

# Identificación del Modelo de Meta-Análisis para Agregación de Experimentos de Distinto Tamaño en Ingeniería del Software

Hernán Amatriain<sup>1,2</sup>, Rodolfo Bertone<sup>3</sup>, Enrique Fernández<sup>2</sup>, Eduardo Diez<sup>2</sup>

<sup>1</sup> Programa de Maestría en Ingeniería del Software. Facultad de Informática. Universidad Nacional de La Plata

<sup>2</sup> Laboratorio de Investigación y Desarrollo en Ingeniería de Software. Grupo de Investigación en Sistemas de Información. Universidad Nacional de Lanús  
<http://www.unla.edu.ar/sistemas/gisi/>

<sup>3</sup> Instituto de Investigaciones en Informática LIDI. Facultad de Informática. UNLP  
<http://weblidi.info.unlp.edu.ar/wp/>  
hamatriain@unla.edu.ar, enriquefernandez@educ.ar

**Resumen.** La síntesis de resultados consiste en integrar los resultados de un conjunto de experimentos, previamente identificados, en una medida resumen. Este tipo de procedimientos recibe el nombre de Agregación o Meta-Análisis. Existen dos estrategias a la hora de agregar un conjunto de experimentos: modelo de efecto fijo y modelo de efectos aleatorios. Los distintos autores no muestran una línea de trabajo unificada. Este hecho hace que sea necesaria la unificación de criterios para la realización de este tipo de trabajos. Por Tanto surge la necesidad de establecer un conjunto de recomendaciones o guías que permitan, a los investigadores en Ingeniería del Software, determinar bajo qué condiciones es conveniente desarrollar un Meta-Análisis mediante el modelo de efecto fijo y cuando utilizar el modelo de efectos aleatorios.

**Palabras clave.** Meta-Análisis. Modelo de efecto fijo. Modelo de efectos aleatorios. Ingeniería de software Experimental.

## 1. Introducción

La cantidad de estudios experimentales en Ingeniería de Software se ha incrementado en los últimos años [Sjoberg, 2005]. Uno de los principales problemas a los cuales se enfrentan los investigadores a la hora de generar evidencias empíricas, es la imposibilidad de desarrollar un experimento de gran envergadura. Si se deja de ver a los experimentos de manera aislada y los mismos se analiza en forma conjunta mediante su agregación, se pasa a tener conocimientos abalados por un mayor número de sujetos, lo cual mejora la calidad de la conclusión.

La síntesis de resultados [Borenstein et al, 2007] consiste en integrar los resultados de un conjunto de experimentos, previamente identificados, en una medida resumen. Existen dos estrategias a la hora de agregar un conjunto de experimentos: modelo de efecto fijo y modelo de efectos aleatorios [Borenstein et al, 2007].

Dentro de este contexto, los autores que han comenzado a trabajar concretamente en Meta-Análisis, no muestran una línea de trabajo unificada. Este hecho hace que sea necesaria la unificación de criterios para la realización de este tipo de trabajos.

## **2. Estado de la cuestión**

En este apartado se explican los inconvenientes de realizar un experimento de pequeño tamaño y la agregación como solución al problema. Se analizan las dos estrategias de agregación, los errores estadísticos a tener en cuenta y la discrepancia entre distintos autores a la hora de elegir una estrategia específica.

### **2.1. Experimentos y Agregación de Experimentos**

En los últimos años puede observarse el incremento de la cantidad de experimentos realizados dentro del ámbito de la Ingeniería del Software (IS) [Sjoberg, 2005]. Si bien los experimentos aportan conocimientos interesantes en cada caso, en general son pequeños (rara vez utilizan más de 20 sujetos experimentales [Davis et al, 2006]), por ello para que la información que aportan sea valiosa los resultados deben agregarse para poder obtener conclusiones avaladas con la mayor evidencia empírica posibles.

