

R. 22582

UNIVERSIDAD COMPLUTENSE DE MADRID

Facultad de Psicología

Departamento de Metodología de las

CC. del Comportamiento

TP
1992
400

BIBLIOTECA UCM



530179692X

62777

**EVALUACION Y MEDIDA DE LA CARGA
DE TRABAJO MENTAL EN UNA TAREA
DE DIAGNOSTICO DE FALLOS**



NOTA: Este libro pertenece a la biblioteca de la Universidad Complutense de Madrid. No se permite su venta, alquiler, préstamo o cualquier otro uso que no sea el de lectura personal.

Susana Rubio Valdehita

Madrid, 1992

Colección Tesis Doctorales. N.º 400/92

b 1596476 s

© Susana Rubio Valdehita

Edita e imprime la Editorial de la Universidad
Complutense de Madrid. Servicio de Reprografía.
Escuela de Estomatología. Ciudad Universitaria.
Madrid, 1992.

Ricoh 3700

Depósito Legal: M-38527-1992



La Tesis Doctoral de D.^a SUSANA RUBIO VALDEHITA...

.....
Titulada ...EVALUACIÓN Y MEDIDA DE LA CARGA DE TRABAJO...
.....
MENTAL EN UNA TAREA DE DIAGNÓSTICO DE FALLOS.....

Director Dr. D.^o ROSARIO MARTINEZ ARIAS.....
fue leída en la Facultad de ...PSICOLOGÍA.....
de la UNIVERSIDAD COMPLUTENSE DE MADRID, el día ¹⁴.....
de JULIO de 19 ⁹²....., ante el tribunal

constituido por los siguientes Profesores:
PRESIDENTE .. DR. D. JOSE ALONSO FORTEZA MENDEZ.....
VOCAL DR. D. SALVADOR URRACA MARTINEZ.....
VOCAL DR. D. MANUEL MONTES PONCE DE LEON.....
VOCAL DRA. D.^a JULIA GARRIGA TRILLO.....
SECRETARIO .. DR. D. JESUS MARTIN GARCÍA.....

.....
habiendo recibido la calificación de *APTO. O.V.C.*
.....
LAUDE POR UNANIMIDAD.....

Madrid, a *14 de Julio* de 19*92*.
EL SECRETARIO DEL TRIBUNAL.

A. Urraca

**EVALUACION Y MEDIDA DE LA CARGA DE
TRABAJO MENTAL EN UNA TAREA DE
DIAGNOSTICO DE FALLOS**

TESIS DOCTORAL DE

SUSANA RUBIO VALDEHITA

DIRIGIDA POR

Dña. MARÍA DEL ROSARIO MARTINEZ ARIAS

**DEPARTAMENTO DE METODOLOGIA DE LAS CC. DEL COMPORTAMIENTO
FACULTAD DE PSICOLOGIA
UNIVERSIDAD COMPLUTENSE DE MADRID**

MAYO DE 1992

Dedicado a la memoria de mi abuela.

AGRADECIMIENTOS

Quiero expresar mi agradecimiento a todas las personas que han hecho posible esta tesis.

En especial, a la Dra. M^a del Rosario Martínez Arias por su encomiable labor de dirección y por haber orientado mi formación al campo de la Ergonomía.

A D. Ramón López Sánchez que me animó y apoyó siempre y que colaboró desinteresadamente para la finalización de esta tesis.

Por último, quiero mostrar mi agradecimiento a todos los alumnos que participaron como sujetos en los diferentes experimentos, pues sin su colaboración no habría sido posible la recogida de los datos.

INDICE

pag.

Dedicatoria

Agradecimientos

CAPITULO 1. DEFINICION DEL CONCEPTO DE CARGA DE TRABAJO MENTAL

1.1. INTRODUCCION.....	2
1.2. DEFINICION DE CARGA DE TRABAJO MENTAL.....	5
1.3. DIMENSIONES DE CARGA.....	6
1.4. IMPORTANCIA DE LA CARGA.....	9
1.4.1. Predicción de la carga.....	10
1.4.2. Evaluación de la carga del sistema actual.....	11
1.4.3. Evaluación de las diferencias del operador.....	11
1.5. RELACIONES ENTRE ATENCION Y CARGA MENTAL.....	12

CAPITULO 2. MODELOS DE CARGA

2.1. INTRODUCCION.....	16
2.2. MODELOS DE FILTRO.....	17
2.3. MODELOS SOBRE LA ARQUITECTURA DEL PROCESADOR CENTRAL.....	23
2.4. MODELOS DEL RECURSO.....	25
2.4.1. Modelos de recurso único.....	26
2.4.2. Modelo de recursos múltiples.....	30
2.4.2.1. Estados de procesamiento.....	36
2.4.2.2. Códigos de procesamiento perceptivo y central.....	37
2.4.2.3. Modalidades de input y de respuesta.....	38
2.4.3. Metodología de análisis de la carga propia de los modelos del recurso.....	40
2.4.3.1. Función ejecución-recurso (PRF).....	40
2.4.3.2. Performance Operating Characteristic (POC)...	42
2.4.4. Recurso único versus múltiple: la influencia de las prioridades.....	46
2.4.5. Recursos versus el procesador central.....	49
2.4.6. Estructura jerárquica de los recursos.....	50
2.4.7. Estrategias y carga.....	52
2.4.7.1. Relaciones recursos/estrategias.....	53



**CAPITULO 3. LA MEDIDA DE LA CARGA : PREDICCIÓN Y
TECNICAS DE EVALUACION**

3.1. INTRODUCCION.....	56
3.1.1. Criterios para la selección de los índices de carga.....	56
3.2. ANALISIS DE LINEA TEMPORAL.....	58
3.2.1. Limitaciones.....	58
3.3. MEDIDAS DE TAREA PRIMARIA.....	59
3.3.1. Limitaciones.....	59
3.4. LA TECNICA DE TAREA SECUNDARIA.....	61
3.4.1. Ventajas e Inconvenientes.....	66
3.5. MEDIDAS FISIOLÓGICAS.....	69
3.5.1. Ventajas y Limitaciones.....	71
3.6. MEDIDAS SUBJETIVAS.....	71
3.6.1. Ventajas y Limitaciones.....	78
3.7. RELACION ENTRE LAS MEDIDAS DE CARGA.....	80

**CAPITULO 4. LA CARGA MENTAL EN TAREAS DE DETECCIÓN Y
DIAGNOSTICO DE FALLOS**

4.1. INTRODUCCION.....	84
4.2. FACTORES QUE AFECTAN A LA DETECCIÓN Y DIAGNOSTICO DE FALLOS.....	86
4.2.1. Características de los displays.....	87
4.2.1.1. Compatibilidad entre el procesamiento central y el estímulo.....	89
4.2.1.2. Principio de compatibilidad de proximidad: Displays integrales versus separados.....	91
4.2.1.3. Displays gráficos versus alfanuméricos.....	95
4.2.1.4. Displays predictivos.....	98
4.2.1.5. Utilización del color: Efectos de la redundancia.....	98
4.2.2. Modelo interno.....	100

4.2.3. Estrategias de muestreo y búsqueda de información.....	104
4.2.3.1. Búsqueda Topográfica.....	104
4.2.3.2. Búsqueda Sintomática.....	107
4.2.4. Complejidad del sistema.....	113
4.2.5. Nivel de automatización del sistema.....	115
4.2.6. Relaciones de correlación y causalidad entre las variables de estado del sistema.....	117
4.2.7. Características de la tarea.....	118
4.2.8. Nivel y tipo de entrenamiento del operador.....	119
4.2.9. Tiempo disponible.....	121
4.2.10. Diferencias individuales.....	122
4.2.10.1. Habilidades y Aptitudes.....	123
4.2.10.2. Estilos Cognitivos.....	124
4.2.10.3. Variables de Personalidad.....	126
CAPITULO 6. CONCLUSIONES.....	128

CAPITULO 7. INTRODUCCION EXPERIMENTAL

7.1. INTRODUCCION Y OBJETIVOS.....	141
7.2. ELABORACION DE LOS MATERIALES.....	144
7.2.1. Pretest.....	144
7.2.2. Resultados.....	145
7.2.3. Selección de los items.....	147
7.2.4. Construcción de los materiales.....	148
7.3. PROCEDIMIENTO DE MEDIDA DE LA CARGA SUBJETIVA UTILIZADO.....	151

CAPITULO 8. EFECTOS DEL FORMATO DEL SISTEMA, DEL TIPO DE FEEDBACK, Y DE LA COMPLEJIDAD DE LA TAREA SOBRE LA CARGA. INFORMACION NUMERICA

8.1. INTRODUCCION.....	157
8.2. METODO.....	160
8.2.1. Sujetos.....	160
8.2.2. Diseño.....	160
8.2.3. Tarea.....	161
8.2.4. Procedimiento.....	162
8.2.5. Instrumentos.....	163
8.3. RESULTADOS.....	164
8.3.1. Análisis de la ejecución.....	164
8.3.1.1. Tiempo de tarea.....	164
8.3.1.2. Aciertos.....	170
8.3.1.3. Número de datos observados.....	171
8.3.2. Medidas de carga subjetiva.....	173
8.3.2.1. Índice global de carga subjetiva.....	173
8.3.2.2. Dimensiones de carga subjetiva.....	174
8.3.3. Diferencias individuales.....	179
8.3.4. Feedback como factor de aprendizaje.....	180
8.4. DISCUSION.....	182

CAPITULO 9. EFECTOS DE LA PRESION TEMPORAL

9.1. INTRODUCCION.....	189
9.2. METODO.....	192
9.2.1. Sujetos.....	192
9.2.2. Diseño.....	192
9.2.3. Tarea y Procedimiento.....	192
9.2.4. Instrumentos.....	193
9.3. RESULTADOS.....	193
9.3.1. Efectos de la presión temporal.....	193
9.3.2. Medidas de la ejecución.....	196

