

Influencia de la Teoría de Roles de Belbin en la Medición de Software: Un estudio exploratorio

Raúl Aguilar Vera¹, Julio Díaz Mendoza¹, Juan Ucán Pech¹

avera@correo.uady.mx, julio.diaz@correo.uady.mx, juan.ucan@correo.uady.mx

¹ Facultad de Matemáticas, Universidad Autónoma de Yucatán, C.P. 97000, Mérida, Yucatán, México.

DOI: 10.4304/risti.31.50-65

Resumen: Se presenta un estudio exploratorio en el que se analiza la influencia de la Teoría de Belbin en actividades de desarrollo de software como mecanismo para la mejora del mismo. El estudio comprende un primer experimento controlado con estudiantes y una réplica similar-interna. La actividad considerada para los equipos, fue la medición de software mediante la técnica de puntos de función. Los resultados obtenidos en el primer experimento no fueron concluyentes debido a que no fueron confirmados en la réplica, no obstante, las lecciones aprendidas con el estudio, permiten identificar y proponer varios experimentos controlados como trabajos futuros.

Palabras-clave: Experimentación en Ingeniería de Software; Integración de Equipos; Medición de Software; Roles de Belbin.

Influence of Belbin's Role Theory in Software Measurement: An Exploratory Study

Abstract: An exploratory study is presented in which the influence of the Belbin Theory on software development activities as a mechanism to improve it is analyzed. The study includes a first controlled experiment with students and a replication. The activity considered for the teams, was the measurement of software using the function point technique. The results obtained in the first experiment were not conclusive because they were not confirmed in the experimental replica, however, the lessons learned with the study allow to identify and propose several controlled experiments as future works.

Keywords: Belbin Roles; Experimentation in Software Engineering; Software Measurement; Teams Integration.

1. Introducción

El desarrollo de software, desde el momento mismo de su concepción como disciplina ingenieril, ha mantenido una dinámica de mejora continua en cuanto a sus métodos, técnicas, herramientas y prácticas. Medio siglo después de la denominada “Crisis del Software” la Ingeniería de Software (IS) cuenta con un cuerpo de conocimientos

aceptado en lo general, por profesionistas e investigadores de la disciplina (Bourque & Fairley, 2014), sin embargo, la dinámica de mejora en cuanto a calidad, tanto en sus procesos, como en los artefactos generados a través de los mismos, sigue siendo fuente de investigación para dicha disciplina ingenieril (p.e. Miramontes et al, 2016; Norambuena y Vega, 2017).

El estudio de los enfoques vinculados con el desarrollo, operación y mantenimiento del software, ha sido analizado con base en la relación de diversas variables, no obstante, el aspecto social intrínseco de la disciplina (Juristo & Moreno, 2001) ha llevado a considerar al factor humano, como un aspecto de singular interés para su investigación. En un estudio reciente, Morales y Vega (2018) proponen un catálogo de factores humanos críticos para el éxito de propuestas de mejora al proceso software; uno de los cuatro factores se relaciona con el rol que desempeña un Ingeniero de Software al interior del equipo de trabajo; Humphrey (2000) afirma que el proceso de formación y creación de un equipo de desarrollo de software no se produce por accidente, si no que el equipo debe establecer relaciones de trabajo, ponerse de acuerdo sobre los objetivos y determinar las funciones de los miembros del grupo. En la última década, incluso elementos de gamificación han sido utilizados como estrategia para integrar equipos de trabajo (Hernández et al, 2016).

El presente estudio, tiene como propósito explorar si el uso de la teoría de roles de Belbin (1993) para la formación de equipos, tiene alguna influencia en una de las tareas vinculadas con los procesos de gestión de software, particularmente, en la tarea de medición. El documento se encuentra organizado de la siguiente manera: la sección 2 presenta trabajos relacionados con el estudio, la sección 3 describe brevemente la técnica de medición de software conocida como puntos de función. La sección 4 presenta un primer experimento realizado. La sección 5 describe una réplica experimental. Finalmente, las conclusiones se presentan en la sección 7.

2. Trabajos relacionados con es estudio

Son varios los investigadores (Belbin, 1981; Briggs-Myers & McCaulley, 1985; Margerison & McCan, 1985; Mumma, 2005) que afirman haber identificado roles que definen el comportamiento de los individuos al interior de un equipo de trabajo, dichos roles son conocidos bajo el nombre de roles de equipo, y aunque no existe el reconocimiento de la existencia de una relación funcional con algún proceso o tarea particular, su ausencia o presencia, al parecer tiene una influencia significativa en los logros del equipo (Aritzeta, Swailes & Senior, 2005).

Entre los estudios existentes sobre roles de equipo, el trabajo de Belbin (1981, 1993) es probablemente el más recurrido entre consultores e investigadores, su popularidad radica en que además de ofrecer una categorización de roles (Impulsor:IM, Implementador:IP, Finalizador:F, Cerebro:C, Monitor-Evaluador:ME, Especialista:E, Coordinador:CO, Investigador de Recursos:IR y Cohesionador:CH), presenta recomendaciones para la integración de equipos con base en roles compatibles, así como mecanismos para la identificación de los roles que un individuo puede asumir en un equipo de trabajo, en función de su comportamiento. Johansen (2003) en su disertación sobre la gestión del recurso humano, validó el inventario de autopercepción propuesto por Belbin para la

identificación del rol natural o roles que un individuo puede asumir en un equipo de trabajo, y entre sus conclusiones, indicó que dicho inventario es definitivamente una herramienta que vale la pena utilizar para evaluar la composición y posible desempeño del equipo.

