparación de Defectos detectados y Tiempo Consumido en las técnicas Ad-Hoc y PBR*

Comparación entre *Ad-hoc* y *PBR*: se observa en la Tabla 5 y Figura 3, la comparación de defectos detectados y tiempos consumidos, expresado en minutos, en cada una de las técnicas de lectura.

Tabla 5. Defectos detectados y tiempo consumido en cada técnica

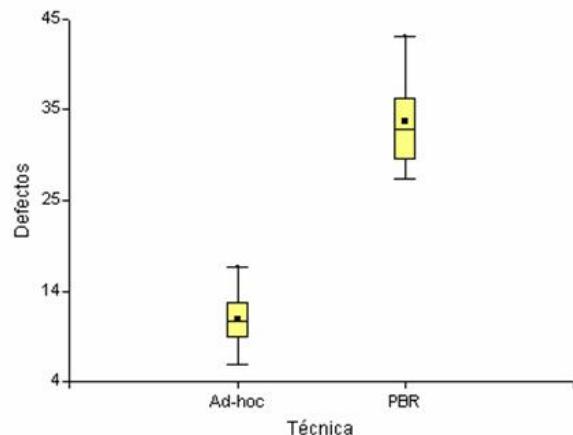
Técnica	Prom. De Defectos Detectados	Prom. Tiempo Consumido (TC)
<i>Ad-Hoc</i>	11	59
<i>PBR</i>	33	260
Prom.	22	159

Y gráficamente:

**Figura 3.** Defectos detectados y tiempo consumido en cada técnica

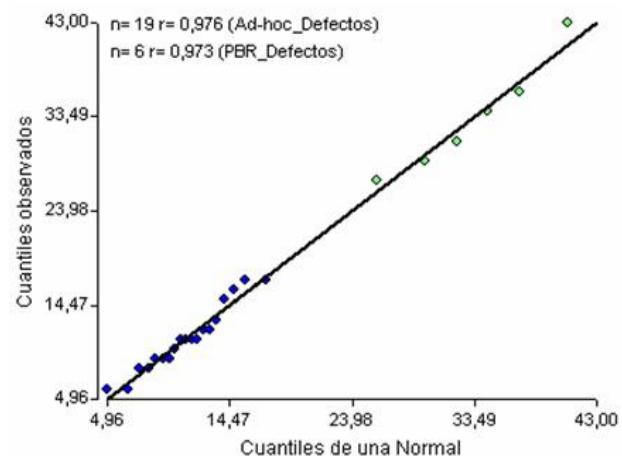
En la Figura 4, haciendo uso del gráfico de caja, se observa cómo la caja de la cantidad de defectos detectados con la técnica *Ad-hoc*

técnica *PBR*, está muy por encima de la caja de la cantidad de defectos detectados con la técnica *Ad-hoc*.

**Figura 4.** Defectos detectados con una y otra técnica de lectura.

Ahora, a efectos de determinar si es posible aplicar un test Paramétrico o No Paramétrico en la docimasia de hipótesis, se lleva a cabo el estudio gráfico (Q-Q Plot) y analítico (Test de Normalidad de Shapiro Wilks) para determinar si la distribución de los mismos es normal o no.

Resultado del Q-Q Plot: es mostrado en la Figura 5.

**Figura 5.** Q-Q Plot de defectos para ambas técnicas

Aparentemente, por lo que se observa, tanto el conjunto de la cantidad de defectos detectados con la técnica de lectura *Ad-hoc*, como con la TL *PBR* siguen una distribución normal, lo que se confirma con el siguiente test:

Tabla 6. Test de Normalidad de Shapiro Wilks

Variable	n	Media	D.E.	W*	p (una cola)
<i>Ad-hoc</i> Defectos	19	11,11	3,33	0,91	0,1892
<i>PBR</i> Defectos	6	33,33	5,75	0,94	0,7359

Como el valor de p es mucho mayor al nivel de significancia considerado en este experimento ($\alpha = 0,05$), se puede deducir que los datos correspondientes a la cantidad de

defectos detectados siguen una distribución normal. Luego se elige el test t como el más adecuado para comprobar o no la veracidad de las hipótesis.

Tabla 7. Test T unilateral

Clasif	Variabre	Grupo(1)	Grupo(2)	n(1)	n(2)
Criterio_Clas	Defectos	{1}	{2}	19	6
media(1)	media(2)	(Var.Hom.)	T	p	prueba
11,11	33,33	0,0787	-11,91	<0,00	UnilatI 01

Como el valor p es menor que el nivel de significancia $\alpha = 0,05$, se rechaza la H_0,a , por lo tanto se puede pensar la cantidad promedio de defectos detectados con la TL *Ad-hoc* es menor a la cantidad promedio de defectos detectados con la TL PBR.

