

CAPÍTULO 9

ANÁLISIS COMPARATIVO DE RESULTADOS

9 ANALISIS COMPARATIVO DE RESULTADOS

9.1 Introducción

Una vez expuestos los resultados obtenidos por cada programa en el capítulo anterior, se presenta a continuación el análisis de los resultados en forma comparativa entre los cuatro tipos software y el grupo testigo (sin software). Al igual que en el caso anterior, se presenta subdividido en análisis de cantidades - tiempos y en análisis de calidades - tendencias. Este tipo de análisis permite determinar si existen diferencias estadísticamente significativas entre los diferentes programas, que conduzcan a conclusiones objetivas.

El análisis estadístico que se emplea es el de varianza²² para un solo factor: «tipo de software utilizado», con el cual es posible determinar si las diferencias de resultados es una diferencia real o es una diferencia debida a las variaciones de cada repetición. Establecer las razones por las que varía la media es importante para este estudio, debido a que la heterogeneidad de los resultados de cada caso individual puede conducir a conclusiones equivocadas. En algunos de los casos el análisis de varianza se complementa con el método de Tuckey de comparaciones múltiples (Ferrán, 2001; Montgomery, 2002), para determinar entre cuáles tratamientos existe diferencia de medias. Con base en este análisis se hacen las conclusiones importantes en este capítulo, presentando los datos más relevantes. En el Anexo 2 se presenta en detalle las tablas con el análisis de varianza realizado con el paquete estadístico SPSS, a las cuales se hará referencia durante este capítulo.

En análisis comparativo se distribuye en cinco secciones a lo largo del capítulo. En la primera se hace referencia a los tiempos empleados en cada etapa y cada actividad del proceso (en %), con lo cual se puede estimar las preferencias de cada participante y la dedicación en cada caso. En la segunda sección del capítulo se presenta el análisis de la cantidad total de ideas, del flujo y del origen de tales ideas, completando de esta manera el análisis puramente cuantitativo de los diferentes tratamientos estudiados.

La tercera sección está dedicada al análisis de lo que se ha denominado la «calidad creativa» de acuerdo con el modelo presentado en el capítulo seis. Para ello se compara cada uno de los índices (flexibilidad, fluidez, elaboración y originalidad) exhibidos a lo largo del proceso, de tal manera que es posible contrastar el comportamiento exhibido durante la ejecución de la sesión de diseño para cada caso. En la cuarta sección se analiza en forma comparativa la trayectoria creativa total, es decir, aquella que combina

²² Al análisis de varianza comúnmente se le llama ANOVA, del término en inglés ANalysis Of Variance.

los cuatro índices mencionados, de acuerdo con el modelo de Shah (2003). Además se presenta como valoración total de cada tratamiento, el índice de «creatividad global».

Finalmente, en la quinta sección se estudia la producción de ideas a lo largo del proceso, mostrando las etapas en las que se produce mayor número de ideas y cómo el proceso en la medida que avanza va convergiendo, reduciendo el número de ideas generadas.

Las conclusiones presentadas al finalizar el capítulo se basan en el análisis realizado en cada una de las cinco secciones y se busca con ellas, ser completamente objetivo con el ánimo de establecer pautas claves para proponer un nuevo software que recoja las principales virtudes de los evaluados y que eviten los inconvenientes detectados.

Con el ánimo de evitar repeticiones y facilitar la lectura, los tratamientos se denominarán así:

- **A** para Axon idea processor©
- **B** para Brainstorming toolbox©
- **C** para CREAX innovation suite©
- **T** para ThoughtPath Corporate Edition©
- **S** para el tratamiento sin utilizar software.

9.2 Análisis de tiempos

Para efectos de comparación entre los diferentes programas este análisis se basa en el porcentaje de tiempo empleado en las diferentes actividades de la sesión, teniendo en cuenta que el tiempo utilizado por los participantes no fue siempre igual y por lo tanto, no es posible hacer un análisis con valores absolutos. La Tabla 9.1 muestra tales valores comparativos.

Las diferencias más significativas entre los tratamientos que utilizan software con respecto al que no utiliza, están centradas en las actividades 3, 4 y 5 (hablar, dibujar y pensar, respectivamente), en las que el tratamiento **S** emplea más tiempo.

En el caso de la actividad 5 (pensar), se puede entender la diferencia tan marcada del caso **S** por el hecho de no estar desarrollando paralelamente la actividad de consulta en el software, como en los otros casos. Así, la persona interactúa solamente con su propia mente y no con agentes externos. Las ideas se generan exclusivamente por la experiencia, el conocimiento previo y su recuperación en la memoria del individuo. Por ello, fue notorio el esfuerzo mental para efectuar la conceptualización de soluciones

potenciales, el cual se evidencia en el mayor tiempo empleado en esta actividad. Es claro, sin embargo, que en todos los demás casos también se desarrolló esta actividad, aunque en forma paralela y con la asistencia de los diferentes estímulos provisto por software.

Tabla 9.1 Porcentaje de tiempo empleado en diversas actividades para todos los programas

Actividad	Software				
	A	B	C	T	S
1 Lectura de problema asignado / guía	3,3	2,2	5,9	3,9	3,9
2 Manipulación de envases de tetrabrick	8,0	4,6	5,3	4,5	4,7
3 Hablar: explica ideas o comentarios	18,0	9,5	12,0	9,8	31,4
4 Dibujar / Escribir sobre el papel	14,7	21,4	18,2	15,5	30,0
5 Pensar	5,2	2,1	14,7	5,3	23,6
6 Navegar en el programa	11,4	29,0	21,4	16,6	0,0
7 Escribir / Editar en el programa	31,9	17,9	18,5	32,9	0,0
8 Leer descripciones anteriores	4,3	3,2	2,3	6,5	3,6
9 Otras acciones dentro del software	0,0	0,0	0,0	3,2	0,0
10 Otras acciones no clasificadas	3,2	10,0	1,6	1,8	2,8

Para la actividad 3 (hablar) se entiende la diferencia por el hecho de que por no interactuar con agentes externos (el ordenador, en este caso), la persona tiene mayor libertad para expresar sus ideas audiblemente en la medida que van surgiendo. La utilización del ordenador implica fijar la atención en su interacción antes que en estar expresando las ideas en voz alta y, aunque en forma regular se les estuvo recordando la necesidad de hacerlo, en todos los casos fue evidente la dificultad para hacer las dos actividades paralelas en forma regular.

La diferencia en la actividad 4 (dibujar) es obvia por no haberse restringido el uso de papel y lápiz durante la primera parte de la sesión, en aquellos casos que se trabajó sin software. Incluso podría esperarse que esta diferencia en el tiempo fuese mayor. Así, la forma de plasmar las ideas fue exclusivamente a través del papel, mientras que en los demás casos existió la posibilidad de escribir en las ventanas que para el efecto cada tipo de software dispone.

El análisis de varianza con comparaciones múltiples con el método Tukey, revela que no existe diferencia significativa entre las medias de los datos de la actividad 4 (dibujar). Sin embargo, se encontró diferencia en la actividad 3 (hablar), con un nivel de significación $\alpha = 0,05$, entre el tratamiento **S** con respecto a cada uno de los demás tratamientos. Algo similar sucede con la actividad 5 (pensar), cuya media difiere significativamente entre el

tratamiento **S** y los tratamientos **A**, **B** y **T**. Así, se confirma estadísticamente la conclusión antes comentada²³.

Se puede afirmar, por lo tanto, que el uso de software implica la necesidad de prestar atención a un agente externo que, por un lado, restringe el tiempo y la atención para realizar otras actividades como pensar y hablar, pero por otro, establece una intermediación en el proceso de recuperación de conocimiento y experiencias, así como de su recombinação para generar nuevas soluciones. Por lo tanto, se puede afirmar que aunque se empleó menos tiempo en pensar, hablar y dibujar, el uso de software puede constituirse en un medio de estímulo en la generación de ideas, al facilitar la recuperación de experiencias previas y la recombinação para generar nuevas propuestas.

9.3 Análisis de cantidades

9.3.1 Cantidad de ideas

El número total de ideas generadas en cada sesión y para cada software se presenta en la Tabla 9.2. El valor promedio de ideas indica más fluidez para el Brainstorming toolbox© con 49,75 ideas, seguido por el tratamiento sin software, con 42,50.