En cuanto a las investigaciones sobre el uso de la teoría de Belbin en el contexto de la Ingeniería de Software, se encuentra el reporte de Henry y Stevens (1999) cuyo esfuerzo de investigación se centró en el análisis de la utilidad de los roles de Belbin para mejorar la efectividad de los equipos, en particular, la influencia del liderazgo en el grupo; dichos autores comentan que los roles de Belbin se pueden utilizar en la formación de nuevos equipos, así como en la evaluación de los equipos existentes, y luego del desarrollo de un experimento controlado con estudiantes, concluyen que un solo líder mejora el rendimiento de un equipo, en comparación con equipos que tienen múltiples líderes o ningún líder. Considerando también a la efectividad como el aspecto a mejorar, Pollock (2009) exploró la influencia de contar con una diversidad de roles y personalidades, como estrategia para mejorar la efectividad del equipo, y concluye, que aunque no encontró evidencia de que dicha diversidad tiene influencia en el equipo, pudo observar que la presencia de ciertos roles como son el caso del Impulsor, Coordinador o Finalizador, pueden incrementar la efectividad.

Otros estudios se han centrado en analizar la relación de los roles de Belbin con tareas vinculadas con el proceso software. Estrada y Peña (2013) reportan un experimento controlado con estudiantes en ciertas actividades vinculadas con las fases de requisitos, diseño y codificación; las autoras concluyen que algunos roles presentan mayor aportación en ciertas actividades, particularmente señalan al rol Implementador en la tarea de codificación. Así mismo, en un estudio previo en el contexto de la fase de codificación, comparamos la calidad de la legibilidad del código generado por equipos integrados con roles compatibles, de acuerdo con la Teoría de Belbin, en contraste con equipos formados por integrantes asignados aleatoriamente, los resultados obtenidos mediante un experimento controlado con estudiantes, nos permitieron concluir que los equipos que fueron integrados utilizando como referencia la Teoría de Belbin, presentan resultados significativamente mejores que los equipos integrados mediante un criterio de aleatoriedad (Aguileta, Ucán y Aguilar, 2017).

3. Metodología

Aunque el cuerpo de conocimientos de la IS comenzó a integrarse a finales de la década de los sesenta, fue hasta la década de los ochenta cuando la comunidad académica comenzó a adoptar y utilizar enfoques de investigación para estudiar de manera rigurosa los diferentes aspectos y problemas involucrado con el proceso de desarrollo de software (Glass, Vessey & Ramesh, 2002).

Una de las metodologías más utilizadas en el campo de la IS, es la experimentación, en particular, la experimentación en entornos controlados; el énfasis en aplicar la experimentación en IS se remonta a la década de los años ochenta (Basili, Selby & Hutchens, 1986) y tiene como objetivo identificar las causas por las cuales se producen ciertos resultados, su aplicación nos ayuda a identificar y comprender las posibles relaciones entre factores y variables dependientes, ambos elementos involucrados en los

procesos de desarrollo de software. Entre los elementos característicos de los estudios exploratorios encontrados en la literatura, se destaca el uso de grupos de estudiantes como sujetos experimentales; Genero, Cruz-Lemus y Piattini (2014) indican al respecto, que dicha muestra académica permite al investigador obtener evidencia preliminar para confirmar o refutar hipótesis que pueden ser contrastadas posteriormente en contextos industriales. Por otro lado, un elemento clave en el paradigma experimental, es el desarrollo de réplicas, su propósito radica en verificar o refutar los hallazgos previamente observados (Gómez, Juristo & Vegas, 2014).

El presente estudio tiene como propósito explorar, mediante la ejecución de un experimento controlado con estudiantes, y una réplica experimental similar-interna, la influencia del uso de la Teoría de Belbin en la integración de equipos de desarrollo de software con integrantes que presentan roles compatibles, en el producto de una actividad, así como en el tiempo invertido en dicha actividad; en el caso del estudio se seleccionó una actividad vinculada con las tareas de gestión, particularmente, la medición del tamaño lógico o funcional de un sistema de software propuesto.

3.1. Factor y Tratamientos

El factor considerado para nuestro estudio, es la estrategia utilizada para la integración de los equipos de trabajo, particularmente, con dos tratamientos:

Equipos Compatibles (EC): Equipos de trabajo de tres integrantes, conformados con roles compatibles —identificados con el inventario de autopercepción— de acuerdo con la Teoría de Belbin.

Equipos Tradicionales (ET): Equipos de trabajo de tres integrantes, seleccionados aleatoriamente para su conformación.

3.2. Hipótesis y Variables Dependientes

De acuerdo con el propósito de nuestro estudio, el objetivo planteado para el experimento así como para su réplica, consiste en comparar métricas —del producto y del proceso— obtenidas mediante equipos integrados con roles compatibles —con base en la Teoría de Belbin— con las obtenidas por equipos integrados en forma aleatoria.

El primer par de hipótesis estadísticas utiliza como variable dependiente, una métrica del producto, es decir, el número de puntos de función sin ajustar obtenidos (PFSA) por los equipos al medir el sistema (Garmus & Herron, 1996).

- H01: La media de los PFSA obtenidos con la Técnica de Puntos de Función por los ET es igual a la media de los PFSA obtenidos por los EC.
- H11: La media de los PFSA obtenidos con la Técnica de Puntos de Función por los ET difiere de la media de los PFSA obtenidos por los EC.

En cuanto a la variable vinculada con el proceso, utilizaremos el tiempo en minutos (TM) utilizado para concluir la actividad, las hipótesis estadísticas planteadas son las siguientes:

- H02: El tiempo medio registrado por los ET en la actividad de medición es igual al reportado por los EC.

