

## Factores que afectan negativamente a la aplicación práctica de las pruebas de software

Luis Fernández-Sanz<sup>1</sup>, Pedro Lara<sup>2</sup>, M.Teresa Villalba<sup>2</sup>, Tanja Vos<sup>3</sup>

<sup>1</sup> Departamento de Ciencias de la Computación, Universidad de Alcalá  
Ctra. Madrid-Barcelona Km 33,600, 28771 Alcalá de Henares  
[luis.fernandezs@uah.es](mailto:luis.fernandezs@uah.es)

<sup>2</sup> Departamento de Sistemas Informáticos, Universidad Europea de Madrid,  
C/Tajo s/n, 28670 Villaviciosa de Odón, Madrid  
{pedro.lara, maite.villalba}@uem.es

<sup>3</sup> Instituto Tecnológico de Informática, Universidad Politécnica de Valencia,  
Camino de Vera s/n, 46022 Valencia  
[tanja@iti.upv.es](mailto:tanja@iti.upv.es)

**Resumen.** En la mayoría de las reuniones y congresos dedicados a la calidad y las pruebas de software existe un sentimiento bastante generalizado de decepción por el poco interés detectado en las organizaciones de desarrollo para sistematizar y mejorar las pruebas de software. En un workshop de expertos en 2007 se sugirieron diversos factores que podrían justificar el sentimiento negativo sobre la práctica real de las pruebas. En este trabajo se presentarán los resultados iniciales del análisis de dichos factores basado en las opiniones obtenidas de una gran cantidad de profesionales de la ingeniería y la calidad del software.

**Palabras Clave:** Pruebas de software, factores, personal.

### 1 Introducción

Las pruebas de software constituyen la técnica de verificación y validación de software más utilizada en los proyectos de desarrollo ya que está presente sistemáticamente en todos ellos en un grado u otro de madurez o eficacia: el resto de técnicas de aseguramiento de calidad del software suele aplicarse en porcentajes relativamente bajos de organizaciones y desarrollos. A pesar de esta presencia generalizada, muchos expertos consideran que las pruebas han quedado tradicionalmente apartadas del interés de las organizaciones y de los profesionales en cuanto a la mejora de los procesos y técnicas de trabajo.

Lamentablemente no es habitual encontrar estudios sistemáticos sobre la práctica real de las pruebas a nivel internacional: en cuanto a datos correspondientes a España simplemente podemos decir que son prácticamente inexistentes. No obstante, sí se han analizado algunas aportaciones al análisis de los factores que influyen en la implantación real de las pruebas más allá de las contribuciones genéricas o indirectas que puedan encontrarse en la literatura de mejora general de procesos de software y de sus modelos habituales como CMMi, ISO15504, etc. Así algún autor justifica que las pruebas sean la actividad probablemente más difícil dentro del desarrollo de software [1] por diferentes motivos. Uno de los temas más tratados en estas contribuciones ha sido el de la estimación de esfuerzo y la complejidad de las pruebas a través de los trabajos presentados en [2] y [3] donde se estudian algunos de los factores que influyen en la existencia de problemas en esta área. Así se distingue entre distintas categorías de factores:

- los ligados a los propios procesos de prueba que no están convenientemente optimizados

- los vinculados al proyecto en cuanto su propia naturaleza y complejidad así como los relacionados con la disponibilidad de recursos y herramientas
- los aspectos relativos al personal, especialmente en cuanto a su formación, cualificación y actitud
- otros factores que incrementan la complejidad y la dificultad de las pruebas.

Por otra parte también existen análisis generales que abordan específicamente la implantación real de prácticas de pruebas en algunos países [4] pero resultan esfuerzos aislados poco útiles para nuestro propósito de conocer más sobre la realidad de la implantación de las pruebas y sus dificultades en España. Por ello, apoyándonos en las acciones de la Red para la promoción y mejora de las Pruebas en la Ingeniería del Software (RePRIS) se procedió a lanzar el estudio presentado en este trabajo. RePRIS (<http://in2test.lsi.uniovi.es/repris/>), iniciada en 2005, contó con el apoyo de la Acción Especial TIN2005-24792-E del Ministerio de Educación y Ciencia y, gracias a los resultados obtenidos, logró a finales de 2007 una renovación de financiación para sus actividades a través de la Acción Especial TIN2007-30391-E.