La agregación de experimentos o síntesis de resultados consiste en combinar los resultados de varios experimentos [Borenstein et al, 2007], que analizan el comportamiento de un par de tratamientos específico, en un contexto determinado, para obtener un único resultado final. El nuevo resultado será más general y fiable que los resultados individuales, por que el mismo estará sustentado por un mayor nivel de evidencia empírica [Cochrane, 2008].

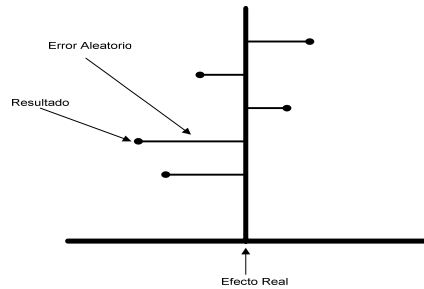
### **2.2. Modelo de efecto fijo y aleatorio**

Existen dos estrategias en la agregación de experimentos: modelo de efecto fijo y modelo de efectos aleatorios. Las Figura 1 y 2 ilustran ambas estrategias, mostrando en la Figura 1 un único tamaño de efecto y un conjunto de resultados que se diferencian del mismo únicamente por un error propio de la experimentación, y en la Figura 2 varios tamaños de efecto, donde los experimentos también poseen error experimental y un promedio general de los tamaños de efecto [Schmidt y Hunter, 2003].

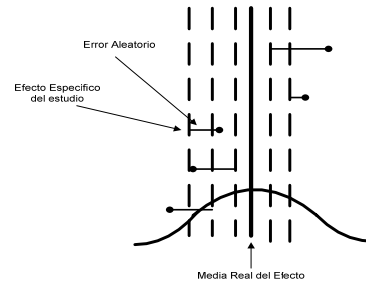
En el modelo de efecto fijo existe un único tamaño de efecto al cual pertenecen todos los experimentos. Cualquier diferencia en los resultados obedece únicamente a un error experimental aleatorio propio de la experimentación [Borenstein et al, 2007]. La ponderación de los experimentos se realiza en base a la inversa de su varianza.

Para el modelo de efectos aleatorios existe más de un tamaño de efecto, por ende existen dos tipos de errores, el error propio de cada uno de los experimentos producto de la experimentación y el error producido por la combinación de estudios

provenientes de distintos tamaños de efecto [Borenstein et al, 2007]. Este se traduce en dos tipos de varianzas, la interna de los estudios y la varianza entre estudios.



**Fig. 1.** Supuestos del modelo de efecto fijo



**Fig. 2.** Supuestos del modelo de efectos Aleatorios

### 2.3. Errores estadísticos

Los métodos estadísticos están sometido a dos tipos de errores [Cohen, 1988]:  $\alpha$ , o error de tipo I, y  $\beta$ , o error de tipo II. Dichos errores se producen por la incertidumbre asociada a estimar parámetros (básicamente medias y desvíos típicos) de una población a partir de una muestra de la misma. Tal y como indica la Tabla 1,  $\alpha$  es el error asociado a aceptar la hipótesis alternativa (H1) cuando en la población se verifica la hipótesis nula (H0); y  $\beta$  es la probabilidad asociada al evento justamente inverso.

**Tabla 1.** Tipos de error de un test estadístico

	H0 verificada en a población	H1 verificada en la población
H0 respuesta del experimento	Decisión correcta ( $1-\alpha$ )	$\beta$ (Tipo II error)
H1 aceptada respuesta del experimento	$\alpha$ (Tipo I error)	Decisión correcta ( $1-\beta$ )

### 2.4. Diferencia de Medias Ponderadas

Existen varias técnicas de agregación, sin embargo, en [Fernández et al, 2009] se comprobó que en contextos experimentales pocos maduros la técnica DMP ha demostrado mejor potencia y fiabilidad general que los otros métodos. El método Diferencias Medias Ponderadas (DMP) [Glass, 1976], es la técnica de estimación de tamaño de efecto (effect size), o mejora de un tratamiento respecto de otro.