Tabla 9.2 Número total de ideas generadas en cada repetición para cada programa

Software	Repetición				Total	Media	D.E.
	1	2	3	4			
Axon	51	32	37	39	159	39,75	8,06
Brainstorming	61	38	60	40	199	49,75	12,45
CREAX	25	43	19	32	119	29,75	10,31
ThoughtPath	37	23	41	34	135	33,75	7,72
Sin software	38	34	47	51	170	42,50	7,85

La Figura 9.1 muestra el diagrama de caja correspondiente a esta variable, donde se puede apreciar más claramente y en forma comparativa, las tendencias de los datos, los valores medios, las dispersiones y los valores extremos. Por ejemplo, el tratamiento **B** tiene una mediana de 50 ideas, un cuartil superior de 60 y uno inferior de 40, y los valores extremos son 61 y 38.

²³ Los detalles del análisis estadístico de las variables hablar, pensar y dibujar se encuentran en el Anexo 2, numeral 1.1 a 1.3.

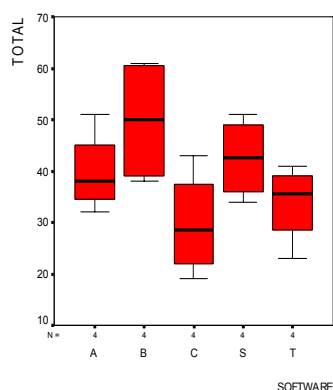


Figura 9.1 Diagrama de caja del número total de ideas para cada tratamiento

Aunque se muestra unas tendencias diferenciales entre tratamientos, es necesario realizar un análisis de varianza²⁴, que conduzca a conclusiones objetivas. El análisis de varianza, que se muestra en la Tabla 9.3, evidencia que se debe aceptar la hipótesis estadística nula, es decir, no existe diferencias significativas entre el total de ideas para cada tratamiento.

El análisis anterior se limita al número total de ideas. Pero el análisis detallado por tipo de idea (funciones modificadores y estructuras)²⁵, revela que solamente se obtienen diferencias significativas (con $\alpha=0,05$) para los modificadores funcionales y, específicamente, entre los tratamientos CREAX y Brainstorming. Esto significa que el Brainstorming permite la obtención de un número mayor de modificadores funcionales que el que se logra utilizando el CREAX. Para todos los demás casos, la cantidad de ideas producidas no son estadísticamente diferentes.

²⁴ Los resultados de este tipo de análisis se suelen presentar como se muestra en la siguiente tabla:

	Suma de cuadrados	Grados de libertad	Media cuadrática	F	Sig.
Inter-grupos	SC_{inter}	$K - 1$	$MC_{inter}^2 = \frac{SC_{inter}}{K - 1}$	$F = \frac{MC_{inter}}{MC_{intra}}$	Valor P del estadístico F
Intra-grupos	SC_{intra}	$N - K$	$MC_{intra}^2 = \frac{SC_{intra}}{N - K}$		
Total	SC_{total}	$N - 1$			

En ella, la variabilidad total de los datos se descompone en la variabilidad debida a las diferencias entre grupos (inter-grupos) y la debida a las diferencias dentro de los grupos (intra-grupos). El estadístico de prueba F permite contrastar la hipótesis nula de igualdad de medias, es decir, compara la variabilidad debida a las diferencias entre grupos con la debida a las diferencias dentro de los grupos. Así, entre mayor sea F, más diferenciados estarán los grupos. Si el valor p asociado es menor que α (error Tipo I) se rechaza la hipótesis nula al nivel de significación α (Ferrán 2001).

²⁵ Este análisis detallado se encuentra en el Anexo 2, numeral 2.

Tabla 9.3 Análisis de varianza para el número total de ideas

	Suma de cuadrados	Grados de libertad	Media cuadrática	F	Sig.
Inter-grupos	965,8	4	241,450	2,699	,071
Intra-grupos	1342,0	15	89,467		
Total	2307,8	19			

9.3.2 Flujo de ideas

Tal como se comentó en el capítulo 8, el tiempo utilizado por los participantes en las diversas sesiones experimentales no siempre fue el mismo. Esto ha obligado a realizar un análisis teniendo en cuenta la cantidad de ideas producidas por unidad de tiempo, a lo que se ha denominado «flujo de ideas».

Tabla 9.4 Flujo de ideas (ideas/min)

Software	Repeticiones				Media	D.E.
	1	2	3	4		
Axon	0,46	0,29	0,51	0,34	0,40	0,102
Brainstorming	0,55	0,50	0,56	0,38	0,50	0,083
CREAX	0,41	0,39	0,34	0,32	0,37	0,042
TrhoughtPath	0,34	0,37	0,38	0,31	0,35	0,032
Sin software	0,59	0,31	0,50	0,56	0,49	0,126

La Tabla 9.4 presenta los resultados para todos los casos. El valor de la media no parece ser muy diferente entre los tratamientos; de hecho, los casos **B** y **S** son prácticamente iguales. Esta afirmación se ratifica al examinar el diagrama de caja mostrado en la Figura 9.2, en la que se establecen dos grupos con medias muy semejantes.

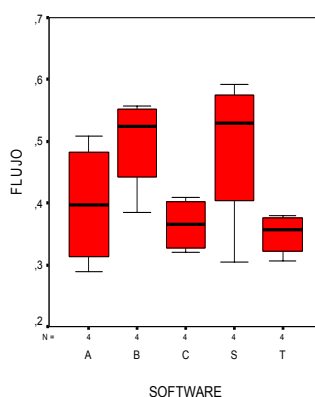


Figura 9.2 Diagrama de caja para la variable flujo de ideas

El grupo de mayor valor corresponde a los dos tratamientos mencionados, mientras que el menor a los otros tres tratamientos. Aunque parece existir un valor atípico (caso S-2), el análisis de residuales estandarizados (Montgomery, 2002, pp.78-79) rechaza esta presunción²⁶.

El análisis ANOVA mostrado en la Tabla 9.5, confirma que efectivamente no existe una diferencia significativa entre las medias del flujo de ideas. Por lo tanto, se puede afirmar que la utilización de software no aumenta, pero tampoco disminuye la producción de ideas por unidad de tiempo. Este resultado es importante teniendo en cuenta que podría haberse esperado que el flujo de ideas disminuyese con el uso de software, debido a la relativamente poca experiencia que cada participante tenía en su utilización, y debido a las observaciones anotadas por Pugh (1990) y otros críticos del uso de software como herramienta de diseño, en el sentido de que el uso de herramientas informáticas entorpecen en lugar de ayudar en el desarrollo del proceso de diseño.

Tabla 9.5 Análisis de varianza para el flujo de ideas

	Suma de cuadrados	Grados de libertad	Media cuadrática	F	Sig.
Inter-grupos	0,077	4	0,019	2,671	0,073
Intra-grupos	0,108	15	0,007		
Total	0,185	19			

9.3.3 Origen de ideas

El análisis comparativo entre tratamientos solo se puede realizar para el grupo de orígenes de tipo general, esto es, aquellos externos al software. Los internos, por ser diferentes para cada tipo de programa no se pueden analizar en conjunto. Sin embargo, a manera de síntesis de los resultados individuales presentados en el capítulo anterior, se presentarán aquí, aunque no sea útil para efectuar comparaciones.

a. Origen externo al software

Se hace aquí referencia al análisis para el primer grupo comentado, es decir, aquellos que se enfatizan en elementos externos al software respectivo. Estos son: enunciado del problema, experiencias y conocimiento acumulados, manipulación de envases de tetrabrik®, dibujos previamente elaborados y módulos del programa (en términos genéricos, sin especificar cuáles). La Figura 9.3 muestra los resultados para todos los casos.

²⁶ Consultar tabla A2.10 del Anexo 2.

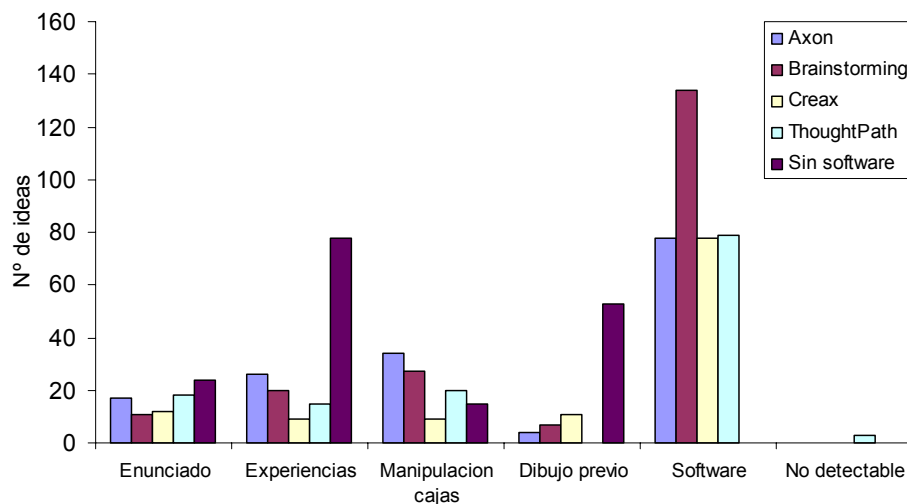


Figura 9.3 Número de ideas por fuente

Evidentemente, cuando no se utiliza software el origen de las ideas se centra en las experiencias (y conocimientos) disponibles y en los dibujos previamente elaborados. Esto significa que el diseño queda supeditado a las habilidades específicas del ingeniero o diseñador, logradas por su experiencia previa, sus conocimientos específicos y su percepción de soluciones ya existentes.