## **2 Estudio sobre factores que influyen en la práctica real de las pruebas**

El estudio de los factores que pueden influir en la práctica real de las pruebas de software en España se ha apoyado en un proceso en 3 fases. Como se ha comentado, la motivación surge al recibir los comentarios de numerosos profesionales e investigadores que insisten en las dificultades de todo tipo existentes para que las pruebas alcancen una sistematización y madurez acorde a su importancia en la práctica real de las organizaciones. Así, en reuniones comerciales, jornadas técnicas, congresos (por ej., las acciones de promoción de REPRIS en diversos eventos y reuniones o los del grupo de Calidad de Software de ATI [www.ati.es/gtcalidadsoft](http://www.ati.es/gtcalidadsoft)) o en acciones de formación abiertas o in-company, se pueden recibir opiniones fundamentadas de profesionales con experiencia sobre diversas causas de la inmadurez de las pruebas, al menos en España.

La Red REPRIS, en su reunión de Zaragoza de septiembre de 2007 a la que acudió una gran variedad de expertos, realizó un análisis colectivo de cuáles podían ser dichos factores. Sintetizados en un documento (confirmado posteriormente en la lista de distribución de la red) estos factores se consideraron como los más influyentes en la situación poco madura de la práctica de las pruebas en España. En total se trata de 23 factores (ver tabla 1) expresados en forma de una sentencia que describe alguna circunstancia o situación que parece ocurrir realmente en la práctica habitual de las pruebas de software en nuestro país. Estos factores se agrupan en 6 categorías relacionadas con las decisiones organizativas sobre pruebas, la carrera profesional en esta disciplina, la formación sobre pruebas de los profesionales, la actitud de distintos tipos de profesionales, las estrategias y técnicas habitualmente utilizadas en las organizaciones y la situación del mercado del software.

Lógicamente, a pesar del valioso resultado obtenido, era necesario obtener algún tipo de contraste con las opiniones de los profesionales que, con sus comentarios, habían motivado este trabajo. Los objetivos planteados en esta acción de comprobación fueron dos:

- Estudiar hasta qué punto se podían considerar como generalmente aceptados los 23 factores o situaciones detectados,
- Determinar hasta qué punto dichos profesionales consideraban que cada uno de los factores era total o parcialmente determinante en las pruebas o, por el contrario, si su influencia era nula a pesar de ser una situación presente en el desarrollo de software.

En ambos casos era necesario que la muestra fuera suficientemente amplia para que los resultados tuvieran una cierta representatividad no basada en muestreos científicos de la población implicada, los desarrolladores (cuyo universo es difícil de caracterizar en España debido a la carencia estadística detallada de las administraciones públicas), sino en variedad de sectores y puestos de la misma. Por ello se procedió a recoger la opinión de profesionales a través de cuestionarios aprovechando la presencia en reuniones y foros especializados, acciones formativas e, incluso, a través de contacto directo electrónico. En la sección 3 se presentan los resultados obtenidos y los datos de la muestra.

Por último, se plantea una fase de trabajo en la que se somete a la difusión pública los resultados obtenidos y se trata de recoger las opiniones que suscita su presentación entre los profesionales y los expertos, Para ello, se procederá a trabajar tanto en las próximas X Jornadas de Innovación y Calidad del Software como, con el presente trabajo, en el Taller PRIS 08.

### 3 Encuesta de opiniones sobre los factores que influyen negativamente en las pruebas del software

Como ya se ha comentado la encuesta propuesta para determinar los factores que contribuyen a que las pruebas de software no alcancen un nivel apropiado de eficacia y eficiencia se basa en un cuestionario en el que, para cada uno de los 23 factores mencionados anteriormente, se pregunta a los profesionales que respondan si, en su experiencia:

- Si se puede considerar como cierto o no que dicho factor o situación ocurra realmente en la prácticas habitual de pruebas de las organizaciones de software
- Si la posible influencia de dicho factor, en caso de producirse, sería totalmente negativa para las pruebas o, por el contrario, tiene una influencia parcial o, incluso, se podría considerar que no influye en la práctica de las organizaciones de software.