9.3.2.1. Tiempo de tarea.....	197
9.3.2.2. Aciertos.....	197
9.3.2.3. Número de datos observados.....	204
9.3.3. Análisis de la carga subjetiva.....	207
9.3.3.1. Índice global de carga subjetiva.....	207
9.3.3.2. Dimensiones de carga.....	207
9.3.4. Diferencias individuales.....	210
9.3.5. Feedback como factor de aprendizaje.....	211
9.4. DISCUSION.....	213

**CAPITULO 10. EFECTOS DE LA COMPLEJIDAD DEL PROBLEMA
Y DE LA PRESENTACION VERBAL DE LA
INFORMACION SOBRE LA CARGA**

10.1. INTRODUCCION.....	217
10.2. METODO.....	218
10.2.1. Sujetos.....	218
10.2.2. Diseño.....	218
10.2.3. Tarea y Procedimiento.....	219
10.2.4. Instrumentos.....	220
10.3. RESULTADOS.....	220
10.3.1. Comparación entre tipos de información.....	220
10.3.2. Medidas de la ejecución.....	225
10.3.2.1. Tiempo de tarea.....	227
10.3.2.2. Aciertos.....	229
10.3.2.3. Numero de datos observados.....	230
10.3.3. Medidas de la carga subjetiva.....	230
10.3.3.1. Índice global de carga.....	230
10.3.3.2. Dimensiones de carga subjetiva.....	231
10.3.4. Diferencias individuales.....	233
10.4. DISCUSION.....	235

**CAPITULO 11. EFECTOS DE LA PRESENTACION GRAFICA DE LA
INFORMACION Y DE LA INTRODUCCION DE UNA
TAREA SECUNDARIA**

11.1. INTRODUCCION.....	239
11.2. METODO.....	242
11.2.1. Sujetos.....	242
11.2.2. Diseño.....	242
11.2.3. Tarea y Procedimiento.....	243
11.2.4. Instrumentos.....	243
11.3. RESULTADOS.....	244
11.3.1. Efectos de la inclusión de la tarea secundaria.....	244
11.3.2. Efectos del tipo de información.....	248
11.3.3. Análisis de la ejecución.....	254
11.3.3.1. Tiempo total de tarea.....	254
11.3.3.2. Aciertos.....	256
11.3.3.3. Número de datos observados.....	258
11.3.4. Análisis de la carga subjetiva.....	259
11.3.4.1. Índice global de carga subjetiva.....	259
11.3.4.2. Dimensiones de carga subjetiva.....	259
11.4. DISCUSION.....	261

CAPITULO 12. EVALUACION DE LA TECNICA SUBJETIVA UTILIZADA

12.1. INTRODUCCION.....	265
12.2. ANALISIS.....	266
12.2.1. Análisis de varianza.....	266
12.2.2. Análisis de regresión logística.....	267
12.3. DISCUSION Y CONCLUSIONES.....	272

CAPITULO 13. CONCLUSIONES

13.1. PRESION TEMPORAL.....	278
13.2. TIPO DE DISPLAY.....	279
13.3. CARACTERISTICAS INDIVIDUALES.....	281
13.4. INCLUSION DE UNA TAREA SECUNDARIA.....	282
13.5. TECNICA SUBJETIVA DE MEDIDA DE LA CARGA.....	283
BIBLIOGRAFIA.....	285
APENDICE 1.....	326
APENDICE 2.....	335
APENDICE 3.....	342
APENDICE 4.....	349
APENDICE 5.....	356

**CAPITULO 1: DEFINICION DEL CONCEPTO DE CARGA DE
TRABAJO MENTAL**

1.1. INTRODUCCION

La Ingeniería psicológica (Engineering Psychology), también denominada Ergonomía, tiene sus orígenes en la Segunda Guerra Mundial, y está interesada en el problema general de la interacción hombre-máquina. Esta disciplina intenta ayudar al diseño de sistemas de forma que la eficacia del operador humano sea máxima. El diseño de los sistemas debe centrarse en las capacidades de procesamiento de la información humanas y acomodarse a las limitaciones del operador (Gopher y Kimchi, 1989). Como la sociedad actual es cada vez más tecnológica, muchos psicólogos han dedicado sus esfuerzos a acomodar los componentes del sistema a las personas. Es necesario diseñar el sistema de forma que las operaciones de mantenimiento sean fáciles, precisas y fiables (Christensen y Howard, 1981). El desarrollo tecnológico ha supuesto un aumento en el número de puestos de trabajo que incluyen actividades de procesamiento y de respuesta que requieren la codificación, organización, representación en la memoria y recuerdo de la información.

Cuando los humanos interactúan con las máquinas surge la cuestión básica de hasta qué punto ambos sistemas son compatibles. Welford (1968) estableció el problema de la compatibilidad como uno de traducción o conversión cognitiva, interna, que interviene entre los estímulos y las respuestas. Algunas conversiones parecen ser construidas mientras que otras, que inicialmente requerían mucha atención y pensamiento, son aprendidas (escribir a máquina, conducir un automóvil, etc.). En cualquier caso, el estudio de la cognición es necesario para poder comprender las destrezas humanas.

Existen tres grandes áreas de investigación psicológica relevante al estudio de la ejecución en actividades múltiples o complejas (Miyata y Norman, 1986):

Estudios sobre memoria: Especialmente los estudios sobre memoria de trabajo y la organización del conocimiento en la memoria a largo plazo.

Estudios sobre atención: Incluyendo los estudios de conducta automática y controlada, estudios de atención simultánea (¿cuántas actividades pueden realizarse al mismo tiempo?), y los de atención selectiva (¿qué actividades interfieren con otras?).

Estudios de la acción: Cómo las personas realizan las tareas.

Estas tres áreas de investigación están íntimamente relacionadas, aunque constituyen tres áreas distintas de estudio dentro de la Psicología. En la investigación sobre memoria, se distingue entre memoria de trabajo y memoria a largo plazo. En cuanto a la primera, lo más relevante para el tema de la ejecución en tareas complejas es reconocer su limitada capacidad, que según Norman (1986) es de 5 ítems a la vez, y que las distracciones (p.e. una llamada telefónica, la conversación de un compañero, etc.) pueden causar pérdidas en el material contenido en la memoria de trabajo (Klapp, 1987). Los aspectos más importantes de la memoria a largo plazo incluyen la dificultad para adquirir nueva información, y los problemas del recuerdo de la información adquirida. (Para una discusión de los modelos de memoria, ver Card, Moran y Newell, 1983; y para una discusión sobre los problemas del recuerdo de información, ver Norman y Bobrow, 1979 y Schank, 1982).

En cuanto a los estudios sobre atención, aquellos relativos a la conducta automática son los menos relevantes para la presente investigación. A partir de los estudios sobre conducta controlada (también denominada de control consciente) y de atención simultánea y selectiva, se ha desarrollado la mayor parte, o la totalidad, de los modelos de carga de trabajo mental que serán descritos posteriormente. El control consciente será utilizado fundamentalmente en aquellas situaciones en las que (1) la tarea ha de ser realizada es nueva para el sujeto, o está mal aprendida; (2) la tarea se percibe como especialmente crítica, difícil o arriesgada; (3) cuando es necesario resolver determinados conflictos imprevistos; o (4) en situaciones en las que la actividad se ve interrumpida por las demandas de otra actividad. Este último aspecto relaciona la conducta controlada con los estudios sobre atención simultánea y selectiva.

La relevancia de los estudios sobre la actividad humana para la investigación sobre la ejecución en actividades complejas es evidente.

Cuando un sujeto tiene que realizar una tarea múltiple o compleja, en función de los recursos de procesamiento que la tarea demande y de los que el sujeto posea, este suele enfrentarse con los efectos de la carga de trabajo mental (*mental workload*). La carga es un importante área problema de la investigación y desarrollo de sistemas hombre-máquina. Para evaluar soluciones alternativas en el diseño de sistemas es necesario medir, no sólo la ejecución del sistema, sino también la carga que sufre el operador humano. En los últimos años, la investigación aplicada ha demostrado considerable interés por el concepto de carga de trabajo mental: ¿En qué medida está ocupado el operador? ¿Hasta qué punto son complejas las tareas que tiene que realizar? ¿Puede manejar tareas adicionales a la vez que realiza su tarea actual? ¿Será capaz de responder a ciertos estímulos? ¿Cómo se siente el operador cuando realiza sus tareas?. El creciente número de artículos, libros y simposios en el campo (Leplat, 1978; Weldford, 1978; Moray, 1979; Roscoe, 1978; Williges y Wierwille, 1979; Wierwille y Williges, 1978, 1980) son testimonio de que el diseñador de sistemas destaca la carga como un aspecto de gran interés que debería ser estudiado. De hecho, un importante tema de debate entre directivos y sus empleados en la industria aérea ha sido el concepto de carga. La Federal Aviation Administration de los Estados Unidos pronto requerirá certificación del vehículo de vuelo en términos de su carga, y el Air Force también impone criterios de carga en los sistemas de nuevo diseño. Todos estos aspectos implican una muy relevante cuestión: ¿Qué es y cómo medir la carga?. Sin embargo, no existe acuerdo en la definición y medida de carga.

1.2. DEFINICION DE CARGA DE TRABAJO MENTAL

Cuando se examina el concepto de carga, la conclusión general que obtienen la mayoría de los autores es que se trata de un concepto **multidimensional** cuya definición no es fácil.

Actualmente, y desde el punto de vista de autores como O'Donnell y Eggemeier (1986) o Gopher y Donchin (1986), la carga de trabajo mental no viene dada unicamente por la dificultad de la tarea en sí misma (aspectos estructurales o físicos de la tarea), sino que surge de la **interacción** entre un operador y una tarea específica.