El método consta de dos pasos: estimar el tamaño de efecto de cada experimento (Fórmula 1, [Hedges y Olkin, 1985]), y el error típico (Fórmula 2), y luego, estimar el tamaño de efecto global (Fórmula 3, [Borenstein et al, 2007; Hedges y Olkin, 1985]).

**Fórmula 1.** Tamaño de efecto individual con factor de corrección

$$d = J \frac{Y^E - Y^C}{S_p}$$

$d$  representa el tamaño de efecto  
 $J$  representa el factor de corrección  
 $Y^E$  representa a las medias del grupo experimental (E) y de control (C)  
 $S_p$  representa el desvío estándar conjunto

$$J = 1 - \frac{3}{4N - 9}$$

$N$  representa el total de sujetos experimentales incluidos en el experimento

**Fórmula 2.** Error típico e intervalo de confianza

$$v = \frac{\tilde{n} + d^2}{2(n^E + n^C)}$$

$v$  representa el error típico  
 $d$  representa el tamaño de efecto  
 $n^E$  representa la cantidad de sujetos experimentales del grupo experimental (E) y de control (C)

$$\tilde{n} = \frac{n^E + n^C}{n^E * n^C}$$

$Z$  representa la cantidad de desvíos estándar que separan, al nivel de significancia dado, la media del límite. En general es 1,96 ( $\alpha = 0,05$ )

$$d - Z_{\alpha/2} \sqrt{v} \leq \lambda \leq d + Z_{\alpha/2} \sqrt{v}$$

**Fórmula 3.** Tamaño de efecto grupal

$$d^* = \frac{\sum d_i / \sigma^2_i(d)}{\sum 1 / \sigma^2_i(d)}$$

$d^*$  representa el tamaño de efecto global  
 $\sum d_i / \sigma^2_i(d)$  es la sumatoria de los efectos individuales

$$v = (1 / \sum 1 / \sigma^2_i(d))$$

$\sum 1 / \sigma^2_i(d)$  es la sumatoria de la inversa varianza

Estimado el efecto global, se debe estimar el intervalo de confianza asociado.  
 Fórmula 4.

**Fórmula 4.** Intervalo de confianza grupal

$$d^* - Z_{\alpha/2} \sqrt{v} \leq \lambda \leq d^* + Z_{\alpha/2} \sqrt{v}$$

$d^*$  representa el tamaño de efecto global  
 $Z$  representa la cantidad de desvíos estándar que separan, al nivel de significancia dado, la media del límite. En general es 1,96 ( $\alpha = 0,05$ )  
 $v$  representa el error típico

Para el modelo de efectos aleatorios se utilizan las fórmulas 5 y 6 ([Borenstein et al, 2007; Hedges y Olkin, 1985]):

**Fórmula 5.** Tamaño de efecto global para el modelo de efectos aleatorios

$$\Delta = \frac{\sum d_i / \gamma^2_i}{\sum 1 / \gamma^2_i}$$

$\Delta$  representa el tamaño de efecto global  
 $\sum d_i / \gamma^2_i$  representa la sumatoria de los efectos individuales  
 $\sum 1 / \gamma^2_i$  representa la sumatoria de la inversa de las varianzas entre-estudios e intra-estudios

**Fórmula 6.** Error típico global e intervalo de confianza para el modelo de efectos aleatorios

$$\Delta - Z_{\alpha/2} \sqrt{v} \leq \lambda \leq \Delta + Z_{\alpha/2} \sqrt{v}$$

$\Delta$  representa el tamaño de efecto global  
 $Z$  representa la cantidad de desvíos estándar que separan, al nivel de significancia dado, la media del límite. En general es 1,96 ( $\alpha = 0,05$ )

$$v = \frac{1}{\sum 1 / \gamma^2_i}$$

### **3. Descripción del Problema**

El Meta-Análisis es hoy día una práctica conocida dentro de la Ingeniería de Software Empírica. Desde que fue propuesto por Basilli [1996] hubo varios trabajos abordados desde esta óptica [Hunt, 1997; Wohlin et al., 2000; Juristo et al., 2004; Jørgensen, 2004; Kitchenham, 2004; Dyba et al, 2007], sin embargo no existe un criterio unificado de que modelo de Meta-Análisis se debe aplicar hoy día en Ingeniería de Software.