Para los demás tratamientos la fuente principal de ideas fue el propio programa. Aquí hay que acotar esta afirmación, diciendo que si bien es cierto el elemento catalizador de las ideas propuestas fue el software, no se puede negar la influencia que tienen las propias habilidades del diseñador mencionadas anteriormente. Aquí solamente se puede afirmar que el software ha servido de «puente» entre aquellas habilidades y el problema a resolver. Es interesante notar que, con excepción del Brainstorming toolbox©, el número de ideas generadas a partir del software fue prácticamente igual en todos los casos. Esto indica una mayor eficacia de aquel programa básicamente por los módulos estimuladores de ideas.

La segunda fuente de ideas para los casos que utilizaron software fue la manipulación de envases de tetrabrik®. Esto parece demostrar que el usuario requiere de algún elemento físico que permita filtrar las diferentes propuestas que se logran con el software. El tocar con las manos sirve para darle forma a aquellas propuestas del software que en principio no pueden asociarse mentalmente. Esta afirmación, sin embargo, debería ser verificada por metodologías de la psicología cognitiva, que se escapan del alcance de esta investigación.

b. Módulos de origen

No es posible hacer una comparación cuantitativa de las ideas generadas dentro de los módulos de cada software debido a las diferentes técnicas que implementan. Sin embargo, sí se puede identificar cuáles módulos han sido los más eficaces y cuáles programas los más eficientes²⁷.

La Figura 9.4 muestra el porcentaje de ideas producido en cada programa. En los tres primeros tratamientos se evidencia el predominio de un módulo en cuanto a la eficacia para asistir en la generación de ideas: *mind map*, para el caso de Axon idea processor©; *random picture*, para Brainstorming toolbox© y *resources*, para Creax innovation suite™.

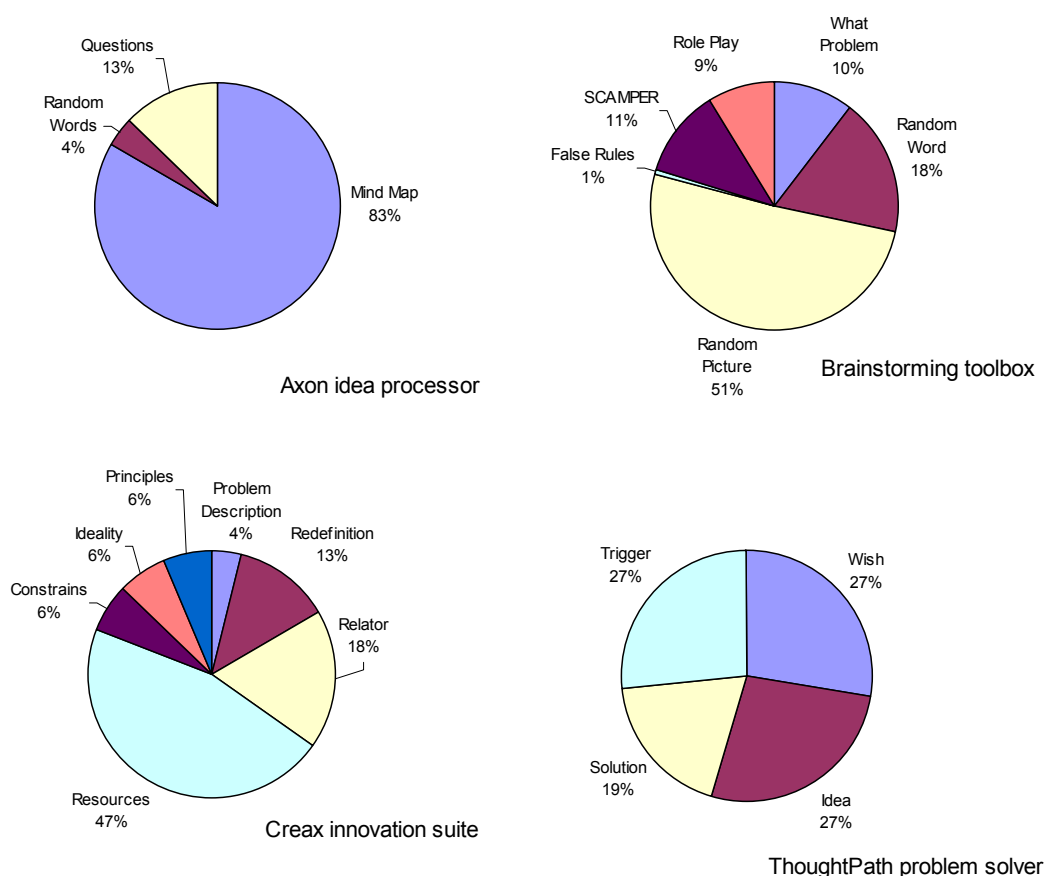


Figura 9.4 Porcentaje de ideas producidas por módulo en cada programa

La característica común de estos tres módulos es la posibilidad de interacción de tipo gráfica, entre el usuario y el programa. En el primer caso, la construcción de mapas

²⁷ Entendiendo como eficacia el rendimiento del módulo para producir gran número de ideas y como eficiencia la característica del programa para producir resultados equilibrados y acordes con los recursos utilizados.

mentales facilita la visualización global del problema con todos sus matices, dejando a un lado la representación lineal que impone el texto escrito. En el segundo, el llamado a forzar relaciones entre figuras o fotografías aleatorias y el problema planteado, ha resultado más efectivo que las otras técnicas disponibles. Finalmente, en el tercer programa (Creax Innovation Suite), la matriz de recursos del sistema, ha resultado ser importante a la hora de identificar ideas para avanzar en la solución del problema planteado. Resulta evidente por estos resultados, la importancia que tiene lo visual, lo gráfico y lo pictórico para potenciar la generación de ideas. Si además, el software posee una interfase que permita una interacción flexible con el usuario, como es el caso del Axon, los resultados seguramente serán mejores.

En el cuarto tratamiento (Thoughtpath) se aprecia un mayor equilibrio entre la cantidad de ideas generadas por los diversos módulos del programa. De hecho, los cuatro módulos centrales del programa generaron aproximadamente el mismo número de ideas. Este programa carece de estímulos de tipo gráfico y, si bien es cierto, algunos de sus «triggers» invita a imaginar situaciones o a consultar elementos visuales externos, su estructura es básicamente de tipo textual y, por lo tanto, lineal. Esto lleva a confirmar las conclusiones de los casos anteriores.

Ahora bien, el equilibrio en la generación de ideas, esto es, la inexistencia de un módulo predominante, también se debe a la estructura secuencial del programa, que obliga a seguir una trayectoria predeterminada de avance a través de todas las herramientas. Esta característica puede ser deseable en ciertas circunstancias, dado que un formato completamente abierto, puede evocar una falta de estructura, a la que normalmente está acostumbrado el ingeniero. Por ello, se puede afirmar que un software adecuado debe mantener un equilibrio claro entre estructura y flexibilidad, entre un camino sugerido y una libertad de elección. Este permitirá al ingeniero tener siempre el control del proceso de desarrollo, pero sin sentir la presión de tener que cumplirlo exactamente de acuerdo con unas pautas preestablecidas.

Al analizar los valores totales, se encuentra que se produjeron 369 ideas a partir de los diferentes módulos de los programas estudiados. Esto equivale al 60% del total de ideas generadas. En cuanto al número y porcentaje de ideas por cada módulo, sobre el total producidas durante las 20 sesiones, los resultados se muestran ordenados en la Tabla 9.6.

Evidentemente, los dos primeros módulos o herramientas marcan una diferencia importante, sumando el equivalente al 36 % de todas las ideas. Estos valores deben ser tenidos en cuenta al formular la propuesta de software en el próximo capítulo.