De este modo, la carga sería un constructo hipotético, resultante de la interacción entre una persona y una tarea. Las tareas se definen en términos de sus propiedades estructurales (estímulos, respuestas y el conjunto de reglas que adaptan las respuestas a los estímulos), por tanto es posible tener una serie de expectativas, relativas a la calidad de la ejecución, las cuales derivan del conocimiento de la relación entre la estructura de la tarea y la naturaleza de las capacidades y habilidades humanas. En la mayoría de los casos, los individuos conocen estas expectativas e intentan alcanzarlas a la hora de llevar a cabo la tarea, de forma que la carga de trabajo tiene lugar cuando las demandas de la tarea exceden las limitadas capacidades de la persona. La carga de trabajo mental es claramente una característica de los sistemas de procesamiento de la información y de control que median entre el estímulo, las reglas y las respuestas; es un atributo del conjunto persona-tarea, y sus efectos sobre la ejecución solamente pueden ser evaluados en relación a un modelo de procesamiento humano de la información. Los autores no toman en cuenta los efectos del aprendizaje puesto que, según ellos, lo interesante es conocer el empleo de las respuestas disponibles del sujeto en un momento determinado, sin considerar cómo ha sido creado su repertorio de habilidades.

Por otro lado, la carga de trabajo mental puede ser vista como la diferencia entre las capacidades del sistema de procesamiento de la información que son requeridas para ejecutar la tarea según lo esperado (nivel de ejecución esperado) y la capacidad disponible del sujeto en un determinado momento (nivel de ejecución real). El nivel esperado de ejecución se establece en función del nivel de ejecución obtenido en la misma tarea bajo las circunstancias menos demandantes. Esta definición implica que la carga es una variable interviniente en vez de un constructo hipotético (Navon, 1984).

Debido a que el único acuerdo sobre la definición de la carga de trabajo mental está en que se trata de un concepto multidimensional, gran parte de las investigaciones en este campo han intentado determinar los componentes o dimensiones que la forman.

1.3. DIMENSIONES DE CARGA

En 1973 Jahns (1973) divide el área global de carga en los tres siguientes atributos relacionados funcionalmente: carga del input, esfuerzo del operador, y ejecución. Conceptos similares de carga han sido propuestos por autores como Rohmert (1971), Rohmert y Laurig (1972) y por Rolfe y Lindsay (1973). Siguiendo este punto de vista, las fuentes de carga del input pueden ser ambientales (ruido, vibraciones, temperatura, etc.), inducidas por diseño o situacionales (características de displays y controles, etc.) y debidas a los procedimientos (instrucciones, secuencia y duración de tareas, etc.).

El esfuerzo del operador es función de la carga, del criterio de ejecución interno y del estado del operador (características psicofísicas, personalidad, experiencia, motivación, atención, etc.). Ejecución y esfuerzo deben ser hallados a la hora de evaluar sistemas hombre-máquina, mientras que la carga del input es generalmente un factor predeterminado.

A partir de una metodología de tarea secundaria, destacan los trabajos de Wickens y sus colaboradores (p.e., Wickens, 1980, 1983, 1984; Fracker y Wickens, 1989). Este grupo distingue diversas dimensiones de carga en función del tipo de recursos de procesamiento que demande la tarea, y que serán descritas en el capítulo siguiente.

En algunas ocasiones, se ha utilizado la técnica de análisis de escalamiento multidimensional (Kruskal y Wish, 1978) para determinar las distintas dimensiones o fuentes de carga, a partir de los datos sobre carga subjetiva obtenidos mediante cuestionarios y escalas (Derrick, 1981).

Siguiendo una metodología de análisis subjetivo de la carga (nivel de carga percibida por el operador), los autores distinguen entre diversas fuentes de carga. Por ejemplo, Hart, Childress y Bortolussi (1981) propusieron las diez dimensiones siguientes: carga global, dificultad de la tarea, presión temporal, ejecución, mental/sensorial, esfuerzo físico, frustración, nivel de estrés, fatiga, y tipo de actividad. La definición dada por los autores a cada una de estas dimensiones es la siguiente (Kantowitz, 1987):

- **Carga global:** La carga total asociada a la tarea, considerando todas las fuentes y componentes.
- **Dificultad de la tarea:** Si la tarea fue fácil o demandante, simple o compleja.
- **Presión temporal:** La presión sentida debida a la tasa de ocurrencia de los elementos de la tarea. ¿Se trata de una tarea lenta o rápida?.
- **Ejecución:** En qué medida el sujeto se siente satisfecho con el nivel de ejecución que ha alcanzado.
- **Mental/Sensorial:** Cantidad de actividad mental y/o perceptiva que fue requerida (p.e., pensar, decidir, calcular, recordar, mirar, buscar, etc.).

- . **Esfuerzo físico:** Cantidad de actividad física que fue requerida (p.e., pulsar, mover, empujar, tirar, girar, controlar, activar, etc.).
- . **Frustración:** En qué medida el sujeto se ha sentido inseguro, irritado, desagradado, frente a seguro, gratificado, contento.
- . **Nivel de estrés:** En qué medida el sujeto se ha sentido ansioso, preocupado, tenso, o calmado, tranquilo, plácido y relajado.
- . **Fatiga:** En qué medida el sujeto se ha sentido cansado, aburrido, agotado, o fresco, vigoroso, y energético.
- . **Tipo de actividad:** El grado al cual la tarea requería actuar en función de rutinas muy aprendidas, o de la aplicación de reglas, o requería toma de decisiones y solución de problemas.

Posteriormente, Hart y Staveland (1988) redujeron estas dimensiones a las seis siguientes: demanda mental, demanda física, demanda temporal, ejecución, esfuerzo, y nivel de frustración. Otros autores, Sheridan (1980), Sheridan y Simpson (1979) o Reid y col. (1981) han propuesto que las dimensiones anteriores se pueden reducir a tres: proporción de tiempo ocupado o carga de procesamiento de la información, esfuerzo mental ejercido o complejidad de la tarea, y "estrés emocional".

En esta misma línea, el grupo de colaboradores de Reid (Reid, Eggemeier y Shingledecker, 1982; Reid, Shingledecker y Eggemeier, 1981; Reid, Shingledecker, Nygren y Eggemeier, 1981) ha asumido que la carga está principalmente compuesta por tres dimensiones: carga debida al tiempo, carga debida al esfuerzo mental y carga debida al estrés. Estas tres dimensiones son adaptaciones de factores propuestos por Sheridan y Simpson (1979) y otros autores (Jahns, 1973; Johanssen y col., 1979; Kahneman, 1973; Moray, 1982), como principales fuentes de carga.

Mezclando las medidas de carga de tipo subjetivo con las de tarea secundaria y con las fisiológicas, Derrick (1981) evaluó las diferencias que percibían los operadores en la carga subjetiva de cuatro tareas, realizadas aisladamente, y después en todas las combinaciones de tarea dual posibles. Aplicando la técnica de

escalamiento multidimensional a los datos subjetivos, Derrick encontró tres dimensiones primarias: 1) competición de recursos, 2) adecuación del feedback, y 3) la medida de la variabilidad del pulso.

En resumen, el número de dimensiones de la carga que distinguen los autores suele venir determinado por supuestos a priori sobre las distintas fuentes de carga, más que partir del análisis estadístico de los datos. Con el tiempo, las distintas dimensiones de carga se han ido reduciendo, de manera que algunas de ellas han entrado a formar parte de una sola dimensión más general (Hart y Staveland, 1988). Aunque se ha propuesto un número variable de dimensiones, parece existir cierto acuerdo en que la carga, fundamentalmente la subjetiva, se debe a tres grandes áreas o fuentes. La primera englobaría todos los aspectos relativos a la presión temporal de la tarea (tiempo disponible, tiempo necesitado). La segunda estaría formada por variables que hacen referencia a la cantidad de recursos de procesamiento que demanda la tarea (mental, sensorial, tipo de tarea,..). Por último, la tercera dimensión general de carga se relacionaría con aspectos de naturaleza más emocional (fatiga, frustración, nivel de estrés, etc.).

1.4. IMPORTANCIA DE LA CARGA

Según los diseñadores y los operadores de sistemas, la ejecución no es siempre lo más importante en el diseño de un buen sistema. Es más importante considerar qué demandas impone una tarea sobre los limitados recursos del operador. Esto puede o no corresponder con la ejecución. Más específicamente, la importancia de la carga puede verse en tres contextos diferentes: la predicción de la carga, la evaluación de la carga impuesta por los equipos, y la evaluación de la carga experimentada por el operador humano. La diferencia entre el segundo y el tercer aspecto radica en sus implicaciones para la acción. Cuando se evalúa la carga de un sistema, la finalidad es optimizarlo. Cuando se evalúa la carga experimentada por el operador, el objetivo final es la selección y/o la formación de operadores.

La carga, en los tres contextos anteriores, puede ser inicialmente representada siguiendo un modelo simplificado de capacidad indiferenciada del sistema de procesamiento humano (el modelo de recursos se expondrá en el capítulo siguiente). Si la realización adecuada de una tarea demanda más recursos de los que el operador dispone, la ejecución fallará. Por otro lado, cuando los recursos son superiores a las demandas, la diferencia entre ambos representa la "capacidad residual".

Se asume que el concepto de carga está expresado en términos de la relación entre los recursos proporcionados y las demandas de la tarea. La capacidad de reserva y el nivel de ejecución son lo que principalmente interesa predecir a los diseñadores de sistemas. Es importante notar que, según esta concepción, los cambios del nivel de carga pueden deberse a las fluctuaciones de la capacidad del operador o a las variaciones de las demandas de recursos de la tarea.

A continuación se describen los tres aspectos en los que la evaluación de la carga es un aspecto clave.