#### 3.1 Muestra de profesionales

Aunque estos cuestionarios son anónimos para evitar cualquier tipo de inhibición de los posibles respondientes, sí se ha incluido un apartado de datos generales relacionados con las siguientes características profesionales de los mismos:

- Sector habitual en el que suelen desempeñar su actividad de desarrollo de software
- Tipo de aplicaciones con las que habitualmente tratan en su trabajo
- Puesto actual, años de experiencia profesional y de experiencia en su puesto actual

La recogida de estos datos nos permite caracterizar la muestra utilizada para este trabajo y que consta de 65 profesionales que respondieron a los cuestionarios (se desecharon algunos cuestionarios por contener respuestas incompletas, confusas o con algún otro tipo de problemas). Por sectores, la distribución es la mostrada en la Figura 1.

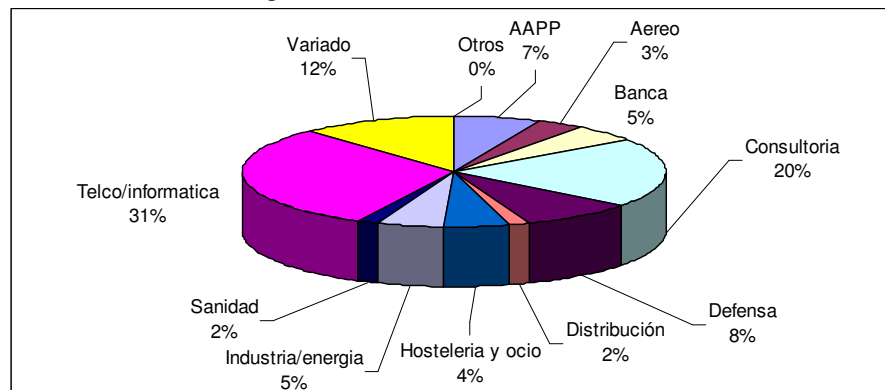


Fig. 2. Distribución sectorial de la muestra de profesionales que respondieron la encuesta

En cuanto a la distribución del tipo de software con el que habitualmente tratan los encuestados, los resultados son los siguientes:

- Software empotrado: 14,06%
- Sistemas Web: 15,63%
- Software de control: 6,25%
- Software de gestión: 15,63%
- Combinación de web y ERPs: 6,25%
- S. de información geográfica: 3,13%
- Software variado no restringido a un solo tipo: 34,38%

Por otra parte, en cuanto al puesto ocupado en el momento de responder al cuestionario, la distribución está reflejada en la Figura 2. El promedio de años de experiencia de quienes han respondido al cuestionario es de 8 mientras que el número medio de años desempeñan las labores de su puesto actual es de 3,62.

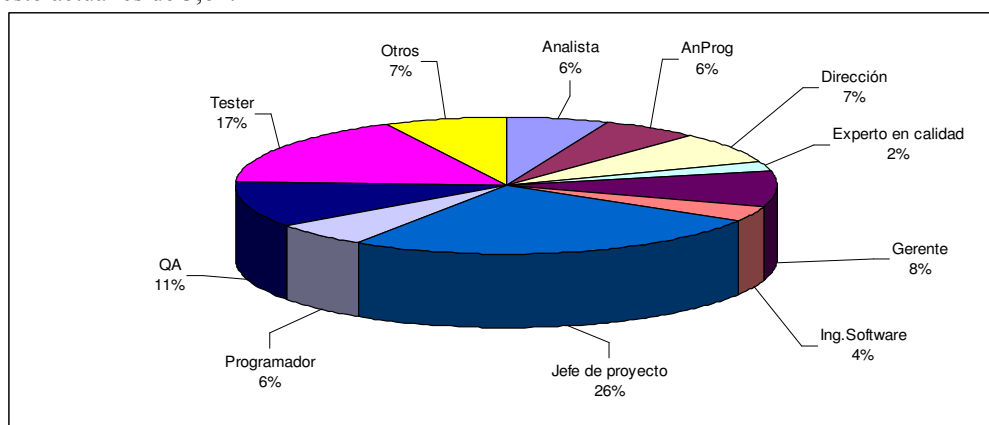


Fig. 2. Distribución de puesto de la muestra de profesionales que respondieron la encuesta