Si se remite a la teoría estadística, esto debería hacerse en base al análisis de la heterogeneidad entre los experimentos, lo cual no es factible realizar debido a que las muestras de los experimentos son pequeñas y los Meta-Análisis agregan pocos experimentos. En este contexto, como se indica en [Hardy y Thompson 1998; Kim, 2000], la heterogeneidad no es medible ya que los test no tienen potencia.

No tener la fiabilidad deseada, implica caer en un error de Tipo I, y la falta de potencia lleva a caer en un error de Tipo II. En el primer caso, se estará afirmando que el tratamiento experimental supera al de control, cuando en realidad no lo hace, llegando a un resultado equivocado. En el segundo caso, se tiene un tratamiento experimental que supera al de control (que es lo que se quiere determinar) pero la experimentación no permite aseverarlo.

Al no existir un consenso sobre que modelo de Meta-Análisis utilizar, se deberá recurrir a la experiencia práctica, para lo que deberán probarse ambos modelos y analizar su comportamiento. Sin embargo, por tratarse de un contexto experimental poco maduro, no es posible contar con los experimentos necesarios para realizar el estudio, por lo que se recurrirá a realizar un proceso de simulación que permita validar el modelo adecuado a utilizar.

Se deberían realizar agregaciones con experimentos de igual tamaño y con experimentos de tamaños distintos. Esto lleva a considerar la existencia de dos escenarios bien definidos: (a) la agregación de experimentos realizados con distinta cantidad de sujetos (experimentos de distinto tamaño); y (b) la agregación de experimentos realizados con la misma cantidad de sujetos (experimentos de igual tamaño). En [Amatriain et al, 2011] se presentaron los resultados de un proceso de simulación que arroja luz sobre el primer escenario, por lo que en la presente investigación se estudiará el caso en que los experimentos a agregar son de distinto tamaño.

### **4. Solución Propuesta**

Por simulación de Monte Carlo se obtendrá la cantidad de estudios experimentales necesarios para poder llegar a una conclusión apoyada por una evidencia empírica adecuada (potencia y fiabilidad estadística de 80% y 95% respectivamente [Cohen, 1988]).

En cada corrida, se generan números aleatorios que representan las medias poblacionales de un tratamiento y otro (tratamiento de control y experimental). Para ello, se realiza la simulación suponiendo que el tratamiento experimental es mejor que el tratamiento de control en una medida igual al tamaño de efecto teórico a analizar.

Posteriormente se calculan la fiabilidad y potencia estadística.

Se considera el siguiente conjunto de variables independientes:

- Cantidad de sujetos por experimentos: entre 4 y 20. Se hará referencia a estas cantidades como baja o pocos (4-9), media (10-12) y alta o muchos (13-20).
- Cantidad de experimentos a agregar/combinar: se tomarán los valores 4, 14 y 20.
- Media poblacional del tratamiento de control: es fijada en 100.
- Desvío estándar teórico: 10% (bajo), 40% (medio) y 70% (alto).
- Tamaño de efecto poblacional teórico: bajo (0,2), medio (0,5), alto (0,8) y muy alto (1,2).

Se han tenido en cuenta el siguiente conjunto de variables dependientes e intermedias:

- Los tamaños de efecto poblacional medido o real, desvío estándar medido e intervalo de confianza (variables intermedias)
- La media poblacional del tratamiento experimental (variable intermedia).
- Potencia y fiabilidad deseadas (errores de tipo I y II, variables dependientes): la fiabilidad indican el porcentaje de veces que el intervalo de confianza estimado contuvo el valor del tamaño de efecto poblacional, la potencia estadística indican el porcentaje de veces que dicho intervalo de confianza no contuvo el valor 0 (cero).

En cada simulación se realizarán 1000 muestreos.