1.4.1. Predicción de la carga

Frecuentemente es deseable conocer si un sistema será satisfactorio antes de que se proceda a su establecimiento final. El sistema debería ser capaz de realizar su misión adecuadamente, sin suponer excesivas demandas para el operador humano. Para tomar una decisión sobre cual de dos sistemas es mejor, es esencial tener algún modelo predictivo que adecue las demandas a las capacidades del operador y determine el grado al cual las primeras exceden o no dichas capacidades. Si las demandas exceden la capacidad del operador, entonces tendrán lugar fallos en la ejecución. Por contra, si las demandas no la exceden, es importante asegurarse de que el sistema está diseñado con un margen suficiente de capacidad residual o recursos, de manera que sucesos inesperados puedan ser resueltos satisfactoriamente.

Por tanto, a este nivel, los modelos de carga son totalmente necesarios.

1.4.2. Evaluación de la carga del sistema actual

Mientras que la predicción de la carga anterior al diseño de un sistema es algo muy deseable, frecuentemente es necesario evaluar la carga de un sistema ya existente. Esta evaluación puede ser realizada con el fin de identificar los periodos temporales o las operaciones en los que las demandas exceden momentáneamente los recursos disponibles.

Alternativamente, la carga puede ser evaluada para comparar dos componentes o piezas del equipo alternativas, que pueden dar lugar a niveles de ejecución similares pero que difieren en sus demandas. A veces los criterios de carga pueden ofrecer el único medio satisfactorio de elección entre alternativas.

Un mayor desafío es determinar el nivel absoluto de carga impuesto por un sistema que es superior o inferior a un criterio dado.

1.4.3. Evaluación de las diferencias del operador

Las medidas de carga pueden también evaluar las diferencias de la capacidad residual disponible de un operador, en lugar de las impuestas por el sistema. Esto puede hacerse en uno de dos contextos:

1. Puede compararse el nivel de destreza o de automatización alcanzado por diferentes operadores quienes pueden obtener niveles de ejecución equivalentes al realizar la tarea primaria, en situaciones de simulación. Por ejemplo, Damos (1978) y Crosby y Parkinson (1979) mostraron que los instructores de vuelo diferían de los pilotos estudiantes en su nivel de atención residual. Aún más, Damos encontró que esta medida aplicada a los estudiantes era un buen predictor del éxito en la formación del piloto.

2. Por otro lado, los operadores pueden ser observados durante la realización de tareas reales. En este caso, sistemas inteligentes basados en computador podrían decidir la responsabilidad para realizar ciertas tareas propias del operador cuando las demandas excedan momentáneamente su capacidad (Enstrom y Rouse, 1977; Wickens, 1979; Wickens y Gopher, 1977).

1.5. RELACIONES ENTRE ATENCION Y CARGA MENTAL

Al definir carga de trabajo mental en términos de las limitaciones del sistema de procesamiento de la información, se encuentra una gran afinidad entre la literatura dedicada a carga y aquella focalizada en atención. Las limitaciones atencionales cuando se realizan dos tareas a la vez parecen estar, como en el caso de la carga de trabajo, en la capacidad del sistema para manejar demandas múltiples. Al igual que con la carga, los investigadores han sugerido que las limitaciones atencionales reflejan restricciones inherentes de la estructura y organización de un **procesador central**. La mayoría de los métodos propuestos y empleados en el estudio de la carga representan intentos para identificar y cuantificar las limitaciones de este procesador central.

El término atención engloba un conjunto de fenómenos diversos, así Posner y Boies (1971) distinguen tres caracterizaciones de la atención: a) La atención como mecanismo de selección de la información; b) La atención como mecanismo de capacidad limitada; y c) La atención como mecanismo de alerta. La segunda concepción mencionada es la más directamente relacionada con la carga mental (Reason, 1990). La capacidad limitada de la atención se pone de manifiesto cuando se intenta la realización simultánea de dos tareas complejas, ya que, o estas se realizan de forma ineficaz, o los recursos atencionales deben distribuirse entre ellas, y por tanto, se produce interferencia y bajo rendimiento.

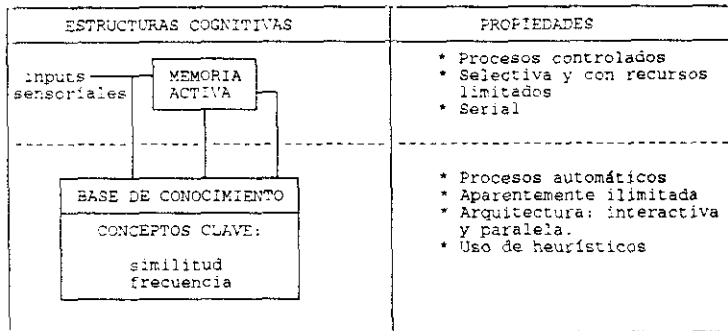


Figura 1.1. Propiedades de las estructuras asociadas con dos modelos del control cognitivo: control atencional asociado con la memoria activa y control central asociado con la base de conocimiento (Adaptado de Reason, 1990).

Con el tiempo se ha producido una convergencia entre los estudios de la memoria basados en la teoría multialmacén y las teorías de la atención (De Vega, 1984). El punto de convergencia se sitúa en la memoria activa (memoria operativa o de trabajo). Desde esta perspectiva, la memoria operativa se considera un espacio de trabajo de capacidad limitada, en donde se ejecutan procesos de control y coordinación. En este sentido, la memoria activa resulta indistinguible de los procesos atencionales controlados. En este punto, es relevante la distinción entre procesos automáticos y procesos controlados (Shiffrin y Schneider, 1977; Posner, 1978; Kahneman, 1973). Los procesos controlados a diferencia de los automáticos, consumen recursos atencionales y dan lugar a fenómenos de interferencia en situaciones de doble tarea.

Resulta especialmente significativo el hecho de que las investigaciones sobre la memoria activa y sobre la atención utilicen frecuentemente los mismos paradigmas experimentales. Por ejemplo, Baddeley y Hitch (1974) utilizaron una técnica experimental que consistía en sobrecargar la memoria activa del sujeto durante un periodo de tiempo en que éste realiza otra tarea concurrente de razonamiento abstracto.

El grado de interferencia entre la tarea de retención y la de razonamiento se consideraba como un índice de que ambas tareas demandan recursos del mismo sistema de capacidad limitada.

El paradigma de atención dividida (Gopher y col., 1982; Posner, 1978; Wickens, 1988) es esencialmente igual que el anteriormente mencionado, con la única diferencia de que los resultados se interpretan en términos de recursos atencionales limitados.

En conclusión, se puede afirmar que la carga mental es un índice de las limitaciones de la memoria activa o de los procesos atencionales controlados que refleja la sensibilidad del sistema a fenómenos tales como: a) el número de alternativas de respuesta o de ítems de información; b) la interferencia entre tareas simultáneas; c) la interferencia entre tipos de respuesta; y, d) el formato de la información de entrada.

CAPITULO 2. MODELOS DE CARGA

2.1. INTRODUCCION

En los últimos años del siglo pasado, James (1890) y sus contemporáneos identificaban las limitaciones del procesador central con las de la conciencia. Sin embargo, muchos de los procesos de interés clave para los teóricos del procesamiento de la información humano no son y no pueden ser accesibles a la conciencia (Kaufman, 1979; Broadbent, 1982; Posner y McCleod, 1982; Kahneman y Treisman, 1983). El fenómeno de la atención y el de la carga de trabajo mental engloban un conjunto de actividades de selección, transformación y procesamiento sólo en parte accesibles a la conciencia. Las limitaciones impuestas a la ejecución derivan de mecanismos que son opacos a la evaluación subjetiva, como por ejemplo la búsqueda en la memoria, la construcción de sentencias gramaticales y la focalización de la atención en uno de varios canales.

Con el paso del tiempo han ido surgiendo diversos modelos sobre el procesador central de capacidad limitada. Todos estos modelos tienen sus orígenes en la Teoría de la Información (Shannon y Weaver, 1949), y fundamentalmente utilizan los conceptos de canal de comunicación, flujo de información y capacidad del canal para desarrollar sus modelos de atención y carga.

Una gran variedad de estudios experimentales intentaron probar la aplicabilidad de la Teoría de la Información al análisis de la ejecución humana. Explícitamente, estos estudios fueron diseñados para evaluar las limitaciones en la capacidad del sistema de procesamiento. Para estos modelos, la dificultad de la tarea es equivalente a la carga del canal de comunicación. Entre los autores que aplican los conceptos y supuestos de la Teoría de la Información a la investigación psicológica destacan Fitts y Posner (1967), Garner (1970, 1978), Keele (1973), y Sheridan y Ferrell (1974), entre otros.

La Teoría de la Información tuvo éxito para modelar las situaciones más simples; sin embargo, no pudo describir muchos aspectos de la ejecución, como por ejemplo el efecto de anclaje. Otra dificultad de la conceptualización del sistema de

procesamiento humano como un canal de comunicación surgió al intentar generalizar las conclusiones de una investigación a otra.

Los estudios realizados bajo el marco de la Teoría de la Información han influido en todos los modelos de carga mental posteriores. De hecho, el concepto de capacidad de procesamiento limitada se debe a este paradigma. Sin embargo, la metáfora del canal de comunicación y las herramientas dadas por la Teoría de la Información para el estudio de la ejecución humana resultaron ser de limitada aplicabilidad.

2.2. MODELOS DE FILTRO

Los intentos de desarrollar modelos siguiendo estrictamente el esquema de la Teoría de la Información fueron abandonados a finales de la década de los 50, en favor de modelos acerca de los mecanismos internos que operan sobre la información y determinan la capacidad del canal. Los nuevos modelos se basaron en el estudio y análisis de uno u otro de los mecanismos responsables de las limitaciones observadas sobre la ejecución, olvidando el interés por la medida de la información.