Tabla 9.6 Número y porcentaje de ideas generadas por módulo de todos los programas

Software	Módulo	Nº de ideas	%
B	Random Picture	68	18,43
A	Mind Map	65	17,62
C	Resources	36	9,76
B	Random Word	24	6,50
T	Wish	22	5,96
T	Idea	21	5,69
T	Trigger	21	5,69
B	SCAMPER	15	4,07
T	Solution	15	4,07
B	What Problem	14	3,79
C	Relator	14	3,79
B	Role Play	12	3,25
A	Questions	10	2,71
C	Redefinition	10	2,71
C	Constrains	5	1,36
C	Ideality	5	1,36
C	Principles	5	1,36
A	Random Words	3	0,81
C	Problem Description	3	0,81
B	False Rules	1	0,27
	Total ideas	369	

9.4 Análisis de calidad creativa

Tal como se explicó en el capítulo siete, el modelo adoptado mide la calidad creativa con base en cuatro índices determinados para cada período del proceso seguido por el usuario. Hacer una comparación de la variación de los índices a lo largo del proceso servirá para determinar las tendencias marcadas para cada tratamiento. También será útil efectuar la comparación del valor total obtenido al final de la sesión calculado simplemente como la sumatoria de los índices parciales.

9.4.1 Índice de flexibilidad

a. Comparación de la flexibilidad del proceso

La Figura 9.5 muestra la curva del valor medio de la flexibilidad durante el proceso para cada tratamiento experimental. La escala del índice de flexibilidad se ha mantenido constante para cada tratamiento con el propósito de facilitar su lectura comparativa.

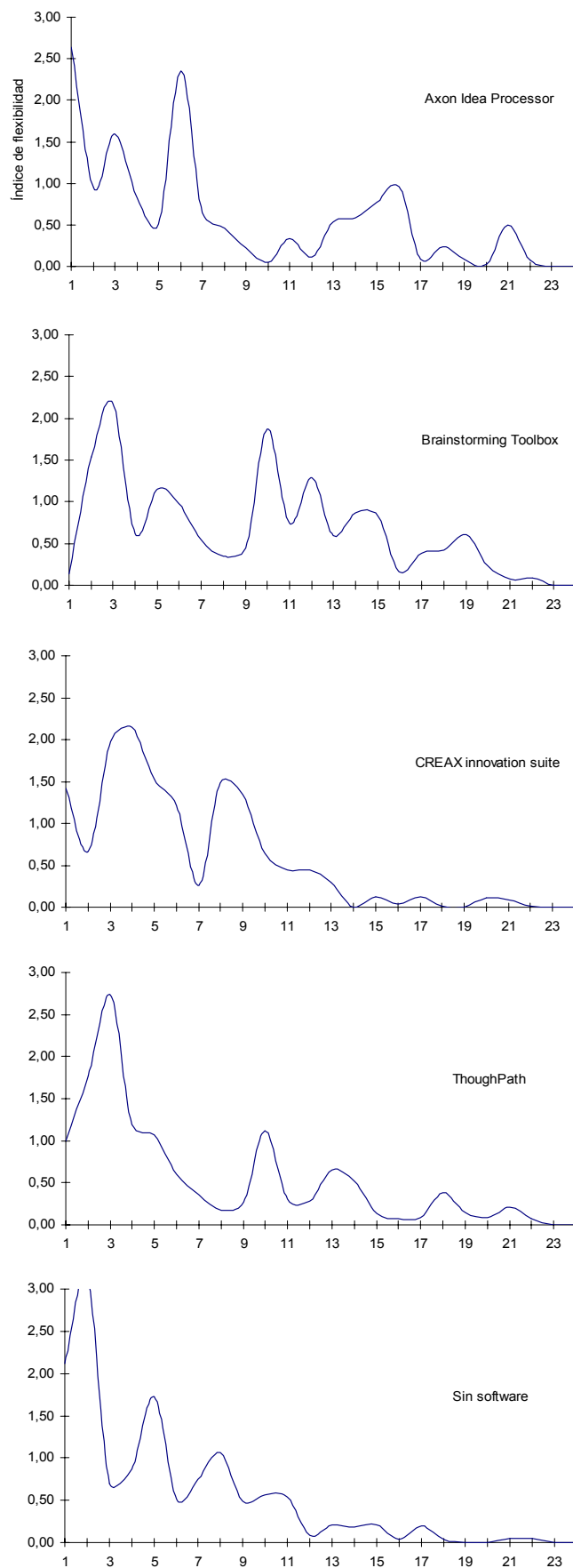


Figura 9.5 Comparación de la flexibilidad del proceso entre tratamientos

Todos los casos muestran un valor alto en los primeros períodos de la sesión, aunque algunos inician con un valor bajo para crecer rápidamente (**B** y **T**), explicado por el tiempo empleado en la lectura inicial del problema asignado. Estos valores altos se sostienen por poco tiempo (tres períodos) para luego comenzar a descender. El tratamiento **B** muestra el comportamiento más flexible a lo largo del proceso, caracterizado no propiamente por su estabilidad, pero sí por tener varios períodos con índices altos, comparables a los iniciales. El **A** también muestra valores altos aunque en menos períodos que el anterior.

El tratamiento **S** (sin software) evidencia un valor inicial muy alto en el período dos para luego caer gradualmente. También muestra picos en algunos períodos, pero siempre menores que los precedentes, de tal manera que la tendencia es claramente descendente.

Se puede concluir que el uso de software (en todos los casos) hizo que la flexibilidad a lo largo del proceso fuese mayor que el obtenido sin utilizar software. Los tratamientos con mejor comportamiento fueron en orden decreciente: B, A, C y T.

Es claro que se requiere de un análisis objetivo basado en los datos específicos para poder ratificar la conclusión anterior, sostenida en la revisión de las gráficas de comportamiento. Por ello se incluye la siguiente sección de análisis de flexibilidad total.

b. Flexibilidad total

La flexibilidad total para cada participante se obtuvo por la suma simple de los índices de cada período, y con base en este valor se realizó el análisis ANOVA para determinar si existen o no diferencias significativas en el valor medio de cada tratamiento.

Los datos para cada tratamiento y cada réplica se muestran en la Tabla 9.7. Se nota la existencia de un valor atípico²⁸ para el caso **T-2**, que puede dar origen a conclusiones equivocadas. La revisión de los datos que llevan al cálculo de este índice, revela que la causa de su particularidad no tiene origen estadístico, sino que se debe al hecho de que este participante utilizó solamente la mitad del tiempo disponible para la sesión. Esto indujo a un desempeño pobre caracterizado por su baja flexibilidad. Conviene en este caso seguir la recomendación de Montgomery (2002, p.78) en el sentido de realizar un análisis incluyendo el dato atípico y otro sin incluirlo.

²⁸ Se comprobó mediante el análisis de residuales estandarizados que este valor es atípico. Ver Anexo 2, numeral 3.1.1

Tabla 9.7 Índice de flexibilidad para cada repetición y cada tratamiento

Programa	Participante			
	1	2	3	4
Axon	15,30	14,50	14,10	14,40
Brainstorming	16,98	16,81	14,63	16,03
CREAX	15,03	14,28	13,25	14,85
Thoughtpath	14,40	7,70	14,30	16,50
Sin-Software	13,56	13,40	13,63	13,40

El primer análisis de varianza²⁹ muestra que la hipótesis nula (valor medio de la flexibilidad es igual para todos los tratamientos) es cierta, dando una probabilidad de error del 0,24 que supera al α establecido para esta investigación ($\alpha = 0,05$).

Al repetir el análisis de varianza eliminando aquel valor atípico, los resultados son diferentes, como lo muestra la Tabla 9.8. Ahora los valores medios de la flexibilidad sí son diferentes entre tratamientos y, al realizar el análisis por comparaciones múltiples³⁰ se identifica que la diferencia de medias ocurre particularmente entre los tratamientos **B** y **S**, confirmado de esta manera las conclusiones del análisis del proceso, antes comentado, de que aquel tratamiento es el más sobresaliente.

Tabla 9.8 Análisis de varianza para el índice de flexibilidad

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de cuadrados	F	Sig.
Entre grupos	14,717	4	3,679	5,547	0,007
Dentro de los grupos	9,285	14	0,663		
Total	24,002	18			

9.4.2 Índice de fluidez

a. Comparación de la fluidez de los procesos

La Figura 9.6 muestra las curvas comparativas de la fluidez para todos los casos. La producción de ideas a lo largo del proceso mostrada por el índice de fluidez vuelve a presentar al tratamiento **B** como el más regular de todos y al **S** como el menos fluido, aunque tenga el valor puntual más elevado al iniciar la sesión.