## 5. Resultados de la simulación

Los resultados de las simulaciones hechas para los modelos de efecto fijo y efecto aleatorio se sintetizan en las Tablas 2 y 3, donde se muestran los resultados de la fiabilidad y potencia para experimentos de distinto tamaño y distintos desvíos estándar. Se hace el análisis de fiabilidad para cada modelo y para cada cantidad de experimentos (la zona en gris indica los casos en que se alcanzó la fiabilidad del 95%):

- Modelo de efecto fijo y 3 experimentos (cuadro marcado en la tabla con una línea continua): se alcanza la fiabilidad para efecto pequeño diferencia de tamaño entre experimentos a agregar máxima.
- Modelo de efecto fijo y 6 experimentos (cuadro marcado en la tabla con una línea de puntos): se alcanza la fiabilidad propuesta para tamaño de efecto.
- Modelo de efecto fijo y 9 experimentos (cuadro marcado en la tabla con una doble línea continua): se alcanza fiabilidad en todos los casos de tamaño de efecto pequeño y para tamaño de efecto medio solo para los casos en que la diferencia de tamaño de experimentos a agregar es máxima.
- Modelo de efectos aleatorios y 3 experimentos (cuadro marcado en la tabla con una línea de segmentos): se presenta fiabilidad para tamaño de efecto pequeño.
- Modelo de efectos aleatorios y 6 experimentos (cuadro marcado en la tabla con una línea de segmentos y puntos): hay fiabilidad para efecto pequeño y medio.

**Table 2.** Cuadro de simulaciones para ambos modelos de efecto de Meta-Análisis, para comparación de fiabilidad con desvío estándar del 10%, 40% y 70%

Sujetos	Experi-mentos	Modelo de efecto fijo				Modelo de efectos aleatorios			
		Efecto 0,2	Efecto 0,5	Efecto 0,8	Efecto 1,2	Efecto 0,2	Efecto 0,5	Efecto 0,8	Efecto 1,2
4-4-14	3	94,3	93,3	91,2	83,8	96,2	87,7	78,0	72,6
4-4-20	3	98,4	98,6	97,9	89,4	99,6	93,4	78,3	68,5
4-14-14	3	91,3	88,2	81,5	70,4	96,4	90,7	85,2	82,0
4-14-20	3	92,1	90,2	80,6	70,3	99,8	98,0	86,7	83,2
4-20-20	3	96,6	93,3	84,5	73,9	97,3	93,0	81,9	76,3
14-14-20	3	91,4	87,1	80,2	71,5	98,9	86,2	83,3	84,0
14-20-20	3	94,1	90,2	82,3	67,8	96,6	88,5	84,8	89,6
4-4-4-14-14	6	97,5	93,0	85,3	70,3	100	98,6	96,0	93,1
4-4-4-4-20-20	6	97,1	92,8	82,1	65,5	100	99,9	96,3	92,7
4-4-14-14-14	6	96,2	92,5	82,5	65,3	99,1	92,4	86,8	86,2
4-4-14-14-20-20	6	94,1	90,2	76,6	59,6	100	98,0	87,7	86,9
4-4-20-20-20	6	100	98,3	81,8	60,7	100	99,0	94,6	92,4
14-14-14-14-20-20	6	97,1	91,2	74,1	57,1	100	98,0	94,4	96,2
14-14-20-20-20	6	99,1	91,7	78,0	57,5	100	93,9	89,1	92,8
4-4-4-4-4-14-14-14	9	96,6	90,9	79,4	65,1	100	99,5	98,2	96,0
4-4-4-4-4-20-20-20	9	100	97,5	83,2	61,6	100	100	100	97,4
4-4-4-14-14-14-14-14	9	99,9	91,8	72,7	53,6	100	99,2	97,1	97,0
4-4-4-14-14-14-20-20-20	9	98,5	91,5	73,5	55,7	100	99,7	96,1	91,8
4-4-4-20-20-20-20-20	9	100	95,3	74,2	51,3	100	99,3	97,1	95,6
14-14-14-14-14-14-20-20-20	9	99,8	88,6	72,6	47,9	100	99,9	96,8	97,9
14-14-14-20-20-20-20-20	9	99,8	91,4	68,4	53,9	100	98,2	93,7	94,1