- H12: El tiempo medio registrado por los ET en la actividad de medición difiere del reportado por los EC.

3.3. Parámetros

Aspectos como la complejidad del problema a resolver, la presión del tiempo disponible para concluir la actividad, así como la experiencia de los participantes, se consideran parámetros que no afectan o sesgan los resultados del estudio, debido a que son parámetros que se mantuvieron homogéneos para todos los equipos de desarrollo.

En el caso de la experiencia de los sujetos, al ser equipos integrados por estudiantes que se encuentran en su proceso de formación y que antes del experimento, aún no habían estudiado la técnica, podemos considerar que son homogéneos —la técnica fue analizada con todos los estudiantes en una sesión de trabajo. Por su parte, en lo referente a la complejidad del sistema por estimar, se redactó un documento de Especificación de Requisitos para un Software Educativo, con base en el estándar IEEE Std. 830 1998, dicho documento fue el mismo que se entregó a todos los equipos en la tarea. Finalmente, en cuanto al tiempo disponible para la realización de la actividad, se realizó una estimación del tiempo requerido para la actividad, y considerando un estimado de 120 minutos, se dispuso de dos sesiones presenciales (180 minutos) para la actividad de medición.

3.4. Diseño Experimental

De acuerdo con las características del estudio, el diseño más apropiado para el experimento es el diseño factorial con una fuente de variación (ver tabla 1), la variable dependiente corresponde, en el primer par de hipótesis, a la métrica obtenida por los equipos en la tarea (PFSA) y en la segunda, al tiempo utilizado por los equipos para su finalización (Tiempo en segundos). En ambos experimentos, los tratamientos experimentales se corresponden con: (1) equipos integrados con base en la teoría de los roles de Belbin (Equipos Belbin) y (2) equipos integrados de manera forma aleatoria (Equipos Tradicionales).

	Tratamientos	Equipos Integrados con Estudiantes
Factor	<i>Equipos Belbin</i>	<i>Mediciones sobre: PFSA/Tiempo</i>
	<i>Equipos Tradicionales</i>	<i>Mediciones sobre: PFSA/Tiempo</i>

Tabla 1 – Diseño factorial con una fuente de variación

3.5. Análisis estadístico

El análisis de los datos que serán recolectados en ambos experimentos, contendrá una sección descriptiva a partir de la generación de gráficos de caja y bigotes, y en el caso del análisis inferencial, se utilizará el modelo estadístico del análisis de varianza simple. Los gráficos de caja y bigotes permiten generar una presentación visual en la que se pueden observar diversos datos de los conjuntos de mediciones por analizar: valor mínimo y máximo, mediana, cuartil inferior y superior, así como valores atípicos; así mismo, permiten observar la forma y dispersión de los conjuntos de datos, pero sobre todo, permiten evaluar empíricamente el comportamiento de los conjuntos de mediciones por comparar.

Para realizar el análisis inferencial, elegimos el análisis de varianza de una vía, debido a que nos permite realizar pruebas de hipótesis para determinar si hay o no diferencias significativas entre las medias y varianzas, de los valores recogidos en la variable dependiente, para los diferentes niveles del factor (tratamientos). El análisis de la varianza (Gutierrez y de la Vara, 2012) permite construir con los datos, un modelo estadístico que describe el impacto de un solo factor categórico sobre una variable dependiente, dicho modelo lineal asociado es el siguiente:

$$Y_{ij} = \mu + \beta_i + \epsilon_{ij};$$

donde Y_{ij} es la ij -ésima medición, μ es un parámetro común para todos los tratamientos, denominado media global, β_i es un parámetro que mide el efecto del i -ésimo tratamiento, y ϵ_{ij} es el error sistemático atribuible a la medición Y_{ij} .

4. Primer Experimento

4.1. Participantes/Sujetos Experimentales

La muestra por conveniencia utilizada para el experimento estuvo conformada por 27 estudiantes de carrera de Ingeniero de Software de la Universidad Autónoma de Yucatán, los participantes cursaban la asignatura “Diseño de Experimentos en Ingeniería de Software” ubicada en el quinto semestre de la carrera. Con los 27 participantes, se formaron 9 equipos de trabajo —sujetos experimentales. La tabla 2 ilustra la asignación de los equipos a los tratamientos.

Tratamientos	Equipos
EB	I, II, III, IV, V
ET	VI, VII, VIII, IX

Tabla 2 – Asignación de sujetos experimentales por tratamiento

La integración de los Equipos Compatibles, con base en los roles identificados mediante el inventario de autopercepción, se ilustra en la figura 1.

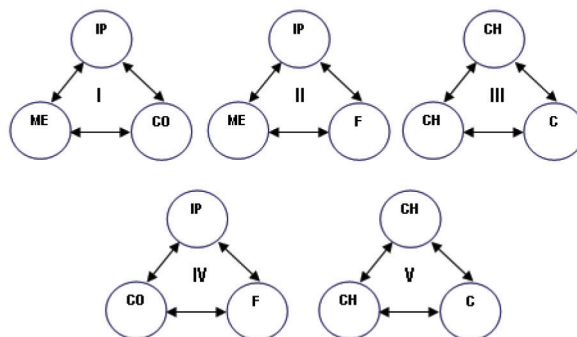


Figura 1 – Equipos integrados con la Teoría de Belbin (Equipos Compatibles)

4.2. Ejecución del Experimento

El experimento se llevó a cabo en tres sesiones de trabajo, en la primera, se administró el inventario de autopercepción, a partir del cual se identificaron los roles primarios de los participantes; posteriormente, en una segunda sesión que duró dos horas, los participantes recibieron capacitación sobre la técnica del análisis de puntos de función; finalmente, la sesión experimental fue realizada.