### 3.2 Resultados de ocurrencia de factores

Los resultados obtenidos de los cuestionarios respondidos para la muestra de profesionales en cuanto a la ocurrencia o no de cada uno de los factores se sintetizan en la tabla 1. Como puede verse, podemos clasificar a los factores en función de su grado de aceptación:

- Ítems sin aceptación general de los profesionales como factores que realmente aparezcan en la práctica profesional: 1.2, 2.1, 2.2, 2.3, 5.1 y 6.2, Entre ellos aparecen casos con más detractores que opiniones a favor (2.1 y 2.2) así como otros en los que no se llega a un 60% de profesionales que los consideran ciertos.
- Ítems bastante aceptados como hecho que ocurren realmente (columna afirmativa entre 60% y 80%): 2.4, 2.5, 3.5, 3.6, 4.1, 5.2, 5.3, 5.6 y 6.1.
- Ítems abrumadoramente considerados como hechos ciertos en la práctica habitual en el desarrollo de software (opiniones afirmativas mayores que el 80%): 1.1, 3.1, 3.2, 3.3, 3.4, 4.2 y 5.5

En forma resumida podemos considerar que, principalmente, son motivo de una clara división de opiniones las afirmaciones relacionadas con la ausencia de carrera profesional (o las malas condiciones de la misma) para los especialistas en pruebas así como el hecho de que se hagan pruebas por agotamiento de recursos o que el mantenimiento de errores propios sea un negocio. En general, el resto de factores se consideran como habituales aunque, con clara diferencia, son los relacionados con la falta de formación sobre pruebas así como la presión de realizarse como última fase del desarrollo y la tentación de recortar en calidad cuando hay problemas financieros.