Principalmente, las investigaciones se centraron en postular y demostrar empíricamente la existencia de filtros en el flujo de información. En su mayoría, estos estudios se focalizaron en la atención, la cual, como ya se ha visto, está íntimamente relacionada con el constructo de carga de trabajo mental y su desarrollo.

Los modelos de filtro representan un esquema del flujo de la información que comprende varios mecanismos, uno de los cuales está más constreñido que los otros. La capacidad de este mecanismo es la que determina los límites del sistema completo al compilar las demandas de la tarea.

El primero de estos modelos fue propuesto por Broadbent (1958). El constructo central de su teoría fue el de un canal de comunicación precedido por un filtro y un almacén a corto plazo. Este procesador se comunica con un almacén a largo plazo y

con un mecanismo que selecciona y controla las respuestas del sistema. El filtro actúa sobre la entrada de información, de forma que solamente aquella que es relevante es analizada al completo, lo cual afecta a la selección de la respuesta y al traspaso de información a la memoria a largo plazo.

Mientras que para los teóricos de la información el canal de comunicación es algo pasivo, Broadbent lo concibió como un mecanismo activo, el cual es un procesador flexible además de un transmisor de información. La función de este filtro selectivo es bloquear la información de entrada que no es buscada por el sujeto, dejando el paso libre sólo a aquellos mensajes que merecen un análisis cognitivo completo. El filtro sólo permite la entrada a los mensajes que tienen ciertas características físicas. Además, entre los sentidos y el filtro existe un almacén de memoria en el que se mantiene toda la información que no ha sido analizada, lo que permite que esta sea atendida más tarde.

Según este modelo, sólo la información consciente entra en la memoria a largo plazo y pasa a formar parte o a modificar nuestro conocimiento del mundo. Su modelo permite algún análisis perceptivo no consciente del input no atendido, pero este análisis es muy simple, sin permitir el análisis por el significado.

La memoria a largo plazo puede influir en el flujo de información a través de todo el sistema. Por tanto, la actuación del filtro está influida por propiedades del input y por la información almacenada en la memoria a largo plazo.

El modelo de Broadbent implica un procesamiento estrictamente serial. Su teoría aparentemente explica el procesamiento en paralelo como *time-sharing* o *multiplexing*: dos o más tareas pueden ser realizadas simultáneamente, simplemente distribuyendo la atención entre ellas. Por tanto, debería ser imposible realizar dos tareas en paralelo si ambas requieren atención continua.

Este modelo se focalizó en los fallos en la ejecución y, por tanto, es un modelo tanto de atención como de carga. Según este modelo, la carga impuesta a un sujeto por una o varias tareas está determinada por el grado al cual las demandas de las tareas exceden la capacidad del filtro. De esta forma, el filtro es un mecanismo activo y temprano que limita la carga de información del sistema.

El modelo sugiere que puede haber costos adicionales cuando las tareas requieren la integración de diversos atributos físicos. Treisman y Gelade (1980) encuentran que la identificación de una conjunción de atributos que requiere procesamiento en paralelo del estímulo en varias dimensiones impone una gran carga sobre el sistema.

Por otro lado, el sistema central de procesamiento puede ser capaz de gestionar tareas que requieren análisis semántico (identificación de palabras, categorización, o toma de decisiones) sin la protección que, según Broadbent, recibe del filtro.

En resumen, el modelo de Broadbent propuso varios principios generales que hacen posible describir el flujo de información en el sistema de procesamiento humano y que sirven como punto de referencia para gran parte de la investigación actual sobre la evaluación de la carga de trabajo. El sugirió que el camino entre estímulos y respuesta podía ser visto como siguiendo tres estados sucesivos y describió las propiedades funcionales de cada estado en términos que podrían ser aplicados al análisis de los componentes de las tareas. Además, asignó una función de costo de la ejecución de las operaciones mentales, identificando las operaciones más y menos costosas para el sistema. Finalmente, Broadbent enfatizó el estudio de la atención selectiva como parte inherente al análisis de la carga de trabajo al sugerir cómo la eficacia de la selección podía afectar al nivel de carga del sistema total.

Otro modelo de filtro es el de Welford (1967). La diferencia entre este y el anterior reside en que mientras el trabajo de Broadbent se focalizó en las diferencias cualitativas en la ejecución, como los fallos en el recuerdo, Welford se interesó más por los aspectos temporales de la misma. Welford diseña su teoría para explicar que la respuesta a un segundo estímulo se ve retardada si el sujeto no ha respondido

todavía al inmediatamente anterior. El atribuye este retraso a las limitaciones de un mecanismo central de decisión (selección tardía) que puede procesar sólo una tarea en cada momento. El tiempo requerido por este mecanismo para procesar un tarea se denomina periodo refractario psicológico (PRP): los datos procedentes de una señal que llegan a este mecanismo durante el tiempo de reacción a una señal previa tienen que esperar a que el mecanismo de decisión se encuentre libre, el cual se encuentra frecuentemente ocupado por el feedback de la ejecución de la respuesta. Sin embargo, el modelo también reconoce que se pueda producir una agrupación de estímulos cuando estos son presentados con un intervalo de tiempo muy corto (alrededor de 30 ms).

Los experimentos realizados por Welford y otros autores mostraron que estos retardos no eran eliminados con el entrenamiento y que estaban localizados en un procesador central mas que en los órganos periféricos sensoriales o motores (Fraise, 1957; Davis, 1959; Welford, 1959, 1967).

En cuanto a las conclusiones derivadas para el estudio de la carga, ambos modelos, el de Broadbent y el de Welford, descansan en gran medida en el paradigma de la Teoría de la Información, ambos proponen tres estados generales en el flujo de información y ambos localizan el filtro en la actuación del mecanismo central de procesamiento. Sin embargo, difieren sustancialmente en sus supuestos acerca de las reglas operacionales que gobiernan este mecanismo. El modelo de Broadbent permite, en cierto modo, el procesamiento en paralelo siempre que las demandas totales de la tarea no excedan el límite del procesador central. Por tanto, si las tareas que se están procesando en un determinado momento no exceden la capacidad del procesador, podrán entrar tareas adicionales en cualquier momento. Los decrementos en la ejecución que son manifestaciones de la carga de trabajo aparecen como tiempos de respuesta prolongados o como omisiones de respuesta. Por contra, el modelo de Welford es consistente con el procesamiento parcial y los errores de interacción.

Ambos modelos estimularon la investigación experimental pero fueron insuficientes para explicar el creciente cuerpo de datos resultante. El modelo de

Welford no pudo explicar, por ejemplo, que la variación de la dificultad asociada con la respuesta a un segundo estímulo afecte al patrón de respuesta de ambos estímulos (Karlín y Kastenbaum, 1968; Keele, 1973).

Otra crítica al modelo de Welford fue sugerida por Kahneman (1973). Según este autor la rapidez con la que un sujeto responde a un segundo estímulo no debe ser examinada como función del intervalo entre estímulos sino como función del intervalo entre respuestas. Cuando se utiliza esta última medida, el procesamiento en paralelo de ambos estímulos es más evidente. Es decir, el segundo estímulo no espera a que se completen el análisis y la respuesta al primero.

Igualmente, las pruebas experimentales de las predicciones derivadas del modelo de Broadbent demostraron la inconsistencia de sus primeras hipótesis y de la idea de que el filtro sirve para proteger al mecanismo central de la posible sobrecarga de las demandas de procesamiento semántico de alto nivel (Fairbank, Guttman y Miron, 1957; Treisman, 1965). Otra serie de estudios examinaron si el filtro descartaba, por características físicas, la información irrelevante. En experimentos de enmascaramiento se encontró que se podía procesar, por el significado, información irrelevante (Treisman, 1960; Moray, 1959, 1967). Además, se demostró que el recuerdo de información era más demandante que la entrada de inputs múltiples (Martín, 1970; Trumbo y Milone, 1971).

Estos hallazgos llevaron a varios investigadores a proponer un filtro tardío, en vez de temprano, en el flujo de información (Deutsch y Deutsch, 1963; Norman, 1968; Keele, 1973). En general, esta alternativa propone que la codificación y entrada de información no es tan demandante, de manera que puede ser procesada en paralelo. Las versiones actuales de esta aproximación se asemejan al modelo de filtro propuesto por Treisman (1964, 1969). El modelo de Treisman difiere del de Broadbent en dos aspectos generales: 1) asume que el filtro atenúa la entrada de información más que bloquearla, y 2) que el filtro está dotado de una estrategia flexible. Este tipo de filtro constituye una "estrategia de atención" más que una estructura fija: es decir, los filtros pueden actuar en diferentes puntos dependiendo de las condiciones en las que se aplica

una tarea y la naturaleza del ambiente de la misma. Según Treisman, el filtro a veces actúa sobre la base de las características físicas del input (como postula Broadbent) pero en otras ocasiones lo hace en función de propiedades más abstractas, como es por el significado. El filtro sería selectivo a lo largo de un continuo de características del input.

Desde este punto de vista la evaluación de la carga impuesta por una tarea debe expresarse en términos de los **componentes de procesamiento implicados**. Esta visión de las limitaciones del sistema de procesamiento implica un **concepto multidimensional de carga mental** puesto que existen varias razones del decremento en ejecución. Aunque el modelo de Treisman sigue siendo un modelo de canal simple, permite una definición más elaborada del mismo, cambiando considerablemente la visión anterior del concepto de canal.

Posteriormente, el interés de los modelos derivados de la Teoría de la Información se centró en el examen de la estructura interna del procesador central con un interés específico en su arquitectura. Uno de los conceptos centrales de este paradigma fue el de "estados de procesamiento". Las investigaciones muestran un sistema más complejo, en el que las limitaciones de procesamiento pueden aparecer en una gran variedad de nodos del sistema, y el coste de diferentes operaciones mentales emerge como un aspecto central al intentar **enlazar las limitaciones del procesador y las demandas de la tarea** (Sternberg, 1969; Sanders, 1980).