²⁹ Este análisis se muestra en el Anexo 2, tabla A2.12

³⁰ El análisis por comparaciones múltiples se encuentra en el Anexo 2, tabla A2.14

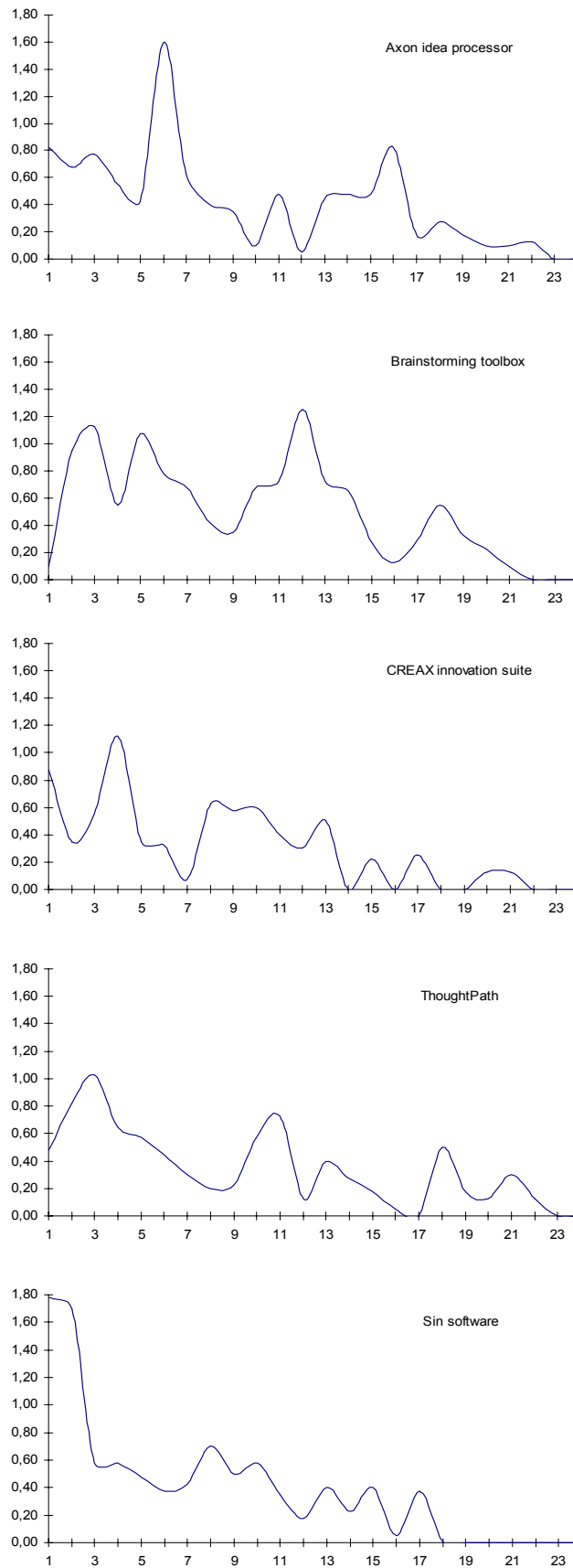


Figura 9.6 Comparación de la fluidez del proceso entre tratamientos

Se puede observar con referencia a este último que el índice llega a un valor nulo antes de terminar la sesión, es decir, la flexibilidad se termina antes que cualquier otro de los tratamientos.

El proceso del tratamiento **T** muestra menos irregularidades, aunque sus índices de fluidez son relativamente bajos, mientras que dentro de los tratamientos con software el **A** es el más irregular con altibajos grandes.

Salvo el tratamiento con CREAX innovation suite, la fluidez mostrada por los tratamientos con software es superior a la del tratamiento sin software. Ello indica que aunque el valor total de ideas fue mayor para éste último caso, el software potencia una mayor estabilidad en la fluidez de ideas mirada a lo largo de todo el proceso de desarrollo de productos.

b. Fluidez total

La fluidez calculada para cada tratamiento y cada participante se muestra en la Tabla 9.9. Los valores más elevados corresponden a los tratamientos **A** y **B**, mientras que los más bajos se asocian al **C**. Sin embargo la variabilidad entre participantes es alta, de manera que conviene, una vez más, hacer un análisis de varianza que permita llegar a conclusiones coherentes estadísticamente.

Tabla 9.9 Índice de fluidez por participante para cada tratamiento

Programa	Participante			
	1	2	3	4
Axon	12,60	9,50	8,70	9,30
Brainstorming	14,00	10,90	13,60	9,30
CREAX	6,20	9,60	5,40	8,30
Thoughtpath	8,70	6,70	10,20	7,50
Sin-Software	8,20	7,70	11,20	11,50

Los resultados del ANOVA para este caso se enseñan en la Tabla 9.10 e indican que sí existe una diferencia significativa entre el valor de las medias del índice de fluidez, con un valor de la variable estadística F superior al crítico.

El análisis por comparaciones múltiples entre todos los tratamientos³¹ revela que la diferencia significativa de las medias se da específicamente entre los tratamientos Brainstorming toolbox y CREAX, con lo cual se verifica las observaciones realizadas en el numeral anterior (análisis del proceso).

Tabla 9.10 Análisis de varianza para el índice de fluidez

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio de cuadrados	F	Sig.
Entre grupos	49,227	4	12,307	3,417	0,036
Dentro de los grupos	54,023	15	3,602		
Total	103,250	19			

9.4.3 Índice de elaboración

a. Comparación de la elaboración del proceso

El índice de elaboración de las ideas producidas a lo largo del proceso se muestra en la Figura 9.7, en forma comparativa entre todos los tratamientos. En esta oportunidad el comportamiento es similar entre el Axon y el Brainstorming, en cuanto a los picos que muestran, aunque el segundo tiene valores más altos. El CREAX muestra mucha inestabilidad y valores comparativamente bajos, mientras que el ThouhPath es, nuevamente, el más estable, pero también el que muestra valores bajos.

El tratamiento sin software se caracteriza nuevamente por un inicio con los valores más altos de todos, pero con un paulatino decrecimiento en el tiempo. Sin embargo para este índice, muestra unos valores puntuales altos promediando la primera parte y al finalizar la sesión (muy similar al caso Brainstorming).

Aquellos valores altos finales de los casos señalados (**B**, **S** y, en menor medida, **A**) coinciden con el momento en el que los participantes quieren dar los últimos detalles a sus propuestas seleccionadas, de manera que se concentran específicamente en mejorar la elaboración de las estructuras finales, especificando los últimos detalles. Por ello el índice de elaboración contrasta con el de fluidez, sobre todo en el tratamiento **S**, que muestra valores nulos en los períodos finales del proceso.