Al inicio de la sesión experimental, se integraron los nueve equipos, de acuerdo con lo indicado en la tabla 2, se les proporcionó el documento de Especificación de Requisitos de un caso, así como una hoja de informe con el esquema de la tabla 1; posteriormente, se describió la actividad y se les pidió que identificaran en la hoja de informe: el número de equipo, hora de inicio y de finalización de la actividad, así como las métricas correspondientes a la técnica. La sesión transcurrió sin problemas.

4.3. Análisis y Resultados del Experimento

En esta sección se presenta tanto el análisis descriptivo, como el análisis inferencial con base en las métricas seleccionadas para el experimento.

Primeramente se analizan los gráficos de Caja y Bigotes generados para las dos variables dependiente consideradas, a efecto de tratar de identificar intuitivamente el comportamiento de los datos (ver figura 2).

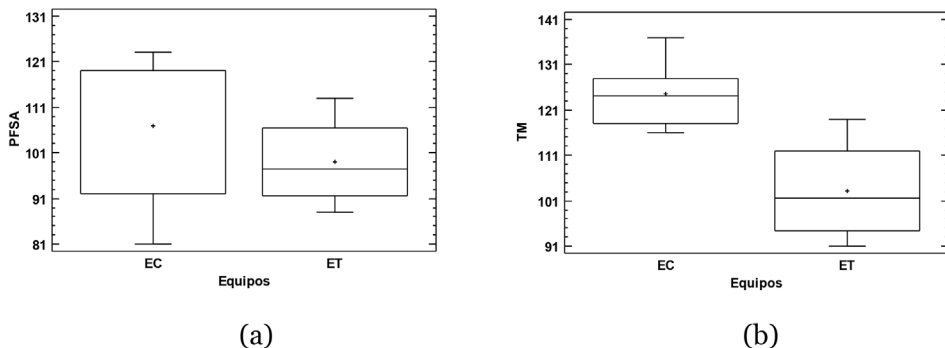


Figura 2 – Gráficos de Caja y Bigotes para las variables PFSA (a) y TM (b)

Como se observa en el gráfico de la figura 2(a), en cuanto a la variable PFSA, los EC presentan mayor dispersión y un sesgo negativo respecto de su mediana, por su parte los ET tienen un comportamiento más simétricos; si observamos ambas gráficas al mismo tiempo, vemos que el área de la caja de los ET se encuentra incluida en el área de los EC, lo cual nos da indicios de que no haber diferencias en cuanto a la variable PFSA. Por su parte, la figura 2(b) ilustra el comportamiento de los datos vinculados con la variable TM, observamos que los ET presentan mayor dispersión que los EC, sin embargo, dado que las áreas de ambas cajas se encuentran desfasadas, podemos suponer que el factor tiene cierta influencia.

Con el objeto de evaluar estadísticamente las diferencias observadas en el análisis descriptivo previo, se plantearon las siguientes hipótesis estadísticas.

$$H_{0_{PFSA}}: \mu_{EC} = \mu_{ET} \quad ; \quad H_{1_{PFSA}}: \mu_{EC} <> \mu_{ET} \quad (I)$$

$$H_{0_{Time}}: \mu_{EC} = \mu_{ET} \quad ; \quad H_{1_{Time}}: \mu_{EC} <> \mu_{ET} \quad (II)$$

El resultado de evaluar con ANOVA ambas hipótesis, se ilustra en la tabla 3.

Métrica	F	p-valor
<i>PFSA</i>	0.53	0.4894
<i>TM</i>	9.96	0.0160 *

*p-valor < 0.05

Tabla 3 – ANOVA para las variable PFSA y TM

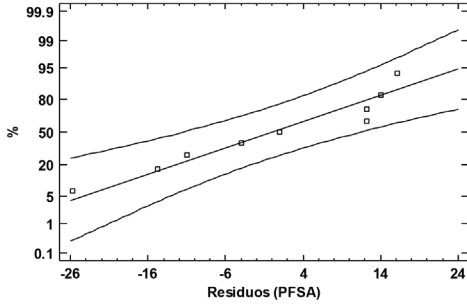
La tabla ANOVA descompone la varianza de la variable en estudio en dos componentes: un componente intergrupo y un componente en grupos. Dado que el valor p de la prueba F —en la variable PFSA— es mayor que 0.05 (PFSA: 0.4894) la hipótesis de nulidad se acepta, es decir, se afirma que no existe diferencias entre las medias de los dos tratamientos.

Por su parte, el valor P de la prueba F —en la variable TM— es menor que 0.05 (TM: 0.016), entonces podemos rechazar la hipótesis de nulidad y podemos afirmar que existe una diferencia estadísticamente significativa entre las medias de la variable TM de los dos tratamientos.

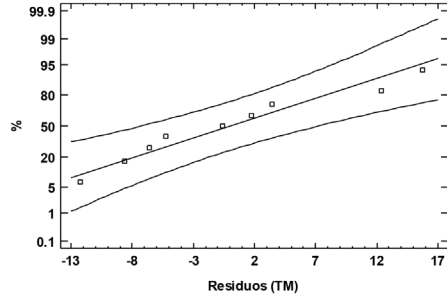
El modelo ANOVA tiene asociado tres supuestos que es necesario validar antes de utilizar la información que nos ofrece; los supuestos del modelo son: (1) Los errores experimentales de sus datos se distribuyen normalmente, (2) No existe diferencia entre la varianza de los tratamientos (Homocedasticidad) y (3) existe independencia entre las muestras.