**Table 3.** Cuadro de simulaciones para ambos modelos de efecto de Meta-Análisis, para comparación de potencia con desvío estándar del 10%, 40% y 70%

Sujetos	Experi-mentos	Modelo de efecto fijo				Modelo de efectos aleatorios			
		Efecto 0,2	Efecto 0,5	Efecto 0,8	Efecto 1,2	Efecto 0,2	Efecto 0,5	Efecto 0,8	Efecto 1,2
4-4-14	3	7,3	25,1	49,1	84,4	0	0,2	0,9	2,7
4-4-20	3	4,3	26,4	67,6	96,0	0	0,1	0,9	2,5
4-14-14	3	10,9	33,4	58,6	90,9	0	0,1	0	0,9
4-14-20	3	8,5	29,3	70,5	94,0	0	0	0	0,4
4-20-20	3	9,0	38,4	79,1	98,6	0	0,5	0,8	0,4
14-14-20	3	17,2	45,4	81,0	98,9	0	0,1	0,7	0,8
14-20-20	3	13,8	54,4	87,4	98,8	0	1,5	1,4	0,4
4-4-4-14-14	6	7,8	35,3	78,6	98,7	0	0	0	0,2
4-4-4-4-20-20	6	6,7	46,1	88,5	99,1	0	0	0,1	0,7
4-4-14-14-14	6	10,6	61,4	92,2	99,7	0	0,3	0,6	0,5
4-4-14-14-20-20	6	17,1	64,8	93,8	99,9	0	0	0,9	0,5
4-4-20-20-20	6	3,4	82,8	99,3	100	0	0,1	0,3	0,5
14-14-14-14-20-20	6	16,7	78,1	99,3	100	0	0	0,1	0,1
14-14-20-20-20	6	10,9	87,5	99,6	100	0	0	0,4	0,3
4-4-4-4-4-14-14-14	9	7,3	55,7	91,5	99,7	0	0	0	0,8
4-4-4-4-4-20-20-20	9	7,4	78,0	99,4	100	0	0	0	0,2
4-4-4-14-14-14-14-14	9	10,4	73,5	98,4	99,9	0	0	0,1	0,2
4-4-4-14-14-14-20-20-20	9	18,4	88,3	99,9	100	0	0	0,1	0,8
4-4-4-20-20-20-20-20	9	7,4	96,3	100	100	0	0	0,1	0,2
14-14-14-14-14-14-20-20-20	9	16,5	93,6	100	100	0	0	0,1	0,1
14-14-14-20-20-20-20-20-20	9	27,5	98,6	100	100	0	0	0,2	0,2

- Modelo de efectos aleatorios y 9 experimentos (cuadro marcado en la tabla con una triple línea continua): se alcanza la fiabilidad en casi todos los casos.

Se puede resumir la tendencia observada como sigue:

- Modelo de efecto fijo: a medida que se incrementa la cantidad de experimentos y el tamaño de efecto es pequeño, se incrementa la fiabilidad. También se observa que mejora la potencia cuando la diferencia de tamaño de experimentos a agregar es máxima (los casos de combinar experimentos de 4 y 20 sujetos).
- Modelo de efectos aleatorios: la fiabilidad se incrementa al subir la cantidad de experimentos y (el desplazamiento sobre la tabla es de cuadro a cuadro de arriba hacia abajo) y crecer el tamaño de efecto (el desplazamiento en la tabla en cada cuadro es de izquierda a derecha).