³¹ Este análisis se encuentra en el Anexo 2, tabla A2.16

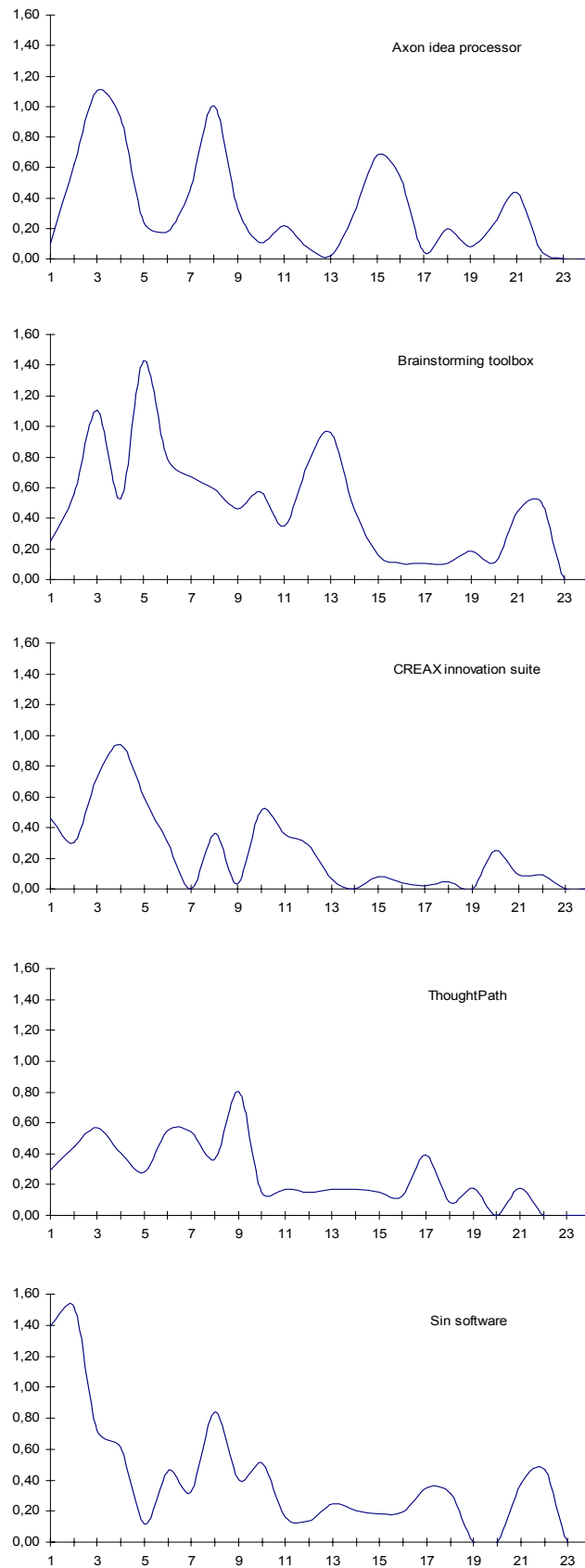


Figura 9.7 Comparación del índice de elaboración entre tratamientos

b. Elaboración total

El índice de elaboración para cada repetición y cada tratamiento se muestra en la Tabla 9.11. Nuevamente se presentan algunos valores que causan distorsión en la media, como son, en orden de importancia, los casos **B-2**, **T-3** y **A-2**. Por ello, el análisis de varianza con grupos homogéneos (del mismo tamaño) arroja una conclusión negativa para la diferencia de las medias.

Tabla 9.11 Índice de elaboración de cada participante para cada tratamiento

Programa	Participante			
	1	2	3	4
Axon	9,84	5,36	9,79	6,77
Brainstorming	14,67	3,67	13,90	12,59
CREAX	4,33	6,95	4,50	6,44
Thoughtpath	4,22	4,33	10,39	5,67
Sin-Software	8,34	7,98	10,87	10,87

La Figura 9.8 muestra el diagrama de cajas para esta variable, donde se evidencia los valores extremos comentados. Por ello se realizó un análisis de residuales estandarizados para determinar si aquellos son o no datos que puedan considerarse atípicos³², encontrándose que no son realmente valores estadísticamente atípicos.

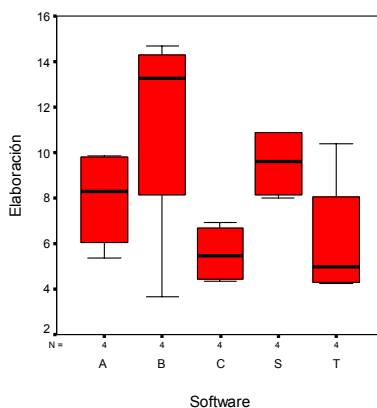


Figura 9.8 Diagrama de cajas para la variable índice de elaboración

La tabla de varianza para este índice permite concluir que no existe diferencias significativas (en el orden de $\alpha=0,05$) entre las medias. Los resultados se muestran en la Tabla 9.12.

³² El análisis de residuos estandarizados se encuentra en el Anexo 2.

Tabla 9.12 Análisis de varianza para el índice de elaboración, con valor atípico eliminado

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio cuadrados	F	Sig.
Entre grupos	87,803	4	21,951	2,513	,086
Dentro de los grupos	131,019	15	8,735		
Total	218,822	19			

Es posible afirmar con base en este análisis que estadísticamente el grado de elaboración de ideas es similar para todos los tratamientos, aunque existe una mejor tendencia para el caso **B**, que muestra una media superior a los otros. De hecho, sería significativa la diferencia si se admitiese un error tipo I del 8,5%.

9.4.4 Índice de originalidad

a. Comparación de la originalidad en el proceso

La Figura 9.9 muestra la variación durante el proceso del índice de originalidad, para cada uno de los tratamientos estudiados.

Se puede ver que los valores puntuales más altos (cerca de 1,2) se dan en los casos Axon idea processor y Sin software. En el primero, ocurre hacia la media hora de trabajo y se repite, aunque con un valor mucho más bajo, promediando el período 16 (trascurre hora y media); mientras que en el segundo sucede al iniciar la sesión y ya no se repite tan marcadamente.

El tratamiento con Brainstorming toolbox muestra un comportamiento más estable en el tiempo con valores entre 0,3 y 0,8 durante la mayor parte de la sesión y solamente disminuye al finalizar el trabajo. El CREAX por otra parte, evidencia valores más bajos, llegando a ser nulo en algunos momentos.

El ThouhtPath también evidencia valores decrecientes, como el CREAX, aunque es menos inestable, mostrando que sus valores altos se prolongan por un poco más de tiempo.

Se puede concluir que una vez más el tratamiento sin software muestra un valor inicial muy alto pero luego decrece siguiendo un comportamiento con valores bajos en el tiempo, mientras que en los tratamientos que utilizaron software, el Brainstorming nuevamente destaca por sus valores altos y una cierta estabilidad en el tiempo.

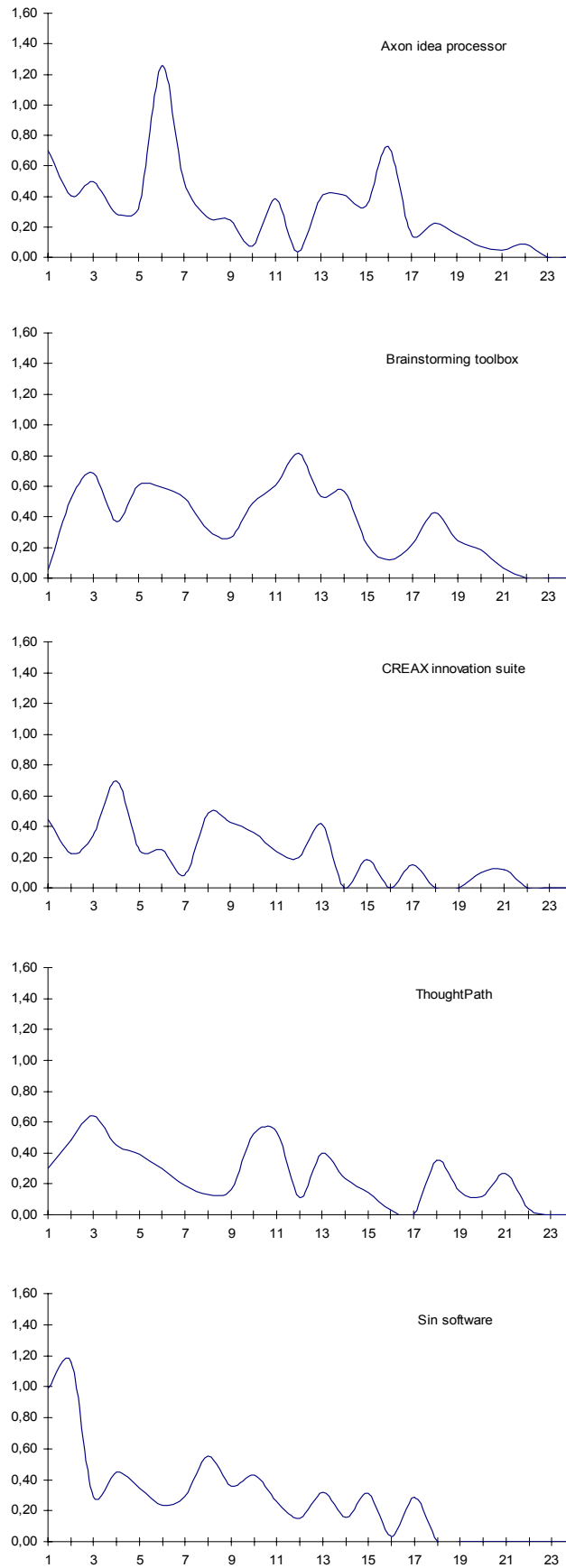


Figura 9.9. Comparación de la originalidad entre tratamientos

El hecho de que la curva de originalidad para el caso sin software llegue a valores nulos antes de finalizar la sesión, está indicando la tendencia de los participantes a concentrarse en detallar aspectos de sus propuestas previamente definidas, antes que en seguir pensando en nuevas posibilidades.

b. Originalidad total

La suma de los índices de originalidad obtenidos por cada participante en cada tratamiento se puede ver en detalle en la Tabla 9.13. A priori se puede ver que de mayor a menor grado la originalidad por tratamiento sigue el siguiente orden: Brainstorming toolbox, Axon idea processor, ThouhtPath, Sin-software y CREAX.