Se concluye entonces, que la técnica de lectura PBR detecta mayor cantidad de defectos que la técnica de lectura *Ad-hoc*.

En el gráfico de caja de la Figura 6, se observa cómo la caja del tiempo consumido con la técnica PBR está bastante por encima de la caja del tiempo consumido con la técnica *Ad-hoc*.

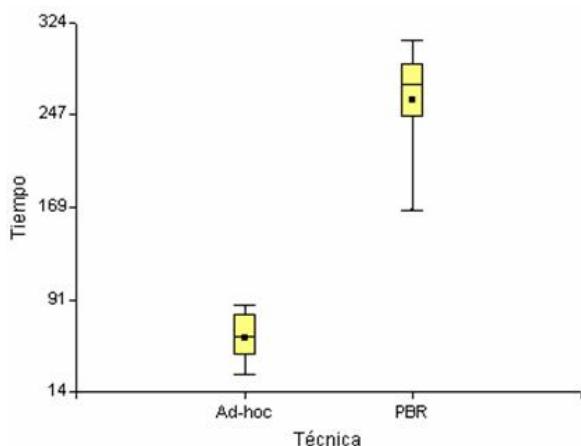


Figura 6. Tiempo consumido con una y otra técnica de lectura.

Resultado del Q-Q Plot

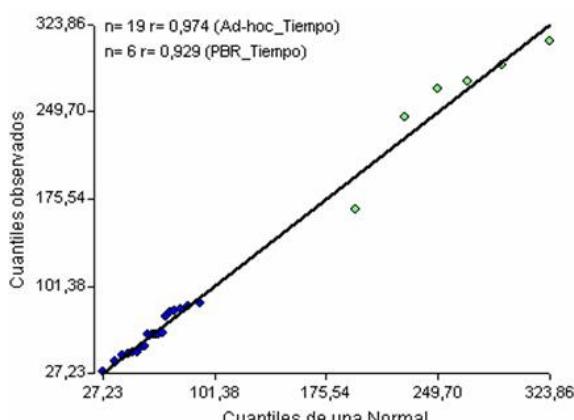


Figura 7. Distribución de los datos sobre la recta Y=X.

Aparentemente, por lo que se observa en la Figura 7, tanto el conjunto de datos del tiempo consumido en el proceso de inspección con la técnica de lectura *Ad-hoc*, como con la PBR siguen una distribución normal, lo que confirma el siguiente test.

Tabla 7. Test de Normalidad de Shapiro Wilks

Variable	n	Media	D.E.	W*	P(unacola)
Ad-hoc Tiempo	19	59,47	17,47	0,91	0,1801
PBR Tiempo	6	259,50	50,22	0,89	0,3641

Se observa en la aplicación de este test sobre los datos, que el valor de p es más grande que el nivel de significancia considerado en este experimento ($\alpha = 0,05$), por lo que se puede deducir que los datos correspondientes al tiempo consumido en el proceso de inspección con una y otra técnica de lectura, siguen una distribución normal. Luego el test T se considera el más adecuado para comprobar o no la veracidad de las hipótesis referentes al tiempo consumido por ambas técnicas de lectura. El resultado de ejecutar el test para el rechazo de la H_0, b es el siguiente:

Tabla 8. Test T unilateral

Clasif	Variable	Grupo(1)	Grupo(2)	n(1)	n(2)
Criterio_Clas	Tiempo	{1}	{2}	19	6
media(1)	media(2)	(p)(Var.Hom.)	T	p	prueba
59,47	259,50	0,0007	-9,58	0,0001	UnilatI 01

Como el valor p es menor que el nivel de significancia $\alpha = 0,05$, se rechaza la H_0,b , por lo tanto se puede pensar que el tiempo promedio consumido en el proceso de inspección con la TL *Ad-hoc* es menor al tiempo promedio consumido en el proceso de inspección con la TL PBR.

Se concluye entonces, que la técnica de lectura *Ad-hoc* consume menos tiempo que la técnica de lectura PBR.

De acuerdo a este análisis se presentan dos alternativas de elección: basada en la cantidad de defectos (PBR) o basada en el tiempo (*Ad-hoc*).

b) Comparación de la Eficiencia

Recordando, que la eficiencia es el esfuerzo requerido en el uso correcto de una TL, la misma se calculó dividiendo la cantidad de defectos detectados por el total de tiempo consumido en el proceso de detección.

Comparación entre *Ad-hoc* y PBR: se puede observar en la Tabla 9.

Tabla 9. Eficiencia lograda en cada técnica.

Técnica	Prom. De Defectos Detectados (DD)	Prom. De Tiempo Consumido (TC)	Eficiencia = DD/TC
Ad - Hoc	11	59	0.19
PBR	33	260	0.13
Prom.	22	15	0.16

En el gráfico de caja de la Figura 8 se observa cómo la caja de la eficiencia lograda con la técnica *Ad-hoc* se ubica por encima de la caja de la eficiencia lograda con la técnica PBR.