Para validar el primer supuesto, usaremos el gráfico de probabilidad normal de los residuos. Es una técnica gráfica para evaluar si un conjunto de datos se distribuye de acuerdo con la distribución normal. Como se puede ver en la gráfica que se ilustra en la figura 3, los puntos, en ambas gráficas, no muestran desviaciones de la diagonal, por lo que es posible asumir que los residuos tienen una distribución normal en ambos casos.

En el caso de la homocedasticidad, generamos un gráfico de residuales vs. predichos y observamos si es posible detectar que el tamaño de los residuales aumenta o disminuye sistemáticamente a medida que aumenta los valores predichos. Como podemos ver en los dos gráficos de la figura 4, no se observa patrón particular alguno, con el cual podamos aceptar que se cumple la hipótesis de varianza constante de los residuos.

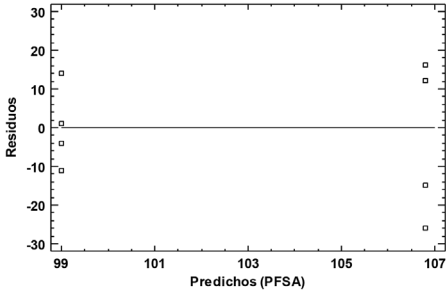


(a)

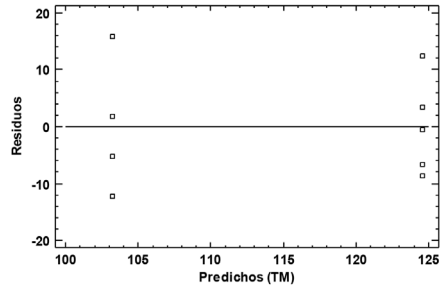


(b)

Figura 3 – Gráficos de Probabilidad Normal de Residuos: PFSA (a) y TM (b)



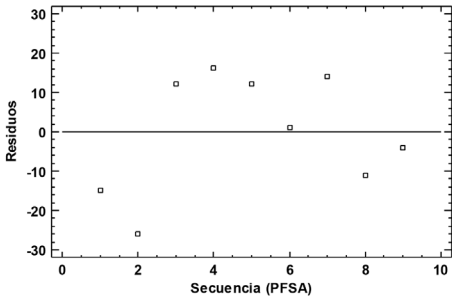
(a)



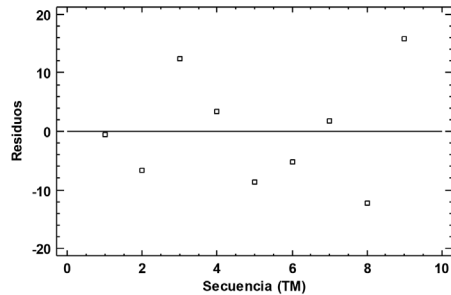
(b)

Figura 4 – Gráficos de Residuos vs. Predichos: PFSA (a) y TM

Finalmente, para validar el supuesto de independencia de los datos, generamos un gráfico de residuales vs. secuencia de los datos; en este caso, observamos si es posible detectar



(a)



(b)

Figura 5 – Gráficos de Residuos vs. Secuencia: PFSA (a) y TM

alguna tendencia a tener rachas con residuos positivos y negativos; en el caso de nuestro análisis, podemos ver en la figura 5 que en ambos casos no se identifica una tendencia, por lo que es posible suponer que los datos provienen de poblaciones independientes.

Con el propósito de profundizar en el análisis, utilizamos el ANOVA para evaluar pruebas de hipótesis relacionadas con cada uno de los cinco componentes de la métrica PFSA (Ficheros Lógicos Internos: ILF, Ficheros de interface externos: EIF, Entradas Externas: EI, Salidas Externas: EO y Consultas Externas EQ); la tabla 4 resume los valores obtenidos para dicho análisis inferencial.

Métrica	F	p-valor
ILF	0.09	0.7730
EIF	0.01	0.9200
EI	0.01	0.9069
EO	10.94	0.0130 *
EQ	0.08	0.7916

*p-valor < 0.05

Tabla 4 – ANOVA para las cinco métricas que componen los PFSA

Como se observa en la tabla 4, únicamente uno de los cinco componentes de la métrica de PFSA, presentó diferencias estadísticamente significativas; el componente vinculado con el elemento de transacciones “salidas externas”.

Resulta importante mencionar, que aunque no se incluye en el artículo, la validación de los tres supuestos del modelo fue realizada por los autores, y en los cinco casos antes citados, el modelo resultaba válido.

5. Réplica Experimental

5.1. Participantes/Sujetos Experimentales

La muestra por conveniencia disponible para la réplica experimental estuvo conformada por 18 estudiantes de la carrera de Ingeniero de Software de la Universidad Autónoma de Yucatán, los participantes cursaban por segunda ocasión —en un curso de verano— la asignatura “Fundamentos de Ingeniería de Software” ubicada en el primer semestre de la carrera. Con los 18 participantes —de primer curso— se integraron 6 equipos de trabajo que fungieron como sujetos experimentales. La tabla 5 ilustra la asignación de sujetos a los tratamientos.

Tratamientos	Equipos
EB	I, II, III
ET	IV, V, VI

Tabla 5 – Asignación de sujetos experimentales por tratamiento

La integración de los Equipos Compatibles, con base en los roles identificados mediante el inventario de autopercepción, se ilustra en la figura 6.