2.3. MODELOS SOBRE LA ARQUITECTURA DEL PROCESADOR CENTRAL

Surge el interés por la identificación de los múltiples componentes del sistema de procesamiento, cada uno de los cuales es concebido como una entidad funcional (cada componente tiene un papel específico en la transmisión de información). Así mismo, se establecen reglas simples que gobiernan la interacción entre estos componentes. Cada componente representa un estado en una secuencia de procesamiento. Según este modelo, la actuación de un estado no comienza antes de que los estados anteriores a él hayan completado sus operaciones (Sternberg, 1969; Sanders, 1980). Otros autores propusieron aproximaciones diferentes. Así, McClelland (1979) propone un modelo en cascada y Eriksen y Schultz (1979) un modelo de flujo continuo según el cual todos los estados de procesamiento operan continuamente transmitiendo información de un estado al siguiente a medida que la información se hace disponible.

El término estado de procesamiento se define como la agrupación de las estructuras de procesamiento que representan una operación mental común. En este sentido, ejemplos de estados de procesamiento son el preprocesamiento de un estímulo, la extracción de características, la elección de la respuesta y su programación y el ajuste motor (Sanders, 1983). El número de estados no es fijo, depende de los requisitos de cada tarea específica.

La herramienta experimental de estos modelos es la prolongación del tiempo de respuesta resultante de la manipulación de las variables de la tarea. En la mayoría de los estudios este análisis está basado en el paradigma de factores aditivos de Sternberg (1969). Según Sternberg, las variables que tienen un efecto aditivo sobre el tiempo de reacción deben operar en estados diferentes, mientras que las variables que interactúan influyen, al menos, en un estado común. La estructura general del flujo de información y la lógica de la metodología de factores aditivos fueron generalizadas a una gran variedad de tareas y a diferentes campos de estudio, como la evaluación de la carga de trabajo y la ejecución en tareas duales (Logan, 1978, 1979; Whitaker, 1979; Wickens, 1980) y el análisis de tareas (Mane, Coles, Wickens y Donchin, 1983).

Estos modelos implican una concepción **multidimensional de la carga mental**. El hecho de que las variaciones de una tarea limiten en forma diferente la ejecución se debe a que se requiera la actuación de uno o varios estados de procesamiento. El efecto de la variación de la estructura de la tarea y de sus demandas puede ser aditivo para algunas variables e interactivo para otras. También puede ocurrir que las demandas de una tarea no se vean afectadas por la manipulación de sus características estructurales. Una de las grandes ventajas de la técnica de factores aditivos es que proporciona una metodología para analizar la estructura de una tarea en función de sus variables.

En conclusión y de acuerdo a este punto de vista estructural, **dos tareas interfieren porque ambas demandan mecanismos de procesamiento o estructuras comunes** (estados de procesamiento, modalidades de input, requisitos de respuesta manual, etc.).

La complejidad de la estructura del procesador y el procesamiento en paralelo, propuestos por estos modelos, hacen muy difícil la medida de la carga. Sin embargo, se ha intentado solucionar este problema adoptando una metáfora nueva. Esta es la metáfora de la **energía**, o también denominada del **recurso**, la cual ha jugado un papel dominante en el análisis de la carga de trabajo mental. Durante los años 70, los tratamientos teóricos de Kahneman (1973), Norman y Bobrow (1975), y Navon y Gopher (1979) han contribuido inmensamente al desarrollo del concepto de capacidad o recurso como una variable implicada en la realización de tareas duales.

2.4. MODELOS DEL RECURSO

Una importante raíz histórica de la teoría del recurso descansa en la medida de la **carga de trabajo** del operador humano. Knowles (1963) presentó un modelo según el cual el sujeto posee un conjunto de recursos de capacidad limitada. A medida que una tarea primaria demanda más de estos recursos (se hace más difícil) menos recursos quedan disponibles para una tarea secundaria concurrente, y su ejecución se verá deteriorada.

Muchos psicólogos industriales adoptaron un modelo conceptual por el cual la carga de trabajo es proporcional a las demandas impuestas por la tarea sobre la capacidad limitada del operador (Rolfe, 1971). Así, desde finales de los años 70 se ha dado un gran interés a la investigación aplicada sobre la representación y medida de la capacidad utilizada y la disponible, y sobre la relación entre las medidas de carga de trabajo basadas en la capacidad y los índices alternativos como escalas subjetivas o parámetros fisiológicos (Moray, 1982; Moray, Johanssen, Pew, Rasmussen, Sanders y Wickens, 1979).

El concepto de capacidad (Moray, 1967) enfatiza la naturaleza flexible de los recursos de procesamiento. Las tareas demandan más de estos hipotéticos recursos ("esfuerzo mental") a medida que se tornan más difíciles o se requiere un nivel de ejecución más elevado. El concepto de recursos de procesamiento es propuesto como una variable hipotética que permite tomar en cuenta las variaciones que se producen en la eficacia con la que se realizan actividades al mismo tiempo; esto es, el grado al que dos tareas pueden ser realizadas a la vez del mismo modo que aisladamente. Se asume que las tareas demandan una serie de recursos para poder ser realizadas, siendo dichos recursos de disponibilidad limitada. Así, cuando el conjunto de recursos demandados por dos tareas excede a los disponibles, la eficacia en tiempo compartido disminuye, descendiendo más a medida que aumenta la dificultad de alguno de los componentes de la tarea.

El concepto de atención como un recurso de procesamiento flexible de capacidad limitada fue elaborado por Kahneman (1973), Norman y Bobrow (1975) y Navon y Gopher (1979).

Se distingue entre modelos de recurso único y modelos de recurso múltiple. Ninguno de ellos representa un modelo estructural del sistema de procesamiento de la información; son afirmaciones sobre la posibilidad de tratar las limitaciones del sistema como si este dependiera de la disponibilidad de algún hipotético recurso. En este sentido, no es muy necesario conocer por qué un operador falla al realizar una tarea, sino que lo importante es cuantificar la diferencia entre la ejecución esperada y la obtenida. En la mayoría de las investigaciones que se basan en el concepto de capacidad, se asume que la capacidad utilizada al realizar una tarea primaria puede ser inferida a partir de la capacidad residual de la tarea secundaria (Eysenck y Eysenck, 1979; Posner y Keele, 1969; Posner y Boies, 1971; Tyler, Herta, McCalum y Ellis, 1979; Underwood, 1976). Por ejemplo, se presume que tiempos de reacción más largos, reflejan mayores demandas de capacidad (menor automatización, procesamiento más profundo) de la tarea primaria. La técnica de tarea secundaria ha sido utilizada en varios estudios sobre los requisitos atencionales de varios procesos de memoria (Johnson, Greenberg, Fisher y Martin, 1970).

2.4.1. Modelos de recurso único

El intento más comprensible para emplear la metáfora de la energía en el estudio de la carga mental y de la atención aparece en el libro de Kahneman "Atención y Esfuerzo" (1973). Kahneman (1973) propone la existencia de un único e indiferenciado conjunto de recursos disponible para todas las tareas y actividades mentales. A medida que aumentan las demandas de una tarea, impuestas por un aumento en su dificultad o por la adición de tareas adicionales, los mecanismos de arousal fisiológico producen un incremento en la provisión de recursos. Sin embargo, en ocasiones, este incremento es insuficiente para compensar por entero el aumento

de las demandas: así, la ejecución descenderá a medida que aumente el desfase entre las demandas y los recursos provistos. Esto es, el límite de los recursos varía con el nivel de arousal de acuerdo a la clásica función de U invertida que relaciona la eficacia de la ejecución con la activación. Los cambios en la capacidad están controlados por el feedback procedente de la ejecución. Por lo tanto, en la formulación de Kahneman, el concepto de recursos de procesamiento está estrechamente relacionado con el de arousal.

Los sujetos pueden focalizar toda su atención en una de las actividades posibles o pueden dividirla entre varias de ellas. La política de asignación de recursos está gobernada por reglas muy aprendidas o inconscientes (*Enduring Dispositions*) y por decisiones conscientes (*Momentary Intentions*). La política de asignación de recursos está también determinada por la dificultad de la tarea y el feedback de las actividades que se están realizando (*Evaluation of Demands on Capacity*). Los sujetos focalizan más su atención en las tareas más difíciles.

En un modelo estructural, los fallos en la ejecución ocurren cuando se requiere que un mecanismo realice operaciones incompatibles. En un modelo basado en la metáfora de la energía, los decrementos de la ejecución se deben a que las demandas de dos actividades concurrentes exceden la capacidad disponible. Así, un modelo estructural implica que la interferencia entre dos tareas es específica, mientras que para un modelo de capacidad esta interferencia es no específica y sólo depende de las demandas totales de las dos tareas.

Por tanto, en el modelo de Kahneman, el procesamiento en paralelo podría ocurrir cuando la capacidad de procesamiento combinada de varias tareas es menor que la capacidad total de la persona. Mientras las tareas no excedan la capacidad de procesamiento del sujeto, estas podrán ser realizadas simultáneamente sin pérdida de eficacia en ninguna de ellas.

El modelo de asignación de recursos de Kahneman sugiere cuestiones como: si la dificultad de la tarea afecta a la asignación de recursos y la tarea se torna más

fácil con la práctica. ¿incrementa la práctica la capacidad?, ¿pueden, con la práctica, realizarse dos tareas simultáneamente?. En este sentido, Underwood (1974) y Ostry, Moray y Marks (1976), entre otros, mostraron como la práctica produce cambios en la asignación de recursos atencionales.