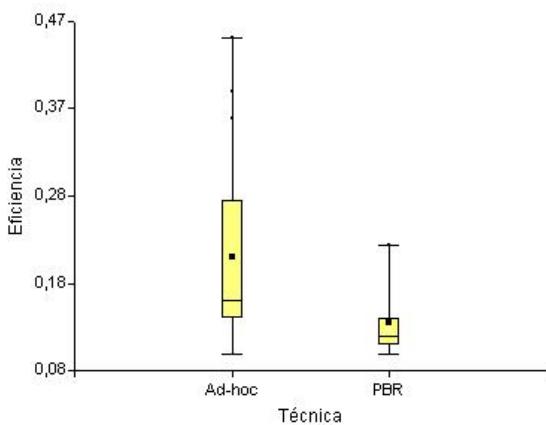


Figura 8. Eficiencia lograda con una y otra técnica de lectura.

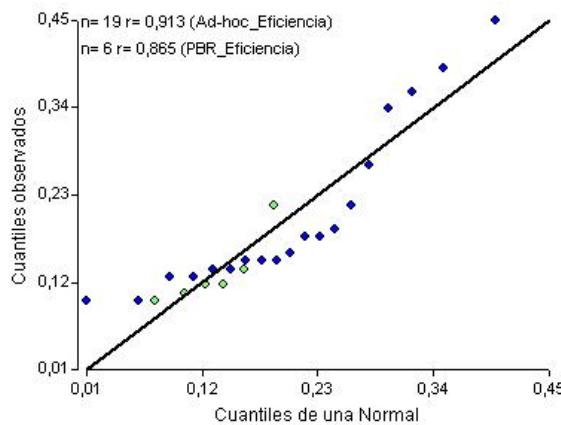


Figura 9. Q-Q plots de eficiencia en ambas técnicas

Observando atentamente el gráfico de la Figura 9, aparentemente el conjunto de datos de la eficiencia lograda con la técnica de lectura *Ad-hoc*, como con la TL PBR no siguen una distribución normal.

Tabla 10. Resultado del Test de Normalidad de Shapiro Wilks

Variable	n	Media	D.E.	W*	p (una cola)
<i>Ad-hoc</i>	19	0,21	0,10	0,81	0,0010
Eficiencia					
PBR Eficiencia	6	0,14	0,04	0,78	0,0443

Se observa en la aplicación de este test sobre los datos, que el valor de p es más pequeño para ambas TL, que el nivel de significancia considerado en este experimento ($\alpha = 0,05$), por lo que se puede deducir que los datos correspondientes a la eficiencia lograda con una y otra técnica de lectura, no siguen una distribución normal. Luego se elige el test no paramétrico de Wilcoxon como el más adecuado para comprobar o no la veracidad de las hipótesis. Los resultados al ejecutar el test para el rechazo de la H_0 , c es el siguiente:

Tabla 11. Test T unilateral

Clasific	Variable	Grupo 1	Grupo 2	n(1)	n(2)
Criterio	Clas	Eficiencia	1	2	19
Media(1)	Media(2)	DS(1)	DS(2)	W	p(2)

colas)					
0,21	0,14	0,10	0,04	46,50	0,0649

Como el valor p no es menor que el nivel de significancia $\alpha = 0,05$, no se rechaza la H_0 , c, por lo tanto se puede inferir que la media poblacional de la eficiencia obtenida con la TL *Ad-hoc* es igual a la media poblacional de la eficiencia obtenida con la TL PBR, con un nivel de significancia del 5%.

Se concluye entonces, que la eficiencia obtenida con la técnica de lectura PBR es igual a la eficiencia obtenida con la técnica de lectura *Ad-hoc* con un nivel de significancia del 5%.

c) Comparación de la Eficacia

En el siguiente gráfico de caja, Figura 10, se puede observar cómo la caja de la eficacia lograda con la técnica de lectura PBR está muy por encima de la caja de la eficacia lograda con la técnica *Ad-hoc*.

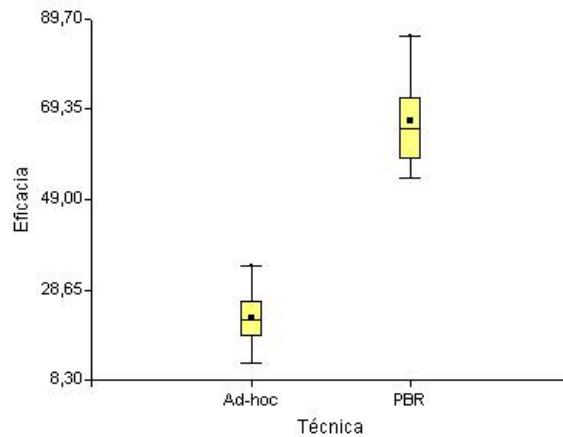


Figura 10. Eficacia lograda con una y otra técnica de lectura.