Tabla 9.13 Originalidad de cada participante para cada tratamiento

Programa	Participante			
	1	2	3	4
Axon	9,10	7,79	6,58	6,73
Brainstorming	9,82	8,82	8,81	6,23
CREAX	3,67	6,37	3,82	5,95
Thoughtpath	6,32	4,69	7,26	5,60
Sin-Software	5,47	5,00	7,75	8,27

Sin embargo, en la Figura 9.10 donde se presenta el diagrama de cajas para esta variable, también se aprecia mucha variabilidad interna en cada tratamiento.

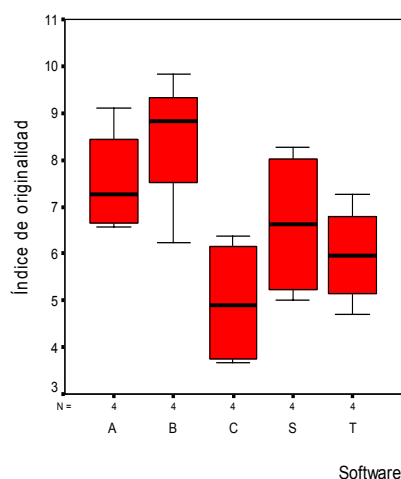


Figura 9.10 Diagrama de cajas para la variable índice de originalidad

Por eso, es necesario hacer un análisis de varianza para validar estadísticamente estas apreciaciones. La Tabla 9.14 muestra los resultados del análisis de varianza para este índice, mostrando que la hipótesis nula ($H_0 : \mu_1 = \mu_2 = \dots = \mu_n$) se rechaza, confirmado la diferencia de las medias entre tratamientos.

Tabla 9.14 Análisis de varianza para el índice de originalidad

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio cuadrados	F	Probabilidad
Entre grupos	29,095	4	7,274	3,811	0,025
Dentro de los grupos	28,628	15	1,909		
Total	57,722	19			

El detalle del análisis por comparaciones múltiples³³, evidencia que tales diferencias significativas suceden solo entre los tratamientos **B** y **C**, que son los que tienen valores extremos (altos el primero y bajos el segundo).

9.5 Análisis de la creatividad global

La creatividad global determinada de acuerdo con el modelo explicado en el capítulo siete, permite dar una lectura de las diferencias globales entre los diferentes tratamientos, cuando se analiza en forma comparativa. Para ello, se presentará dos análisis diferentes. El primero realizado por medio de la comparación entre los índices globales de creatividad y el segundo mediante la comparación con la creatividad individual determinada por el test CREA.

9.5.1 Creatividad global del proceso

En el capítulo ocho se presentaron estos valores en forma individual para cada participante en particular. Ahora se presentan en forma global por tratamiento, determinadas con base en los valores medios de todos los participantes, de manera que es posible hacer una comparación entre tratamientos.

La Tabla 9.15 muestra los valores para todos los tratamientos bajo estudio. La revisión de estos datos permite ver que la desviación estándar, sobretodo en los tratamientos **B** y **T**, son grandes. En particular en el tratamiento **T** se debió al caso 2 con un valor atípico, del cual ya se habló anteriormente.

³³ Ver este análisis en el Anexo 2, tabla A2.20

Tabla 9.15 Índice global de creatividad

Programa	CREATIVIDAD GLOBAL				Media	D.E
	1	2	3	4		
A	3,40	2,69	2,79	2,66	2,88	0,35
B	4,00	2,92	3,68	3,10	3,42	0,50
C	2,00	2,65	1,86	2,51	2,26	0,38
T	2,40	1,70	3,02	2,46	2,40	0,54
S	2,51	2,39	3,13	3,19	2,81	0,41

El análisis de varianza para este índice se muestra en la Tabla 9.16. Allí se puede concluir que existen diferencias significativas entre los valores de las medias, en particular entre los tratamientos B – C y B – T, confirmando el mejor desempeño del proceso utilizando el software Brainstorming Toolbox.

Tabla 9.16 Análisis de varianza para el índice global de creatividad

Origen de las variaciones	Suma de cuadrados	Grados de libertad	Promedio cuadrados	F	Probabilidad
Entre grupos	3,392	4	0,848	4,308	0,016
Dentro de los grupos	2,952	15	0,197		
Total	6.344	19			

La Figura 9.11 muestra la gráfica de cajas para esta variable con el fin de ilustrar mejor las diferencias de los datos obtenidos.

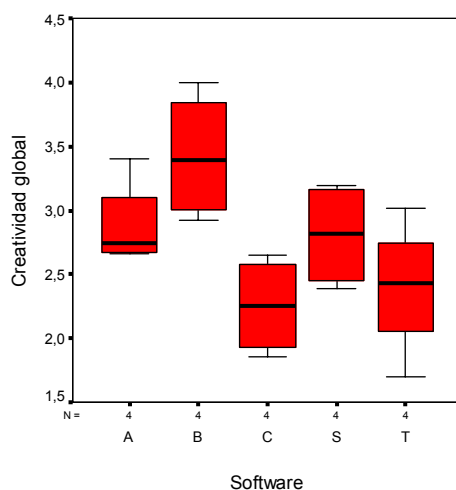


Figura 9.11 Diagrama de cajas para la variable creatividad global

9.5.2 Creatividad individual

La creatividad determinada a partir de la aplicación del test CREA (Corbalán, 2003), se presenta para cada participante en la Tabla 9.17. Aunque la variación entre índices es muy grande, se puede decir que el valor promedio de los diseñadores que trabajaron sin software fue mayor que en los otros casos y solamente el caso del ThoughtPath se aproxima. Los otros casos manejan en promedio valores similares. Esta característica refleja con mayor énfasis la influencia ejercida por los diferentes programas sobre el proceso creativo, teniendo en cuenta los análisis previos de producción y calidad de ideas, lo cual es una conclusión importante de este estudio.

Tabla 9.17 Índice de creatividad absoluta

Programa	Participante				Media
	1	2	3	4	
Axon	90	n.d	20	80	63,33
Brainstorming	80	n.d	97	30	69,00
CREAX	35	90	40	25	63,33
Thoughtpath	80	40	90	30	80,00
Sin-Software	80	50	75	55	86,67

9.6 Análisis de producción de ideas

Aquí se presenta un análisis comparativo entre cada tratamiento que utilizó software y el tratamiento sin-software. Con ello se pretende ver la influencia que ha tenido el respectivo programa en la generación de ideas, tomando como referencia o curva «testigo», la construida para el caso sin software. La Figura 9.12 muestra esta comparación en forma gráfica.

Los tratamientos **A** y **B** presentan un proceso de generación de ideas más activo que el tratamiento **S**, salvo durante los dos primeros períodos de tiempo (10 minutos). A partir de allí es notorio el predominio de los casos mencionados, particularmente en tres regiones concretas: entre los períodos 3 al 7 y del 11 en adelante. Esto se debe a la tendencia existente en el tratamiento **S** a seguir una trayectoria hiperbólica descendente, tal como se ha explicado, mientras que en los otros tratamientos, aunque va descendiendo tal tendencia no es tan marcada, presentando picos de producción de ideas importantes.

Para los otros dos tratamientos con software, **C** y **T**, el proceso en cuanto a generación de ideas es ligeramente más pobre. Estos dos siguen la misma tendencia decreciente mostrada por el tratamiento sin software y salvo por los “picos” mostrados entre los períodos 3 al 5 para el primero y 3 al 6 para el segundo, podría decirse que los tratamientos son equivalentes.