Análisis de potencia (la zona en gris indica que se alcanzó la potencia del 80%):

- Modelo de efecto fijo y 3 experimentos (cuadro marcado en la tabla con una línea continua): se alcanza la potencia propuesta para efecto muy grande, o para efecto grande si la cantidad total de sujetos es elevada
- Modelo de efecto fijo y 6 experimentos (cuadro marcado en la tabla con una línea de puntos): hay potencia cuando el efecto es grande y muy grande. También hay potencia para efecto medio si la cantidad de sujetos totales es alta.
- Modelo de efecto fijo y 9 experimentos (cuadro marcado en la tabla con una doble línea continua): se alcanza la potencia del 80% para tamaño de efecto grande y muy grande. Para tamaño de efecto medio, se alcanza la potencia mínima deseada para una cantidad de sujetos totales media/alta.
- Para el modelo de efectos aleatorios no se observa potencia en ningún caso

Analizando la tendencia observada en la Tabla 3, puede resumirse:

- Modelo de efecto: la potencia se incrementa al subir la cantidad de experimentos (desplazamiento en la tabla de arriba hacia abajo de cuadro a cuadro) y al crecer el tamaño de efecto (desplazamiento dentro de cada cuadro de izquierda a derecha).
- Para el modelo de efectos aleatorios no se observa potencia en ningún caso

## 6. Conclusión

Los resultados obtenidos en la sección anterior, deben agruparse bajo un criterio que haga simple su análisis visual, para lo cual se diseñó la Tabla 4. Cuando se hace referencia a baja o alta potencia y a baja o alta fiabilidad se estará indicando que se alcanzó el 95% de fiabilidad y 80% de potencia estadística. La alta fiabilidad se indica con (+) y la baja con espacio en blanco. En tanto que la alta potencia se indica con (//) y la baja potencia con espacio en blanco. De esta manera, pueden agruparse los resultados similares en una tabla, donde no se indica la potencia y fiabilidad alcanzada, sino simplemente si se llegó a la cota superior propuesta como objetivo.



**Table 4.** Cuadro comparativo de ambos modelos para tamaños de distinto tamaño.

Sujetos	Experi- mentos	Modelo de efecto fijo				Modelo de efectos aleatorios			
		Efecto 0,2	Efecto 0,5	Efecto 0,8	Efecto 1,2	Efecto 0,2	Efecto 0,5	Efecto 0,8	Efecto 1,2
4-4-14	3				//	+			
4-4-20	3	+	+	+	//	+			
4-14-14	3				//	+			
4-14-20	3				//	+	+		
4-20-20	3	+			//	+			
14-14-20	3				//	+			
14-20-20	3				//	+			
4-4-4-4-14-14	6	+			//	+	+	+	
4-4-4-4-20-20	6	+		//	//	+	+	+	
4-4-14-14-14-14	6	+		//	//	+			
4-4-14-14-20-20	6			//	//	+	+		
4-4-20-20-20-20	6	+	//	//	//	+	+		
14-14-14-14-20-20	6	+		//	//	+	+		
14-14-20-20-20-20	6	+	//	//	//	+			
4-4-4-4-4-4-14-14-14	9	+		//	//	+	+	+	+
4-4-4-4-4-20-20-20	9	+	+	//	//	+	+	+	+
4-4-4-14-14-14-14-14-14	9	+		//	//	+	+	+	+
4-4-4-14-14-14-20-20-20	9	+	//	//	//	+	+	+	+
4-4-4-20-20-20-20-20-20	9	+	+	//	//	+	+	+	+
14-14-14-14-14-14-20-20-20	9	+	//	//	//	+	+	+	+
14-14-14-20-20-20-20-20-20	9	+	//	//	//	+	+		

En forma de cuadro (Tabla 5), se resumen las recomendaciones a las que se arriban en la presente investigación, y que busca ser una herramienta de apoyo al investigador que se dedica al campo de la Ingeniería de Software Empírica.

**Table 5.** Recomendaciones para elección del modelo de Meta-Análisis

Características de la experimentación					Recomendación
Tamaño de efecto	Desvío estándar	Sujetos experimentales	Cantidad de experimentos	Modelo recomendado	Nota a la recomendación
Pequeño	Indiferente	Indiferente	>6	FIJO	Fiabilidad deseada y potencia baja
Indiferente	Indiferente	Indiferente	>6	NINGUNO (ALEATORIO)	Fiabilidad deseada y potencia NULA
Pequeño	Indiferente	Indiferente	Indiferente	NINGUNO (ALEATORIO)	Fiabilidad deseada y potencia NULA

## 7. Financiamiento

Las investigaciones que se reportan han sido financiadas parcialmente con presupuesto de la Licenciatura en Sistemas del Departamento de Desarrollo Productivo y Tecnológico de la Universidad Nacional de Lanús (Argentina).