Resultado del Q-Q Plot

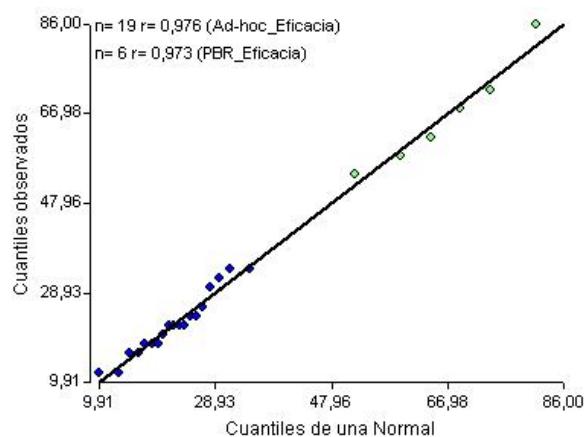


Figura 11. Q-Q Plots de la Eficacia en ambas técnicas

Por lo que se observa en el gráfico, de la Figura 11, tanto el conjunto de datos de la eficacia lograda con la técnica de lectura *Ad-hoc*, como con la TL PBR siguen una distribución normal.

Tabla 12. Resultado del Test de Normalidad de Shapiro Wilks

Variable	n	Media	D.E	W*	p (una cola)
Ad-hoc	19	22,21	6,66	0,91	0,1892
Eficacia	6	66,67	11,50	0,94	0,7359

Se observa en la aplicación de este test sobre los datos (Tabla 12), que el valor de p es mayor que el nivel de significancia considerado en este experimento ($\alpha = 0,05$), por lo que se puede deducir que los datos correspondientes a la eficacia alcanzada con una y otra técnica de lectura, siguen una distribución normal. Luego se elige el test T como el más adecuado para comprobar o no la veracidad de las hipótesis.

Los resultados del test para el rechazo de la H0, d se pueden observar en la Tabla 13.

Tabla 13. Test T unilateral

Clasif	Variable	Grupo(1)	Grupo(2)	n(1)	n(2)
Técnica	Eficacia	{Ad-hoc}	{PBR}	19	6
media(1)	media(2)	p(Var.Hom.)	T p	prueba	
22,21	66,67	0,0787	-11,91	<0,0001	UnilatI

Como el valor p es menor que el nivel de significancia $\alpha = 0,05$, se rechaza la H0, d, por lo tanto se puede pensar que la eficacia promedio obtenida con la TL Ad-hoc es menor a la eficacia promedio obtenida con la TL PBR.

Se concluye entonces, que la técnica de lectura PBR logra una mayor eficacia que la técnica de lectura Ad-hoc con un nivel de significancia del 5%.

d) Análisis de las Variables Subjetivas

Al finalizar el experimento se les proporcionó, como estaba previsto, la encuesta final sobre las técnicas de lectura utilizadas con el fin de identificar las medidas de estas tres variables: facilidad, preferencia e intensión de uso.

Los resultados de la encuesta mencionada se indican en la Tabla 14:

La falta de control que es posible ejercer sobre las variables subjetivas en cuestión hace sumamente necesaria la observación, a través de la matriz de correlación, del grado de similitud o correspondencia entre los datos recolectados.

Así, la Tabla 15 muestra tal parecido entre las preguntas realizadas:

Tabla 14. Resultados de la encuesta post-tareas

Inspector Nº	Facilidad de Uso														Preferencia de Uso								Intensión de Uso	
	Ad-hoc				PBR										Ad-hoc				PBR				Ad-hoc	PBR
	P1	P2	P3	P4	Prom	P6	P7	P8	P9	P11	P12	Prom	P5	P13	P14	Prom	P10	P13	P14	Prom	P15	P15		
1	4,0	4,0	3,5	3,5	3,8	3,5	3,5	3,5	3,0	3,8	4,0	3,6	3,0	1,0	1,0	1,7	3,0	5,0	5,0	4,3	1,0	5,0		
2	4,0	5,0	5,0	5,0	4,8	4,0	4,0	5,0	5,0	4,0	5,0	4,5	4,0	1,0	1,0	2,0	5,0	5,0	5,0	5,0	1,0	5,0		
3	5,0	5,0	5,0	5,0	5,0	4,0	4,0	5,0	5,0	4,0	4,0	4,3	4,0	1,0	1,0	2,0	4,0	5,0	5,0	4,7	1,0	5,0		
4	4,0	4,0	3,0	2,0	3,3	5,0	3,0	4,0	4,0	2,0	4,0	3,7	2,0	1,0	1,0	1,3	4,0	5,0	5,0	4,7	1,0	5,0		
5	5,0	4,0	4,0	4,0	4,3	4,0	4,0	5,0	4,0	3,0	5,0	4,2	4,0	1,0	1,0	2,0	5,0	5,0	5,0	5,0	2,5	2,5		
6	3,5	4,0	2,0	2,0	2,9	4,0	4,0	4,0	3,5	3,0	5,0	3,9	3,0	1,0	1,0	1,7	3,5	5,0	5,0	4,5	2,0	4,0		
7	5,0	4,0	3,0	2,0	3,5	4,0	4,0	5,0	5,0	4,0	5,0	4,5	3,0	1,0	1,0	1,7	5,0	5,0	5,0	5,0	1,0	5,0		
8	5,0	5,0	4,0	4,0	4,5	4,0	4,0	4,0	3,0	2,0	5,0	3,7	5,0	1,0	2,5	2,8	4,0	5,0	2,5	3,5	1,0	5,0		
9	2,0	2,0	4,0	4,0	3,0	4,0	2,0	2,0	2,0	4,0	5,0	3,2	4,0	2,5	1,0	2,5	4,0	2,5	5,0	3,5	1,0	5,0		
10	3,0	4,0	2,0	2,0	2,5	4,0	4,0	4,0	4,0	4,0	4,0	4,0	2,0	1,0	1,0	1,3	4,0	5,0	5,0	4,7	1,0	5,0		
11	4,0	4,0	4,0	3,0	3,8	4,0	3,0	4,0	4,0	2,0	5,0	3,7	4,0	2,5	2,5	3,0	4,0	2,5	2,5	3,0	1,0	5,0		
12	4,0	4,0	4,0	3,0	3,8	4,0	4,0	4,0	5,0	3,0	4,0	4,0	3,0	1,0	1,0	1,7	4,0	5,0	5,0	4,7	1,0	5,0		
13	4,0	3,0	4,0	4,0	3,8	4,0	5,0	5,0	5,0	4,0	5,0	4,7	3,0	1,0	1,0	1,7	5,0	5,0	5,0	5,0	1,0	5,0		
14	4,0	5,0	4,0	4,0	4,3	3,0	2,0	4,0	4,0	2,0	5,0	3,3	4,0	1,0	1,0	2,0	4,0	5,0	5,0	4,7	1,0	5,0		
15	4,0	5,0	4,0	4,0	4,3	4,0	5,0	5,0	4,0	4,0	5,0	4,5	3,0	1,0	1,0	1,7	5,0	5,0	5,0	5,0	1,0	5,0		
16	4,0	4,0	4,0	3,5	3,9	3,7	3,2	3,0	3,5	2,0	4,0	3,2	4,0	2,5	2,5	3,0	3,0	2,5	2,5	2,7	1,0	5,0		
17	3,0	3,0	1,0	2,0	2,3	1,0	5,0	2,0	4,0	2,0	5,0	3,2	3,0	1,0	1,0	1,7	5,0	5,0	5,0	5,0	1,0	5,0		
18	3,0	5,0	4,0	2,0	3,5	4,0	2,0	5,0	4,0	1,0	5,0	3,5	3,0	2,5	2,5	2,7	4,0	2,5	2,5	3,0	2,5	2,5		
19	1,0	4,0	1,0	1,0	1,8	5,0	5,0	5,0	5,0	5,0	3,0	4,7	1,0	1,0	1,0	1,0	5,0	5,0	5,0	5,0	1,0	5,0		
Prom. y Tot.	3,8	4,1	3,4	3,2	3,6	3,9	3,7	4,1	4,1	3,1	4,6	3,9	3,3	1,3	1,3	2,0	4,2	4,5	4,5	4,4	1,2	4,7		

Tabla 15. Matriz de correlación entre las preguntas formuladas en la encuesta.