**Tabla 1.** Porcentajes de respuestas sobre la ocurrencia de los factores

Factor	Sí	No	NC
1.1. Se recorta en calidad y pruebas cuando hay problemas bien de retrasos y de proyecto o bien de finanzas de la organización.	92,2%	7,8%	0,0%
1.2. No son raros los casos en los que especialistas en QA y pruebas ven desaparecer su puesto y pasan a desarrollo.	51,6%	42,2%	6,3%
2.1 Son poco creativas, un fastidio, algo que hay que hacer, sin atractivo, negativo (buscar defectos) y destructivo (desprestigia el trabajo de desarrolladores)	46,9%	53,1%	0,0%
2.2 Es un área sin carrera profesional o promoción, etc.	42,2%	57,8%	0,0%
2.3 El plan de carrera en pruebas no promete que se pague lo mismo, más bien lo contrario, que en otras trayectorias.	59,4%	37,5%	3,1%
2.4 No se suele reconocer el trabajo, no se suele facturar ("es un gasto obligado del proyecto"), el servicio es más bien interno y no frente al cliente externo, etc.	71,9%	26,6%	1,6%
2.5 Los testers más básicos no requieren un título universitario: seguramente un título de FP es suficiente (incluso más que para programadores): imagen de carrera poco atractiva.	62,5%	37,5%	0,0%
3.1 Muchos titulados no tienen formación específica en pruebas.	98,4%	1,6%	0,0%
3.2 Muchos profesionales tampoco la tienen.	89,1%	9,4%	1,6%
3.3 Tampoco los profesionales suelen recibir formación en pruebas. Se insiste más en tecnologías, nuevas versiones de productos o, quizás, en métodos de desarrollo y herramientas.	93,8%	6,3%	0,0%
3.4 En las universidades no es un tema estrella: muchos lo mencionan y piden su uso pero pocos comprenden su filosofía y la manera apropiada de probar	92,2%	7,8%	0,0%
3.5 En la formación se presta poca atención a pruebas funcionales y de sistema y más a pruebas unitarias o de detalle.	64,1%	31,3%	4,7%
3.6 Poca importancia o tradición de formación y cualificación específica en pruebas: existe poco material de cursos, certificaciones, etc.	78,1%	20,3%	1,6%
4.1 Los junior (o recién titulados) vienen con actitud muy centrada en programar y en código y no desean aplicarse en otras actividades	76,6%	21,9%	1,6%
4.2 Muchos directivos no aprendieron sobre pruebas y no aprecian su interés o potencial para la eficiencia y la calidad	84,4%	14,1%	1,6%
5.1 Se tiende a las pruebas por agotamiento de recursos: si probamos mucho, aunque sea de manera descontrolada, terminaremos cubriendo todos los aspectos del sistema.	57,8%	40,6%	1,6%
5.2 No se suele planificar y a diseñar casos eficientes para un coste mínimo, ni se vincula a prioridades o riesgos, no se controla el riesgo incurrido según la cantidad o calidad de pruebas, no se controlan evidencias, etc.	76,6%	23,4%	0,0%
5.3 El diseño de pruebas significa muchas veces rehacer un trabajo que el analista no remató o documentó bien ya que las pruebas son totalmente dependientes de la especificación.	79,7%	18,8%	1,6%
5.4 Las pruebas quedan atrapadas al final del ciclo y muchas veces con un calendario escaso. Suelen sufrir los retrasos de fases anteriores de desarrollo y no pueden retrasar su final ya que lo siguiente es la entrega a cliente	92,2%	6,3%	1,6%
5.5 Falta de relación entre modelos de desarrollo y actividades de prueba, especialmente para su diseño: "Las pruebas son algo que se realiza al final, una vez que tengo código".	89,1%	10,9%	0,0%
5.6 Es muy poco habitual diseñar pruebas en cuanto tengo una especificación aunque, incluso, el diseño de pruebas puede simultanearse con el análisis puesto que documenta el conocimiento de la funcionalidad y los requisitos.	78,1%	20,3%	1,6%
6.1 El mercado es poco serio (aún en proporción que no es baja) y no se penalizan problemas de calidad del software entregado.	71,9%	25,0%	3,1%
6.2 El negocio redondo aparece cuando se cobra mantenimiento de elementos defectuosos entregados en un desarrollo (cobrar por reparar los defectos que uno ha insertado).	56,3%	40,6%	3,1%

### 3.3 Grado de influencia de los factores

En la encuesta no sólo se pretendía medir la aceptación de los factores presentados como hechos o situaciones reales que ocurren en el mundo profesional del software sino evaluar también cuál es, en caso de producirse, el grado real de influencia negativa de los mismos en las prácticas cotidianas de los proyectos. Para ello se pidió a los encuestados que indicaran si dicha influencia se consideraba total en la situación y las prácticas de las pruebas o, por el contrario, en grado parcial o realmente inexistente en cuanto que, a pesar de que dicha situación pudiera producirse, ello no perjudicaría a la madurez, eficacia y eficiencia de las pruebas. Los resultados obtenidos sobre esta valoración de la influencia los incluimos en la Tabla 2. Como se puede apreciar en dicha tabla, suelen asociarse los valores que más frecuentemente se consideran como ciertos con los valores de influencia más elevados (considerando tanto influencia total como parcial).

**Tabla 2.** Porcentajes de respuestas sobre el grado de influencia de los factores

Factor	% Total	% Parcial	% NC
1.1.	92,2%	7,8%	0,0%
1.2.	51,6%	42,2%	6,3%
2.1	46,9%	53,1%	0,0%
2.2	42,2%	57,8%	0,0%
2.3	59,4%	37,5%	3,1%
2.4	71,9%	26,6%	1,6%
2.5	62,5%	37,5%	0,0%
3.1	98,4%	1,6%	0,0%
3.2	89,1%	9,4%	1,6%
3.3	93,8%	6,3%	0,0%
3.4	92,2%	7,8%	0,0%
3.5	64,1%	31,3%	4,7%
3.6	78,1%	20,3%	1,6%
4.1	76,6%	21,9%	1,6%
4.2	84,4%	14,1%	1,6%
5.1	57,8%	40,6%	1,6%
5.2	76,6%	23,4%	0,0%
5.3	79,7%	18,8%	1,6%
5.4	92,2%	6,3%	1,6%
5.5	89,1%	10,9%	0,0%
5.6	78,1%	20,3%	1,6%
6.1	71,9%	25,0%	3,1%
6.2	56,3%	40,6%	3,1%