## 8. Bibliografía

- Amatriain, Dieste, Fernández, García-Martínez, 2011: evaluación de la aplicabilidad del Meta-Análisis de efectos aleatorios en Ingeniería de Software, Proceedings XV Congreso Argentino de Ciencias de la Computación. Workshop de Ingeniería de Software
- Basili, Green, Laitenberger, Lanubile, Shull, Sörumgård y Zelkowitz, 1996: *The empirical investigation of perspective-based reading*, International Journal on Empirical Software Engineering, Vol. 1, No. 2; pp. 133–164.
- Borenstein, Hedges y Rothstein, 2007: *Meta-Analysis Fixed Effect vs. random effect*; www.Meta-Analysis.com
- Ciolkowski, 2009: *What do we know about perspective-based reading? An approach for quantitative aggregation in software engineering*; 3rd International Symposium on Empirical Software Engineering and Measurement, pp. 133-144.
- Cochran, 1954: *The combination of estimates from different experiments*; Biometrics, 10, 101–129.
- Cohen, 1988: *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.); ISBN 0-8058-0283-5.
- Davis, Dieste, Hickey, Juristo y Moreno, 2006: *Effectiveness of Requirements Elicitation Techniques: Empirical Results Derived from a Systematic Review*; 14th IEEE International Requirements Engineering Conference (RE'06) pp. 179-188
- Dieste y Griman, 2007: *Developing Search Strategies for Detecting Relevant Experiments for Systematic Reviews*; IEEE Press
- Dyba, Aricholm, Sjöberg, Hannay y Shull, 2007: *Are two heads better than one? On the effectiveness of pair programming*. IEEE Software;12-15.
- Dyba, Kampenes y Sjöberg, 2006: *A systematic review of statistical power in software engineering experiments*; Information and Software Technology; vol. 48, ejemplar 8, página 745-755
- Fernández, Pollo, Amatriain, Dieste, Pesado y García-Martínez, 2009: *Aplicabilidad de los Métodos de Síntesis Cuantitativa de Experimentos en Ingeniería de Software*. Proceedings XV Congreso Argentino de Ciencias de la Computación. Workshop de Ingeniería de Software. Págs. 752-761. ISBN 978-897-24068-4-1.
- Glass, 1976: *Primary, secondary, and meta-analysis of research*; Educational Researcher 5: 3-8
- Hedges y Olkin, 1985: *Statistical methods for meta-analysis*. Academic Press
- Hunt, 1997: *How Science takes stock: the story of meta-analysis*; Russell Sage Foundation: New York
- Jørgensen, 2004: *A Review of Studies on Expert Estimation of Software Development Effort*; Journal of Systems and Software; (70): 1-2, pp. 37-60.
- Juristo, Moreno y Vegas, 2004: *Towards building a solid empirical body of knowledge in testing techniques*; Acm Sigsoft Software Engineering Notes (Sigsoft) 29(5):1-4
- Kitchenham, 2004: *Procedures for performing systematic reviews*. Keele University; TR/SE-0401. Keele University Technical Report.
- Schmidt y Hunter, 2003: *Handbook of Psychology, Research Methods in Psychology*; Chapter 21, “Meta-Analysis”; Schinka, J., Velicer, W., Weiner, I. Editors, Volume 2.
- Sjöberg, 2005: *A survey of controlled Experiments in Software Engineering*; IEEE Transactions on Software Engineering; Vol 31 Nro. 9.
- Wohlin, Runeson, Håst, Ohlsson, Regnell y Wesslén, 2000: *Experimentation in Software engineering: An Introduction*; International Series in Software Engineering; id: 29; Record: 5370; Volume: 6