			Facilidad de Uso										Preferencia de uso						Intención de Uso	
			AH					PBR					AH			PBR			AH	PBR
			P1	P2	P3	P4	P6	P7	P8	P9	P11	P12	P5	P13	P14	P10	P13	P14	P15	P15
FU	AH	P1	1,00	0,43	0,58	0,52	-0,04	0,04	0,31	0,16	-0,23	0,36	0,59	-0,26	0,12	-0,05	0,26	-0,12	0,05	-0,06
		P2	0,43	1,00	0,36	0,22	0,19	-0,03	0,62	0,32	-0,20	-0,02	0,18	-0,23	0,26	-0,05	0,23	-0,26	0,15	-0,16
		P3	0,58	0,36	1,00	0,84	0,22	-0,37	0,28	0,00	-0,10	0,32	0,70	0,25	0,25	-0,14	-0,25	-0,25	0,02	-0,08
		P4	0,52	0,22	0,84	1,00	-0,06	-0,09	0,09	-0,08	0,12	0,34	0,76	-0,02	-0,02	-0,02	0,02	0,02	-0,15	0,12
	PBR	P6	-0,04	0,19	0,22	-0,06	1,00	-0,09	0,56	0,15	0,32	-0,38	-0,26	0,05	0,05	-0,05	-0,05	0,08	-0,08	
		P7	0,04	-0,03	-0,37	-0,09	-0,09	1,00	0,28	0,49	0,51	-0,18	-0,33	-0,63	-0,36	0,55	0,63	0,36	-0,21	0,23
		P8	0,31	0,62	0,28	0,09	0,56	0,28	1,00	0,68	0,28	-0,04	-0,16	-0,35	-0,07	0,42	0,35	0,07	0,28	-0,30
		P9	0,16	0,32	0,00	-0,08	0,15	0,49	0,68	1,00	0,28	-0,23	-0,34	-0,43	-0,27	0,55	0,43	0,27	-0,09	0,07
		P11	-0,23	-0,20	-0,10	0,12	0,32	0,51	0,28	0,28	1,00	-0,34	-0,35	-0,41	-0,65	0,36	0,41	0,65	-0,34	0,35
		P12	0,36	-0,02	0,32	0,34	-0,38	-0,18	-0,04	-0,23	-0,34	1,00	0,60	0,15	0,15	0,26	-0,15	0,30	-0,29	
PU	AH	P5	0,59	0,18	0,70	0,76	-0,26	-0,33	-0,16	-0,34	-0,35	0,60	1,00	0,28	0,42	-0,19	-0,28	-0,42	0,05	-0,07
		P13	-0,26	-0,23	0,25	-0,02	0,05	-0,63	-0,35	-0,43	-0,41	0,15	0,28	1,00	0,68	-0,38	-1,00	-0,68	0,17	-0,20
		P14	0,12	0,26	0,25	-0,02	0,05	-0,36	-0,07	-0,27	-0,65	0,15	0,42	0,68	1,00	-0,38	-0,68	-1,00	0,17	-0,20
	PBR	P10	-0,05	-0,05	-0,14	-0,02	-0,05	0,55	0,42	0,55	0,36	0,26	-0,19	-0,38	-0,38	1,00	0,38	0,38	0,01	-0,06
		P13	0,26	0,23	-0,25	0,02	-0,05	0,63	0,35	0,43	0,41	-0,15	-0,28	-1,00	-0,68	0,38	1,00	0,68	-0,17	0,20
IU	AH	P14	-0,12	-0,26	-0,25	0,02	-0,05	0,36	0,07	0,27	0,65	-0,15	-0,42	-0,68	-1,00	0,38	0,68	1,00	-0,17	0,20
	PBR	P15	0,05	0,15	0,02	-0,15	0,08	-0,21	0,28	-0,09	-0,34	0,30	0,05	0,17	0,17	0,01	-0,17	-0,17	1,00	-0,98
		P15	-0,06	-0,16	-0,08	0,12	-0,08	0,23	-0,30	0,07	0,35	-0,29	-0,07	-0,20	-0,20	-0,06	0,20	0,20	-0,98	1,00

La sección sombreada representa la relación de una pregunta con otra en una misma variable.

Como se mencionó anteriormente surge la necesidad de analizar la fiabilidad de los datos obtenidos por medio de la encuesta post-tareas que se les facilitó a los sujetos que participaron en el experimento. Este Alpha de Cronbach considera que los datos de una variable subjetiva son fidedignos si es mayor a 0,70, y son medianamente fidedignos si es mayor a 0,50 [1]. De los valores resultantes de aplicar este coeficiente a los datos correspondientes a la facilidad de uso, se infiere que pueden ser considerados como datos medianamente fidedignos, situación que no se cumplió para la preferencia e intención de uso.

Tabla 16. Datos para la variable Facilidad de Uso

Reliability Coefficients		
Nº of Cases =	Nº of Items =	Alpha =
19	10	,6788

IV. CONCLUSIONES DEL EXPERIMENTO

Se presenta en este trabajo la descripción del proceso para diseñar e implementar un experimento controlado que tiene como objetivo medir, por un lado, la performance que tienen las técnicas de lectura de documentos de requisitos y por el otro, la percepción de los usuarios de dichas técnicas. Los resultados que se obtuvieron, específicamente respecto a las

hipótesis planteadas y analizadas estadísticamente, fueron los siguientes:

- Cantidad de defectos detectados: se pudo comprobar que la técnica de lectura PBR detecta muchos más defectos que la técnica de lectura *Ad-hoc*. Para este experimento la cantidad de defectos encontrados por la técnica PBR fue tres veces superior a la técnica *Ad-hoc*.