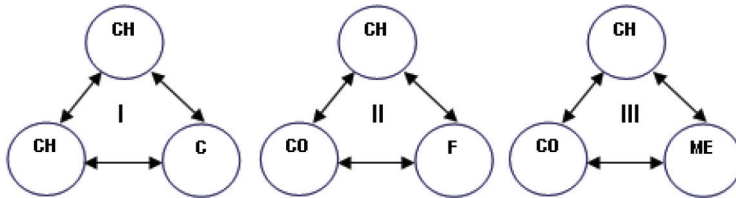


Figura 6 – Equipos integrados con la Teoría de Belbin (Equipos Compatibles)

5.2. Ejecución del Experimento

La ejecución de la réplica experimental se desarrolló en tres sesiones, en forma similar al primer experimento, es decir, para la sesión de capacitación, se utilizó el mismo material de instrucción, incluso fue el mismo instructor quien impartió dicho taller; en cuanto al problema por resolver, se utilizó el mismo Documento de Especificación de Requisitos, y en relación con el tiempo disponible para la actividad, también se dispuso de 180 minutos. Un parámetro que se consideró homogéneo respecto del primer experimento, fue el de la experiencia de los participantes, ya que en el caso de la réplica, los alumnos se encontraban en su primer año de la carrera.

5.3. Análisis y Resultados del Experimento

En esta sección se presenta tanto el análisis descriptivo, como el análisis inferencial con base en las variables seleccionadas para el experimento. Los gráficos de Caja y Bigotes generados para las variables dependientes consideradas se presentan en la figura 7. Como se observa en el gráfico de la figura 7(a), en cuanto a la variable PFSA, los ET presentan mayor dispersión y un sesgo negativo respecto de su mediana, por su parte los EC tienen poca dispersión pero también presentan un sesgo negativo respecto de su mediana.; si observamos ambas gráficas al mismo tiempo, vemos que el área de la caja de los EC se encuentra totalmente desfasada respecto de los ET, lo cual nos da indicios de diferencias con la variable PFSA.

Por su parte, la figura 7(b) ilustra el comportamiento de los datos vinculados con la variable TM, observamos que los EC presentan mayor dispersión que los ET, y ambos mantienen un sesgo positivo, también se puede observar que el área de la caja de los ET se encuentra incluida en el área de los EC, lo cual nos da indicios de que no haber diferencias en cuanto a la variable TM.

Con el objeto de evaluar estadísticamente las diferencias observadas en el análisis descriptivo previo, se plantearon las mismas hipótesis estadísticas que en el primer experimento.

$$H_{0_{PFSA}}: \mu_{EC} = \mu_{ET} \ ; \ H_{1_{PFSA}}: \mu_{EC} <> \mu_{ET} \tag{I}$$

$$H_0_{Time}: \mu_{EC} = \mu_{ET} ; H_{1_{Time}}: \mu_{EC} <> \mu_{ET} \quad (II)$$

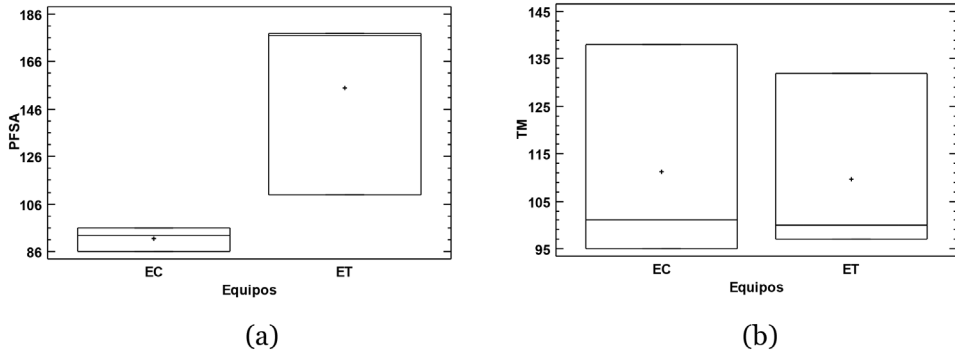


Figura 7 – Gráficos de Caja y Bigotes para las variables PFSA (a) y TM (b)

El resultado de evaluar con ANOVA ambas hipótesis, se ilustra en la tabla 6.

Métrica	F	p-valor
PFSA	7.79	0.0494 *
TM	1.01	0.9287

*p-valor < 0.05

Tabla 6 – ANOVA para las variable PFSA y TM

En virtud de que el valor P de la prueba F —en la variable PFSA- es menor que 0.05 (PFSA: 0.0494), entonces podemos rechazar la hipótesis de nulidad y afirmar que existe una diferencia estadísticamente significativa entre las medias de la variable PFSA de los tratamientos. Por su parte el valor p de la prueba F —en la variable TM— es mayor que

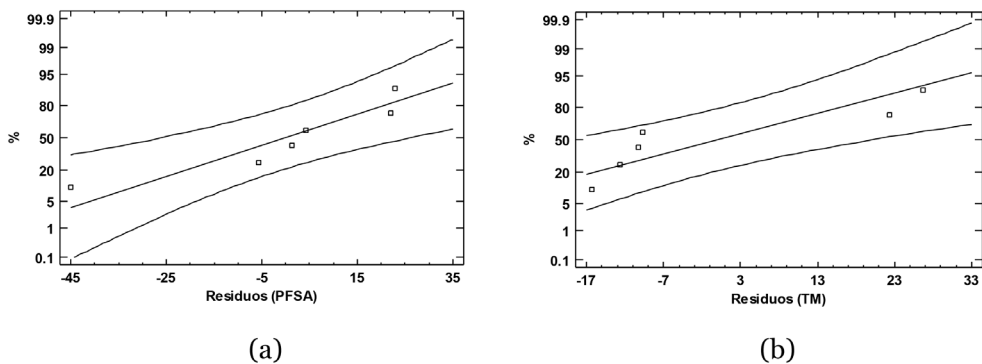


Figura 8 – Gráficos de Probabilidad Normal de Residuos: PFSA (a) y TM (b)

0.05 (TM: 0.9287) la hipótesis de nulidad se acepta, es decir, se afirma que no existe diferencias entre las medias de los dos tratamientos.