Este concepto de recurso de procesamiento implica la utilización de medidas fisiológicas. En este sentido, destacan estudios que utilizan el diámetro pupilar como medida del esfuerzo (Kahneman y col., 1967, 1968, 1969, 1981; Beatty, 1982). Datos procedentes de una gran variedad de condiciones experimentales han mostrado que la ejecución de algunas tareas interfiere con la realización de cierto tipo de tareas pero no con las de otro, mientras que una tercera clase de tareas se ve igualmente afectada cuando se realiza con miembros de los dos primeros grupos (Navon y Gopher, 1979; Ogden, Levine y Eisner, 1979; Wickens, 1980, 1983). Otros estudios han mostrado que algunas manipulaciones de las características de tareas que se realizan a la vez afectan a ambas, mientras que otras degradan la ejecución solamente de una de ellas, no pudiendo ser compensada por la transmisión de recursos de una a otra tarea (Wickens y Kessel, 1981; Gopher y Navon, 1980; Gopher, Brickner y Navon, 1982). Estos resultados son inconsistentes con la noción de un único y no diferenciado conjunto de recursos o energía de procesamiento. Más bien, ellos sugieren la existencia de varias fuentes de interferencia y competición más específicas. De nuevo, el concepto de una única y general causa de las limitaciones en la ejecución atribuidas a la carga no parece adecuado.

Como ya se ha visto anteriormente, la teoría de la capacidad ha sido desarrollada en la investigación sobre tareas duales. Hay cuatro fenómenos que difícilmente pueden ser explicados por el modelo del recurso simple. Estos son los siguientes:

1. **Insensibilidad a la dificultad:** Se pueden citar varios ejemplos en los cuales los incrementos en la dificultad o demandas de la tarea primaria (presumiblemente mayor consumo de recursos) no influye en la ejecución de

una tarea secundaria. Es necesario asegurar que el decremento en la tarea primaria causado por la realización de la secundaria no aumenta en función de los niveles de dificultad de la tarea primaria. Ejemplos de insensibilidad a la dificultad pueden encontrarse en los trabajos de North (1977), Isreal, Chesney, Wickens y Donchin (1980), o Wickens y Kessel (1979).

2. **Ejecución perfecta en tareas concurrentes:** Ejemplos de realización perfecta de dos tareas a la vez pueden encontrarse en los trabajos de Allport, Antonis y Reynolds (1972, Shaffer (1975), Spelke, Hirst y Neisser (1976) y Wickens (1976), entre otros.

Una posible explicación a los efectos de insensibilidad a la dificultad y a la ejecución perfecta en tareas realizadas al mismo tiempo es que las tareas utilizadas contengan amplias regiones limitadas de datos. Sin embargo, esta explicación no parece plausible ya que en los estudios mencionados se utilizaron tareas que dudosamente pueden ser muy limitadas de datos (tareas de seguimiento (*tracking*), de procesamiento de dígitos, de detección de señales auditivas, de enmascaramiento auditivo, etc...).

3. **Efectos de la alteración estructural:** Estos efectos hacen referencia a los casos en los que la modificación en la estructura de procesamiento (modalidad de display, código de memoria, modalidad de respuesta) produce un cambio en la interferencia con una tarea concurrente (siempre que la dificultad se mantenga constante). Así, en cuanto a las variaciones en la modalidad del input destacan los estudios realizados por Treisman y Davies (1973), Isreal (1980), Martin (1980), Rollins y Hendricks (1980), Vidulich y Wickens (1981) y Wickens y col. (1983). Igualmente, estudios sobre la modalidad de respuesta son los de McLeod (1977), Harris, Owens y North (1978), Wickens (1980) y Wickens y col. (1983). Finalmente, estudios sobre los efectos del cambio en el código de procesamiento son los realizados por McFarland y Ashton (1978),

Friedman, Polson, Dafoe y Gaskill (1982), Wickens y Sandry (1982) y Wickens y col. (1983). Si la dificultad de la tarea modificada no cambia, las demandas de recursos deberían ser similares o idénticas para todas las tareas. Por lo tanto, bajo el supuesto de la existencia de un recurso no diferenciado no se podría predecir ningún cambio en la interferencia de tareas concurrentes.

4. **No emparejamiento entre dificultad y estructura:** Este efecto fue notado por Wickens (1976) y hace referencia a aquellos casos en los que la más difícil de dos tareas interfiere menos con una tercera tarea de lo que lo hace la más fácil.

2.4.2- Modelo de recursos múltiples

La debilidad del modelo anterior lleva al planteamiento de modelos de recurso múltiple. Según estos modelos, el sistema humano posee un número de mecanismos de procesamiento cada uno de los cuales requiere su propia provisión de recursos. La capacidad de cada una de estas estructuras depende del nivel de arousal y de su dependencia específica de dicho nivel, y puede ser desplegada para varias tareas. De este modo, existe una competición continua entre las tareas que requieren los mismos recursos (Norman y Bobrow, 1975; Navon y Gopher, 1979; Wickens, 1980, 1983; Sanders, 1983; Gopher y Sanders, 1984; Fracker y Wickens, 1989).

Norman y Bobrow (1975), los primeros en utilizar el término recursos, definieron, utilizando una analogía global con los sistemas de computador, los recursos de forma muy general, con referencia a todas las facilidades de procesamiento.

Navon y Gopher (1979, 1980) no fueron más específicos en su análisis teórico inicial, en el cual utilizaron una metáfora económica. En su posterior trabajo experimental (Gopher y Brickner, 1980; Gopher y col., 1982) concluyeron que es razonable asumir la existencia de al menos dos tipos de recursos relativamente independientes. Uno está relacionado con los procesos perceptivos y computacionales, y el otro con la selección y la generación de la actividad motora.

Las figura 2.1. y 2.2. muestran *gráficamente* las ideas centrales del modelo de recursos múltiples propuesto por Navon y Gopher (1979), en referencia a las dos medidas de ejecución más usuales: latencia y precisión de la respuesta.

En las figuras se puede apreciar que, como postula el modelo, para una misma cantidad de recursos de procesamiento invertidos por el sujeto en la realización de la tarea (R_0), el aumento en el nivel de carga de la tarea produce un incremento en la latencia de la respuesta (figura 2.1), mientras que disminuye su precisión (figura 2.2.). Además, la figura 2.2. muestra que el aumento en el tiempo disponible para realizar la tarea mejora la precisión de la respuesta.

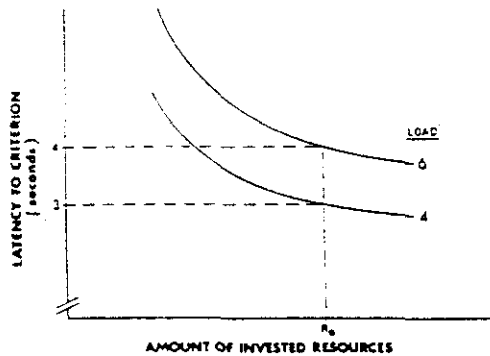


Figura 2.1. Representación de dos funciones latencia-recurso. Tomado de Navon y Gopher (1980).

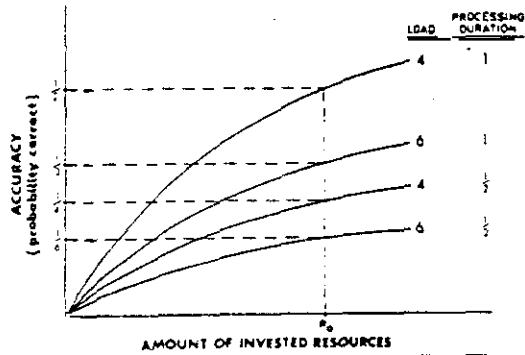


Figura 2.2. Representación de cuatro funciones precisión-recurso. Tomado de Navon y Gopher (1980).

Las implicaciones de esta teoría en cuanto a la ejecución de tareas duales son cuatro:

1. En la medida en que dos tareas demanden más recursos diferentes que comunes, ellas podrán ser realizadas a la vez más eficazmente.
2. Si dos tareas demandan recursos comunes la POC será una curva suave; mientras que en el caso contrario se generará entre ambas tareas una POC (ver apartado 2.4.3) en forma de caja (ver figura 2.3). Esto es así porque si los recursos de ambas tareas son independientes entre sí, los recursos de una de las tareas no pueden ser utilizados para mejorar la ejecución en la otra. Por lo tanto, cuando dos tareas concurrentes demandan recursos totalmente diferentes

tiene lugar la realización perfecta de ambas. Además, el desemparejamiento entre estructura y dificultad tendrá lugar cuando se compara la ejecución en dos tareas muy demandantes pero de recursos diferentes con la de otras dos menos demandantes pero de recursos comunes.

3. El cambio en la dificultad de una tarea se define como el incremento en la demanda de uno o más de los recursos de los que depende la ejecución. Si parte de estos recursos son también requeridos para realizar una tarea concurrente, esta última se verá afectada. Si, por el contrario, estos recursos afectados por la modificación de la dificultad de la tarea no son necesarios para realizar la tarea secundaria, esta no se verá afectada. De este modo, la insensibilidad a la dificultad tiene lugar cuando los recursos adicionales no pueden ser transferidos desde la tarea secundaria para compensar las demandas impuestas sobre la tarea manipulada, o cuando los recursos son transferidos pero la ejecución de la tarea manipulada no puede beneficiarse de su mayor disponibilidad.

4. Si una tarea se hace más difícil porque demanda más recursos de los compartidos con otra concurrente, la ejecución será más dependiente de dichos recursos, y en consecuencia, habrá un mayor efecto de las prioridades en la asignación de recursos. En otras palabras, el efecto de la dificultad debería interactuar con el de las prioridades siempre que la dificultad surja de la *manipulación de la demanda de los recursos comunes*. La extensión a la que las tareas comparten un recurso común dicta el grado al cual la ejecución en la tarea primaria (para la cual se varía la dificultad) puede ser preservada mediante la reasignación de los recursos a dicha tarea en contra de la ejecución de la tarea secundaria.