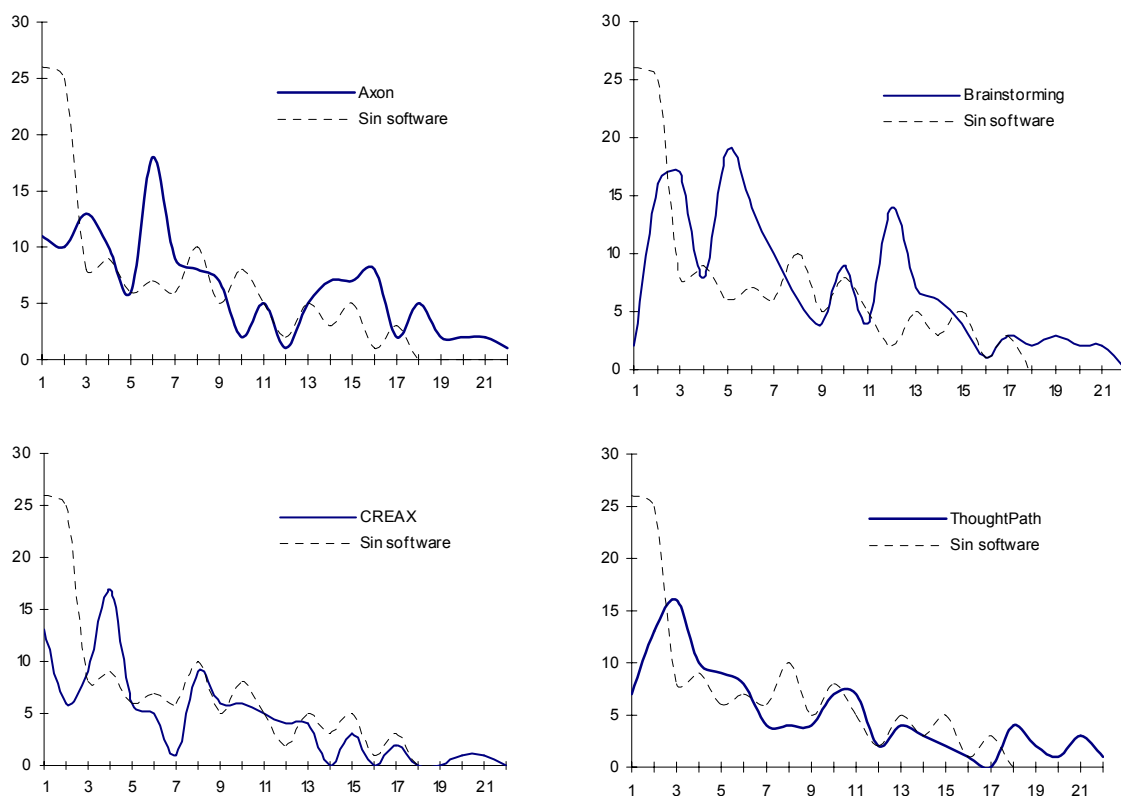


Figura 9.12 Curvas comparativas entre los tratamientos con software con respecto al tratamiento sin software

9.7 Conclusiones del capítulo

Se ha presentado un análisis comparativo detallado, entre los cinco tratamientos bajo estudio. Evidentemente no existe una diferencia marcada y regular para todas las variables estudiadas, de manera que no es posible afirmar categóricamente el predominio de algún tratamiento sobre los otros. Sin embargo, a lo largo de la discusión se presentaron las diferencias más significativas para cada variable, de manera que las particularidades representadas por ellas sí pueden llevar a conclusiones importantes a la hora de determinar características deseables en una herramienta de tipo informático para asistir el proceso de diseño de nuevos productos, en su fase conceptual.

Las diferencias de uso del tiempo entre los cinco tratamientos (cuatro utilizando software y uno sin software) se centran en las actividades de pensar, dibujar sobre papel y hablar. Efectivamente el no utilizar software libera al individuo para realizar otras actividades paralelas, como por ejemplo expresar sus ideas audiblemente y hacer gestos o imitaciones corporales para comunicar sus ideas. Mientras que el uso de software reclama mucha atención y concentración del usuario quien entonces se ve limitado para otras acciones externas al software.

Así, pues, una conclusión importante es que el software acapara buena parte de la atención del usuario limitándolo para realizar otras actividades. Se supone que entre mayor dominio tenga del manejo del software tal limitación se reduciría.

En cuanto al número total de ideas generadas, aunque estadísticamente no hay diferencia significativa en los valores medios, se aprecia una mejor eficacia con el uso del Brainstorming toolbox. Esta diferencia es particularmente clara en la formulación de modificadores funcionales, donde sí muestra estadísticamente ventaja, sobretodo sobre el CREAX. Esto se explica por dos razones. El primer programa ofrece la técnica que resultó más utilizada de todas: el «random picture» (figuras aleatorias) debido a que ésta facilita la realización mental de combinaciones, transformaciones y exploraciones en forma rápida y sin las limitaciones del lenguaje escrito. El segundo programa implementa la metodología TRIZ que es compleja de aplicar sin un entrenamiento y practicas exhaustivas.

A manera de conclusión se puede decir que los medios gráficos transmiten más y de mejor manera la información necesaria para establecer puentes, para realizar transformaciones y para explorar posibles soluciones.

En cuanto al origen de las ideas cada programa tiene sus propias virtudes y sus defectos. El tratamiento **B**, por ejemplo, el módulo Random Picture tuvo especial protagonismo, ya que propició el 51% de las ideas, lo que resalta la potencialidad de tal medio de estímulo frente a las demás opciones, menos efectivas debido al uso intensivo de elementos del lenguaje complicados de traducir del inglés al castellano. En el tratamiento **A** el 83% de las ideas corresponden a los mapas mentales, lo cual demuestra la poca utilidad de otras alternativas, básicamente por problemas de diseño de la interfase. Por su parte en el **C** el 47% de ideas obedece a la identificación de recursos del sistema, mientras que el resto se reparte con mayor equilibrio entre otros módulos. Destaca, en este caso, el hecho de la nula utilidad de la herramienta más conocida del TRIZ como es la matriz de contradicciones. El caso **T** a diferencia de los anteriores, muestra un equilibrio entre las diferentes fuentes de ideas. Esto se debió a la forma en que el programa guía al usuario a través de un proceso preestablecido y un tanto rígido.

La calidad creativa medida por medio de los cuatro índices revela un comportamiento a lo largo del proceso caracterizado por tendencias diferentes. Por ejemplo, para el caso de la flexibilidad los tratamientos **A**, **B** y **C** muestran mejores indicadores y tendencias más estables a través del tiempo, comparados con el tratamiento **S**, el cual evidencia una tendencia descendente en todo el proceso.

El índice de fluidez, por otra parte, marca diferencia entre los tratamientos **B** y **C**, a favor del primero de ellos. La diferencia en la cantidad de ideas generadas para los otros casos no significó diferencias estadísticamente relevantes. La razón para aquella diferencia se debe buscar en la complejidad del CREAX, software que utiliza la metodología TRIZ, mucho más elaborada y compleja que las otras técnicas estudiadas, por lo cual, hubiese sido preferible haber empleado mucho más tiempo de entrenamiento para este caso, para que los participantes llegasen a las sesiones experimentales con el conocimientos suficiente del manejo de la metodología.

En cuanto al índice de elaboración se encontró que con un error del 5% no hay diferencias entre los tratamientos, pero que la tendencia marca un mejor comportamiento para el caso **B**, hasta el punto de tener una probabilidad del 8.5% de error (muy cercano al 5% predefinido).

Para el último de los índices de creatividad, la originalidad, se evidencia un mejor comportamiento de los tratamientos **B** y **A**, existiendo una diferencia estadísticamente significativa entre el primero de ellos y el que mostró el peor de los índices que fue el **C**.

Finalmente, la trayectoria creativa seguida para los cinco tratamientos estudiados presenta algunas diferencias importantes. Es evidente que durante los primeros momentos de la sesión de diseño el tratamiento sin software muestra los mejores resultados. Pero a partir de los diez minutos, algunos de los tratamientos con software, específicamente el **B** y el **A** evidencian mejores valores. No así los casos **C** y **T**, que se parecen mucho al tratamiento sin software.

Así, se puede generalizar diciendo que el software ayuda en la medida que se conozca adecuadamente y que sea sencillo de utilizar. Por ello, casi todos los resultados muestran al tratamiento **B** como el mejor, siendo el más simple de todos los programas estudiados. Sus características de sencillez y de utilización de gráficos como una de sus herramienta preferida por los participantes, influyó definitivamente en los resultados. Puede esperarse, por lo tanto, que un programa diseñado especialmente para estos propósitos deba tener esta característica de sencillez.

La anterior afirmación no significa que los demás tratamientos que utilizaron software fueran muy diferentes. De hecho, entre estos y el tratamiento sin software existe muy poca diferencia, por lo que podría afirmarse que el uso de software no entorpece la labor del diseñador y, además, que en la medida que se tenga mejor manejo de los programas, pueden aportar una asistencia útil para la tarea de generación de conceptos.

Los resultados del análisis presentados en este capítulo, que pretendió ser exhaustivo con el tratamiento de los datos disponibles, permiten tener mayor claridad sobre las herramientas que debe y las que no debe tener un software de este tipo y sobre la estructura que pudiese tener para facilitar su uso y potenciar su efectividad, al menos en el entorno socio-cultural estudiado.