Para validar el primer supuesto del modelo ANOVA, generaremos el gráfico de probabilidad normal de los residuos para ambas variables. Como se puede ver en la gráfica de la figura 8, los puntos, en ambas gráficas, no muestran desviaciones graves de la diagonal, por lo que es posible asumir que los residuos tienen una distribución normal en ambos casos.

Para analizar la homogeneidad de varianzas, observamos el gráfico de residuales vs. predichos; y como podemos observar en ambas variables, no se tiene un comportamiento de aumento o disminución de valores predichos, por lo que consideramos homogeneidad de varianzas.

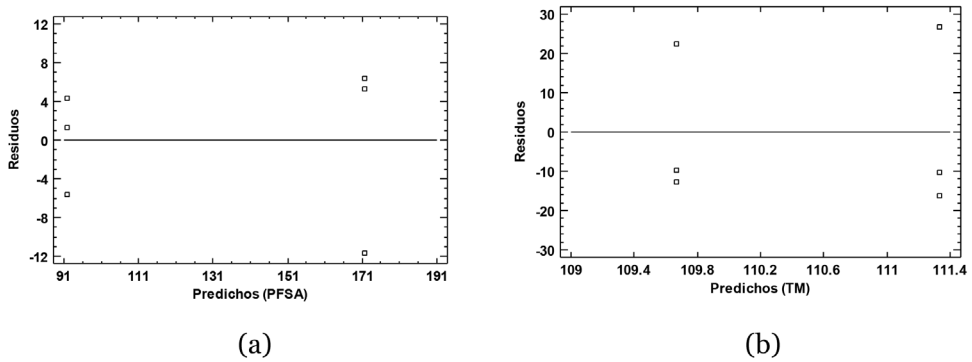


Figura 9 – Gráficos de Residuos vs. Predichos: PFSA (a) y TM

Finalmente, para validar el supuesto de independencia de los datos, generamos un gráfico de residuales vs. secuencia de los datos, y como podemos ver en la figura 10, en ambos casos no se identifica tendencia alguna en los datos, por lo que podemos suponer que los datos provienen de poblaciones independientes.

De manera similar a lo reportado en el primer experimento, se realizaron los cinco análisis de varianza —para cada componente de la métrica PFSA— con sus respectivas pruebas de validez al modelo, la cuales, por las limitaciones de espacio no serán incluidas en el artículo; la tabla 7 resume los valores obtenidos para el ANOVA.

Métrica	F	p-valor
<i>ILF</i>	13.31	0.0218 *
<i>EIF</i>	0.84	0.4104
<i>EI</i>	1.50	0.2879
<i>EO</i>	0.13	0.7406
<i>EQ</i>	0.23	0.6542

*p-valor < 0.05

Tabla 7 – ANOVA para las cinco métricas que componen los PFSA

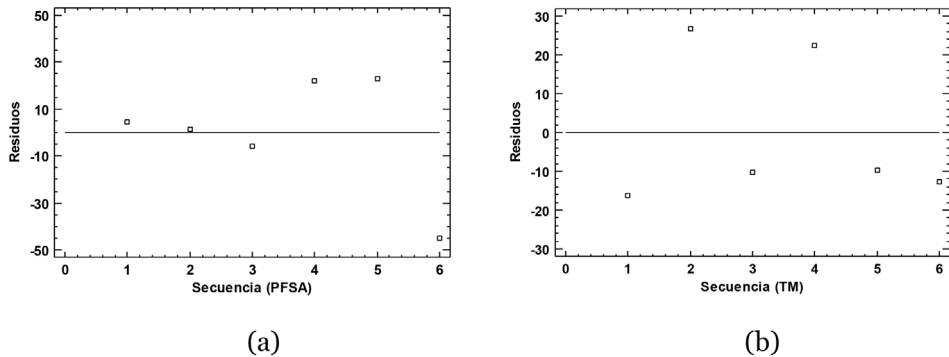


Figura 10 – Gráficos de Residuos vs. Secuencia: PFSA (a) y TM

Como se observa en la tabla 7, aunque de manera global la variable PFSA presenta diferencias significativas entre ambos tratamientos, en cuanto a sus componentes, solamente uno de los cinco presentó diferencias estadísticamente significativas, particularmente, el componente vinculado con el elemento de datos “archivos lógicos internos”.

6. Conclusiones y Trabajos Futuros

En este estudio presentamos un primer experimento controlado con el propósito de explorar si el uso de la teoría de roles de Belbin puede ser de utilidad para la integración de equipos que pudiesen evidenciar mayor calidad, tanto en el proceso de desarrollo, como en los productos generados. En particular, se seleccionó la actividad de medición de software en el contexto de un entorno académico. Con el propósito de conformar o refutar los hallazgos del primer experimento, se realizó una réplica similar-interna. En el primer experimento se encontró con que la variable PFSA no presentó diferencias entre tratamientos, sin embargo, la variable Tiempo en minutos para concluir la actividad, presentó diferencias significativas. Con el objetivo de profundizar en los componentes de la variable PFSA, se analizaron posibles diferencias entre sus cinco elementos, presentando diferencias solo uno de éstos. Posteriormente, con la réplica experimental, los resultados no fueron confirmados, en este caso, la variable PFSA presentó diferencias significativas. Por su parte, lo obtenido en la variable Tiempo para concluir la tarea, en la réplica no se identificaron diferencias entre tratamientos. Siendo un estudio exploratorio, y luego de un razonamiento abductivo, podemos comentar que posiblemente la madurez de los participantes en el estudio tuvo alguna influencia en los resultados, si bien, en ambos experimentos se mantuvo un mismo nivel de experiencia en la técnica, la madurez en procesos de software en el primer experimento, es mayor que la de los sujetos de la réplica.