• Tiempo consumido: para esta variable, se pudo revelar que la técnica de lectura *Ad-hoc* consume menos tiempo de ejecución que la técnica PBR. Antes de interpretar fríamente estos valores, es conveniente destacar que la técnica de lectura PBR de por sí trabaja con tres perspectivas diferentes (diseñador, tester y usuario), las cuales son llevadas a cabo por especialistas en cada área, que generalmente suelen ser diferentes personas, por lo tanto, se consume mucho más tiempo en el proceso por el hecho de repetir la inspección tres veces por cada documento.

• Eficiencia: No pudo probarse fehacientemente, por medios estadísticos, cual de las dos técnicas es más eficiente. No obstante, al estar la eficiencia relacionada con el tiempo, y siendo la técnica *Ad-hoc* la que menos tiempo consume, se puede afirmar que es interesante la aplicación de la técnica *Ad-hoc* principalmente para encontrar los defectos mas sencillos de un documentos de requisitos.

• Efectividad: en este caso, la técnica de lectura PBR pudo ser probada estadísticamente como la que logró mayor

efectividad que la técnica *Ad-hoc*. Claramente la técnica PBR, para nuestro experimento, mostró encontrar un alto porcentaje de defectos, lo cual haría a esta técnica de gran utilidad con el fin de encontrar defectos tempranamente en el proceso de desarrollo de software.

Respecto a las variables subjetivas, a pesar de no haber podido validar la totalidad de los datos recolectados, se pudo observar que por parte de los sujetos experimentales hay una inclinación (percepción) hacia la técnica de lectura PBR. Esto parece deberse a la simpatía que provoca el mayor formalismo y mejor definición procedural que tiene la técnica PBR.

Los resultados expuestos pretenden dar a los ingenieros de software algunos datos que puedan mejorar la toma de decisión al momento de aplicar técnicas de lecturas de documentos de requisitos.

El diseño experimental presentado puede verse como un paquete de replicación. Potenciales replicas futuras de este experimento en otras universidades mejorarán los resultados precisando las conclusiones, incrementando de esta manera el conocimiento científico en el área de la ingeniería de software.