## 4 Conclusiones

El presente estudio sobre los factores que influyen negativamente en la implantación efectiva y eficiente de las mejores prácticas sobre pruebas de software constituye una aproximación más a la realidad de esta disciplina en España. Como ya se ha comentado, con el apoyo de la red REPRIS, se pretende obtener una imagen más acertada de la situación actual de las pruebas como ya se ha hecho a través del análisis de las prácticas realmente implantadas en las organizaciones [4] así como de la manera en que los profesionales abordan el diseño de casos [5].

Por supuesto, es intención de los autores seguir incrementando la representatividad de la muestra elevando el número de profesionales que contesten al cuestionario en España a la vez que se extenderá su aplicación a profesionales de otros países para poder confirmar si las conclusiones obtenidas son globales o, por el contrario, existen importantes matizaciones en función de la localización geográfica.

Para ello se cuenta ya con la colaboración de diversos eventos nacionales así como una red de contactos internacionales para la difusión de la encuesta con soporte web.

En cuanto a los resultados obtenidos, es importante indicar que ninguno de los factores ha sido realmente rechazado (incluso los situados con menor porcentaje de confirmación rondan el 50% de las opiniones) sino que puede ser discutible su ocurrencia generalizada en todos los ámbitos de la muestra. Centrándonos en aquellos que cuentan con mayor confirmación como situaciones generalmente presentes en las organizaciones, podemos apreciar que son las relacionadas con la formación y cualificación en pruebas, la presión de calendario por su situación en el ciclo de vida y la falta de eficiencia e integración de las pruebas con el resto de modelos y productos del ciclo de vida las que reúnen mayor grado de confirmación y de influencia. Esta idea es coherente con el diagnóstico que el grupo de investigadores y expertos que lanzó REPRIS en 2005 realizó y consideró como motivación esencial para la creación de esta red.

## **Agradecimientos**

Los autores agradecen el apoyo de la Acción Especial TIN2007-67843-C06-01 para este trabajo.

## **Referencias**

1. De Baar, B., Software testing difficult?, <http://www.softwareprojects.org/software-testing.htm> (visitado 14-7-08) (2008)
2. Taipale, O.; Smolander, K.; Kalviainen, H.: Factors affecting software testing time schedule. In: Wohlin, C. (ed.), Proceedings of the 17th Australian Software Engineering Conference (ASWEC 2006), pp. 283-291, IEEE Computer Society (2006)
3. Black, R., Factors that Influence Test Estimation, StickyMinds.com, [http://www.stickyminds.com/s.asp?F=S5992\\_ART\\_2](http://www.stickyminds.com/s.asp?F=S5992_ART_2) (visitado 14-7-08) (2007)
4. Ng, S.P.; Murnane, T.; Reed, K.; Grant, D.; Chen, Y.: A Preliminary Survey on Software Testing Practices in Australia, In: Kurchten, P. (ed.) Proceedings of the 2004 Australian Software Engineering Conference (ASWEC'04), pp. 116-2251, IEEE Computer Society (2004)
5. Fernández-Sanz, L., Lara, P.: Un sondeo en España sobre la práctica profesional en procesos para pruebas de software. In: Tuya, J. (ed.) Taller PRIS'06 (Sitges, 3 de octubre). <http://in2test.lsi.uniovi.es/pris2006/PRIS2006-FernandezLara.pdf> (2006)
6. Fernández-Sanz, L., Lara, P.: Un experimento sobre hábitos de pruebas artesanales de software: Resultados y Conclusiones. In: Tuya, J. (ed.) Actas de Talleres de Ingeniería del Software y Bases de Datos, vol. 1, no. 4, pp. 23-30. Sistedes, Zaragoza (2007)