Con lo obtenido en el estudio, los autores proponen continuar con el desarrollo de experimentos controlados con estudiantes; una variante es analizar el factor madurez académica, para identificar si éste tiene alguna influencia en las variables respuesta. En cuando a la incorporación de una nueva variable dependiente, se tiene el interés

de incorporar una métrica vinculada con la colaboración de los integrantes del equipo, durante la realización de una tarea. Por otro lado, considerando los estudios previos, y siendo la actividad de medición de software, una actividad que pudiese realizarse de manera individual, nos resulta interesante investigar si alguno de los roles de Belbin presenta diferencias significativas en cuanto a la calidad de la medición de software.

Referencias

- Aguileta, A. Ucán, J., & Aguilar, R. (2017). Explorando la influencia de los roles de Belbin en la calidad del código generado por estudiantes en un curso de ingeniería de software. *Revista Educación en Ingeniería*, 12(23), 93–100.
- Aritzeta, A., Swailes, S., & Senior, B. (2005). *Team roles: Psychometric evidence, construct validity and team building*. Hull, UK: University Hull.
- Basili, V., Selby, R., & Hutchens, D. (1986). Experimentation in Software Engineering. *IEEE Transactions on Software Engineering*, 12(7), 733–743.
- Belbin, M. (1981). *Management teams: Why they succeed or fail*. New York: John Wiley & Sons.
- Belbin, M. (1993). *Team roles at Work*. Oxford, USA: Elsevier Butterworth Heinemann.
- Bourque, P., & Fairley, R. (2014). *Guide to the Software Engineering Body of Knowledge (SWEBOK V3.0)*. IEEE Computer Society.
- Briggs-Myers, I., & McCaulley, M. (1985). *Manual, A Guide to the Development and Use of the Myers-Briggs Type Indicator*. Palo Alto, USA: Consulting Psychologists Press.
- Estrada, E., & Peña, A. (2013). Influencia de los roles de equipo en las actividades del desarrollador de software. *Revista Electrónica de Computación Informática, Biomédica y Electrónica*, 2(1), 1–19.
- Garmus, A. & Herron, D. (1996). *Measuring the Software Process: a practical guide to functional measurements*. New Jersey, USA: Prentice Hall.
- Genero, M., Cruz-Lemus, J., & Piattini, M. (2014). *Métodos de investigación en ingeniería de software*. Madrid, España: Ra-Ma.
- Glass, R., Vessey, I., & Ramesh, V. (2002). Research in Software Engineering: An analysis of the literature. *Information and Software Technology*, 44(8), 491–506.
- Gómez, O., Juristo, N., & Vegas, S. (2014). Understanding replication of experiments in software engineering: A Classification. *Information and Software Technology*, 56(8), 1033-1048. doi: 10.1016/j.infsof.2014.04.004.
- Gutiérrez, H., & De la Vara, R. (2012). *Análisis y Diseño de Experimentos*, (3ª Ed). Ciudad de México: McGraw Hill.
- Henry, S., & Stevens, K. (1999). Using Belbin's leadership role to improve team effectiveness: An empirical investigation. *The Journal of Systems and Software*, 44, 241–250.

- Hernández, L., Muñoz, M., Mejía, J., Peña, A., & Rangel, N. (2017). Una Revisión Sistemática de la Literatura Enfocada en el uso de Gamificación en Equipos de Trabajo en Ingeniería de Software. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, (21), 33-50. doi: 10.17013/risti.21.33-50.
- Johansen, T. (2003). *Predicting a Team's Behaviour by Using Belbin's Team Role Self Perception Inventory*. (Dissertation at Department of Management & Organisation), University of Stirling, UK.
- Juristo, N., & Moreno, A. (2001). *Basics of Software Engineering Experimentation*. Boston, USA: Kluwer Academic Publishers.
- Margerison, C.J., & McCann, D.J. (1985). Team Management Profiles: Their use in Managerial Development. *Journal of Management Development*, 4(2), 34-37. doi: 10.1108/eb051580.
- Miramontes, J., Muñoz, M., Calvo-Manzano, J., & Corona, B. (2016). Establecimiento del estado del arte sobre el aligeramiento de procesos de software. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, (17), 16-25. doi: 10.17013/risti.17.16-25.
- Morales, N., & Vega, V. (2018). Factores Humanos y la Mejora de Procesos de Software: Propuesta inicial de un catálogo que guíe su gestión. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, (29), 30-42. doi: 10.17013/risti.29.
- Mumma, F.S. (2005). *Team-work & team-roles: what makes your team tick? facilitator guide*. King of Prusia, PA: HRDQ Press.
- Norambuena, B., & Vega, V. (2017). Minería de procesos de software: una revisión de experiencias de aplicación. *RISTI - Revista Ibérica de Sistemas y Tecnologías de Información*, (21), 51-66. doi: 10.17013/risti.21.51-66.
- Senior, B. (1997). Team roles & Team performance: Is there 'really' a link?. *Journal of Occupational and Organizational Psychology*, (70), 85-94.