REFERENCIAS

- [1] Abrahão, S., Poels, G., Pastor, O. (2004). "Comparative Evaluation of Functional Size Measurement Methods: An Experimental Analysis". Working Papers of Faculty of Economics and Business Administration, Ghent University, Belgium 04/234.
 - [2] Ackerman A., Buchwald L. and Lewski F. (1989). "Software Inspections: An Effective Verification Process". IEEE software
 - [3] Basili V., Green S., Laitenberger O., Lanubile F., Shull F., Sørumgård S. and Zelkowitz M.V. et al. (1996) "The Empirical Investigation of Perspective-Based Reading," Empirical Software Eng.: An Int'l J., Vol. 1, No. 2, , pp. 133-164.
 - [4] Bernardz, B. "Una Aproximación Empírica a la Verificación de Especificaciones de Requisitos para Sistemas de Información". Tesis de Doctorado. Univ. de Sevilla – Dpto. de Lenguajes y Sistemas Informáticos (2003).
http://www.lsi.us.es/docs/doctorado/memorias/Memoria_ProyectoInvestigador.pdf
 - [5] Cronbach, L. J. (1951). "Coefficient alpha and the internal structure of tests." Psychometrika, 16(3), 297-334.
 - [6] ESA Software Engineering Standards – ESA PSS-05-0 Issue 2, Feb 1991.
 - [7] Fagan, M.E. (1976). Design and code inspections to reduce errors a program development. IBM System Journal, 15(3):182–211.
 - [8] Goal Question Metric: Para más información sobre la métrica, visitar <http://www.gqm.nl>
 - [9] Juristo, N., Moreno Ana M., Vegas, Sira (2004). Técnicas de evaluación de Software. Visitar http://is.ls.fi.upm.es/udis/docencia/erdsi/Documentacion_Evaluacion_7.pdf
 - [10] Shull F., Rus I., and Basili V., "How Perspective-Based Reading Can Improve Requirements Inspections", Computer, 33(7): 73-79(July 2000).
 - [11] Travassos G., Shull F., Fredericks M., Basili V. "Detecting defects in object-oriented designs: using reading techniques to increase software quality." 1999
 - [12] Wohlin C., Runeson P., Höst M., Ohlson M., Regnell B., Wesslén A."Experimentation in Software Engineering. An Introduction."
 - [13] www.iso.org
 - [14] Zelkowitz M.V. y Wallace D. (1997). Experimental validation in software engineering. Information and Software Technology. 39(11).
 - [15] Biffl St., Halling M. "Software Product Improvement with Inspection", Proc. Of Euromicro 2000 Workshop on Software Product and Process Improvement, Maastricht, IEEE Comp. Soc. Press, Sept. 2000.
 - [16] Gilb T., Graham D., "Software Inspection", Addison- Wesley, 1993. www.iese.fhg.de, May 2000.
 - [17] Halling M., Biffl St., (2001) "Using Reading Techniques to Focus Inspection Performance", Proc. of Euromicro 2001 Workshop on Software Product and Process Improvement, Warsaw, IEEE Comp. Soc. Press, Sept. 2001.
 - [18] Carver, J.; Jaccheri, L.; Morasca, S.; Shull. "Issues in using students in empirical studies in software engineering education", F. Software Metrics Symposium, 2003. Proceedings. Ninth International Volume , Issue , 3-5 Sept. 2003 Page(s): 239 – 249
- María Inés Lund.** Nacida en la ciudad de San Juan, Argentina, en el año 1966, obtuvo el título de grado en Licenciatura en Informática de la Universidad Nacional de San Juan, Argentina en el año 1991 y el de posgrado en Especialista en Sistemas de Información para Intranets de la Universidad Nacional de San Juan, en el año 2001.
Ella se desempeña como Consultor Independiente en Ingeniería de Software, específicamente en el área de Análisis y Diseño de Software. También es Profesor e Investigador, desde el año 1987, en la Universidad Nacional de San Juan.
Prof. Lund actualmente es co-directora del proyecto de investigación denominado "Adaptación de Modelos Internacionales de Procesos de Software a Empresas Locales Productoras de Software" (AMIPROSOFT).
- Myriam Beatriz Herrera.** Nació en la ciudad de San Juan, Argentina, en el año 1962, obtuvo el título de grado en Profesorado en Matemática en la Universidad Nacional de San Juan, San Juan, Argentina en el año 1984 y Licenciatura en Ciencias Matemáticas en la Universidad Nacional de San Luis, San Luis, Argentina en el año 1986. Obtuvo el título de posgrado de Magíster en Matemática en la Universidad Nacional de San Luis, San Luis, Argentina en el año 2001.
Ella se desempeña como Consultor independiente en Estadística. También es Profesor e Investigador de la Universidad Nacional de San Juan desde el año 1986. Actualmente está desarrollando investigaciones en clasificación digital en el Instituto de Informática de la Universidad Nacional de San Juan, Argentina.
Prof. Herrera es miembro del proyecto de investigación, financiado por la Universidad Nacional de San Juan, denominado "Adaptación de Modelos Internacionales de Procesos de Software a Empresas Locales Productoras de Software" (AMIPROSOFT).
- Laura Nidia Aballay.** Nació en la ciudad de Chepes, Provincia de La Rioja, Argentina, en el año 1974. En el año 1998 obtuvo el título técnico de Programador Universitario en la Universidad Nacional de San Juan, Argentina. Ella se desempeña como profesional independiente en el área de ingeniería de software. Es docente-investigadora del Instituto de Informática de la Universidad Nacional de San Juan desde el año 1999.
Prof. Aballay es miembro del proyecto de investigación, financiado por la Universidad Nacional de San Juan, denominado "Adaptación de Modelos Internacionales de Procesos de Software a Empresas Locales Productoras de Software" (AMIPROSOFT).
- Sergio Gustavo Zapata.** Nació en la ciudad de San Juan, Argentina, en el año 1964, obtuvo el título de grado en Licenciatura en Ciencias de la Computación en la Universidad Nacional de San Luis, San Luis, Argentina en el año 1989.
El se desempeña como Consultor independiente en Ingeniería de Software, específicamente en el área de Calidad de Software. También es Profesor e Investigador de la Universidad Nacional de San Juan desde el año 1987.
Actualmente es director del Instituto de Informática de la Universidad Nacional de San Juan, Argentina.
Prof. Zapata es director del proyecto de investigación denominado "Adaptación de Modelos Internacionales de Procesos de Software a Empresas Locales Productoras de Software" (AMIPROSOFT).

Universidad Nacional de Colombia Sede Medellín

Facultad de Minas



Escuela de Ingeniería de Sistemas

Misión

La misión de la Escuela de Ingeniería de Sistemas es fomentar y apoyar la generación o la apropiación de conocimiento, la innovación y el desarrollo tecnológico en el área de ingeniería de sistemas e informática sobre una base científica, tecnológica, ética y humanística.



Escuela de Ingeniería de Sistemas
Dirección Postal:
 Carrera 80 No. 65 - 223 Bloque M8A
 Facultad de Minas. Medellín - Colombia
 Tel: (574) 4255350 Fax: (574) 4255365
 Email: esistema@unalmed.edu.co
<http://pisis.unalmed.edu.co/>

Visión

La formación integral de profesionales desde el punto de vista científico, tecnológico y social que les permita adoptar, aplicar e innovar conocimiento en el campo de los sistemas e informática en sus diferentes aspectos, aportando con su organización, estructuración, gestión, planeación, modelamiento, desarrollo, procesamiento, validación, transferencia y comunicación; para lograr un desempeño profesional, investigativo y académico que contribuya al desarrollo social, económico, científico y tecnológico del país.