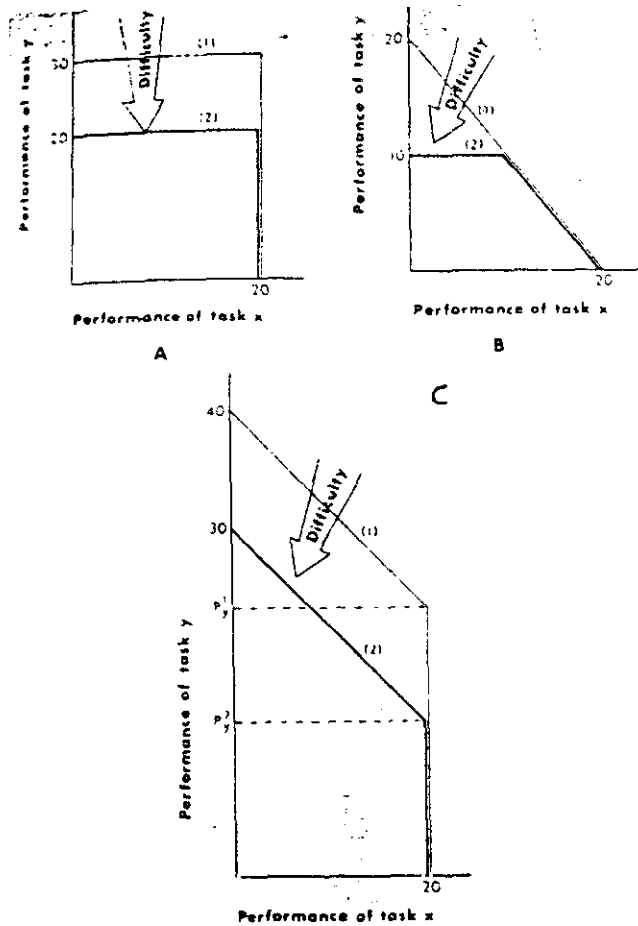


Figura 2.3. Ilustración de tres posibles efectos de la manipulación de un parámetro de los recursos relevantes sólo a la ejecución de la tarea Y. A representa el caso de recursos completa ausencia de recursos comunes entre las tareas. B representa el caso de existencia de recursos tanto comunes como independientes utilizados en proporciones fijas. C representa el caso de existencia de recursos tanto comunes como independientes utilizados sustitivamente. Tomado de Navon y Gopher (1979).

La naturaleza de los recursos es una cuestión central en el trabajo de Wickens (1980, 1981, 1983, 1988) y sus colegas (Isreal, Chesney, Wickens y Donchin, 1980; Fracker y Wickens, 1989; Vidulich y Wickens, 1983; Wickens y col., 1981, 1982, 1983). Examinando un gran número de estudios sobre tarea dual en los que se producían efectos de la alteración estructural e insensibilidad a la dificultad, Wickens (1980) encuentra que los recursos pueden ser definidos a partir de tres dimensiones: a) *estados de procesamiento* (*perceptual-central versus respuesta*), b) *códigos de procesamiento perceptivo y central* (*verbal versus espacial*) y c) *modalidades de input* (*visual versus auditivo*) y *respuesta* (*manual versus oral*). Asumiendo que las respuestas manuales generalmente están guiadas espacialmente y las orales son verbales, las dimensiones modalidad de respuesta y código de procesamiento central son similares. Si este es el caso, la "estructura" de los recursos puede representarse conceptualmente según la figura 2.4.

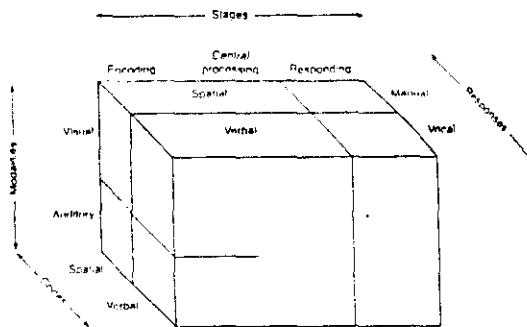


Figura 2.4. Estructura de recursos de procesamiento. Tomado de Wickens (1984).

La clasificación por estados de procesamiento tiene su procedencia en la metodología de factores aditivos (Sternberg, 1969; Sanders, 1979, 1980). Los códigos de procesamiento son sugeridos a partir del análisis teórico y trabajo experimental que ve a los hemisferios como reservas de recursos que actúan separadamente en virtud de su separación funcional y estructural (Kinsbourne y Hicks, 1978; Friedman y Polson, 1981; Friedman, Polson, Defoe y Goskill, 1982). Los estudios que compararon las formas de presentación visual y auditiva y los modos de respuesta verbal versus manual en los paradigmas de tarea dual (Gopher, Brickner y Navon, 1982; McLeod, 1972, 1978; Wickens, 1988; Wickens y Kessel, 1981; Wickens y Sandry, 1982; Wickens, Sandry y Vidulich, 1983) son la justificación de las modalidades de procesamiento y de respuesta como criterios de clasificación.

Para organizar estas categorías, Wickens (1981, 1983) propone un esquema descriptivo tridimensional que incluye una distinción entre los tipos de códigos de representación (verbal y espacial). A continuación se describen estas tres dimensiones:

2.4.2.1. Estados de procesamiento

Según esta dimensión, el procesamiento perceptivo y el central requieren recursos comunes y estos, en cambio, son funcionalmente independientes de los recursos propios de los procesos de respuesta. Evidencia de esta afirmación se obtiene cuando la dificultad de respuesta de una tarea es manipulada sin que esto afecte a la ejecución de una tarea concurrente para la que las demandas son más cognitivas o perceptivas. Esta insensibilidad a la dificultad ha sido demostrada en los experimentos realizados por Isreal, Chesney, Wickens y Donchin (1980) y por Isreal, Wickens, Chesney y Donchin (1980), que utilizaron medidas fisiológicas (potenciales evocados) de la carga mental. En estos trabajos, la amplitud del componente P300, que se supone depende de recursos de procesamiento perceptivo y central, se vio afectada por las manipulaciones de la carga perceptiva de una tarea concurrente, pero no quedó afectada por los requisitos para generar respuestas manuales ni por las manipulaciones de la frecuencia de respuesta en una tarea concurrente de seguimiento.

La relevancia de esta dimensión también quedó demostrada por Wickens (1976), en cuyo experimento se produjo un no emparejamiento estructura-dificultad cuando se realizaban a la vez tareas de detección de señales y de seguimiento. Otra evidencia resulta de la insensibilidad a la dificultad encontrada por Kantowitz y Knight (1976) y Wickens y Kessel (1980). Aún más, Shaffer (1971, 1975) comprueba que procesos perceptivos, de traducción y de respuesta, pueden proceder en paralelo de forma muy efectiva. Finalmente, parece que la manipulación de la competición de recursos entre tareas al nivel de codificación es relativamente independiente de esta manipulación al nivel de respuesta (Vidulich y Wickens, 1981; Wickens, 1980; Wickens y col., 1983). Una independencia de este tipo de los efectos de codificación y de respuesta indica una independencia de los recursos correspondientes.

2.4.2.2. Códigos de procesamiento perceptivo y central

La idea de que los procesos espacial y verbal pueden ser descritos a partir de recursos funcionalmente diferentes, y que estos pueden relacionarse anatómicamente en la mayoría de los sujetos con los hemisferios cerebrales derecho e izquierdo respectivamente, está avalada por la investigación y teoría de Kinsbourne y sus colegas (Kinsbourne y Hicks, 1978). En un típico experimento, ellos observaron una mayor interferencia de una tarea verbal con otra de equilibrio cuando esta última era realizada con la mano derecha (controlada por el hemisferio de procesamiento verbal) que con la izquierda (controlada por el hemisferio espacial). McFarland y Ashton (1978) mostraron que esta asimetría manual de interferencia se invertía cuando la tarea verbal era sustituida por una de memoria espacial. Por tanto, existen condiciones en las que las funciones de procesamiento y de respuesta están bajo el control de diferentes hemisferios. Demostraciones similares de esta interferencia "específica de código" se encuentran en Baddeley, Grant, Wight y Thompson (1975) y Baddeley y Lieberman (1980).

En el dominio perceptivo, Moscovitch y Klein (1980) observaron que la ejecución en tareas de reconocimiento era peor si se presentaba simultáneamente dos estímulos diana espaciales que cuando uno de ellos era espacial y el otro verbal.

Finalmente, se han obtenido conclusiones sobre los recursos definidos hemisféricamente a partir de tareas de tiempo de reacción. En este caso, se obtienen latencias de respuesta más largas cuando el hemisferio de procesamiento del estímulo es el mismo que el que controla la respuesta (Allwitt, 1981; Dimond y Beaumont, 1972; Wickens y Sandry, 1982).

El supuesto de que el procesamiento central espacial y verbal requieren recursos distintos permite explicar los resultados obtenidos por Allport y colaboradores (1972), quienes observaron una ejecución perfecta en dos tareas de procesamiento de información realizadas a la vez y que envolvían todos los estados de procesamiento (lectura muy rápida de música y enmascaramiento verbal).

2.4.2.3. Modalidades de input y de respuesta

Parece evidente que en muchas ocasiones se puede dividir la atención entre ojo y oído mejor que entre los dos ojos o los dos oídos. Esta observación es claramente cierta si se acude al concepto de interferencia periférica, la cual domina en condiciones intramodales.

Isreal (1980), Martin (1980), Rollins y Hendricks (1980) y Treisman y Davies (1973), han sugerido que existen ventajas en la presentación entremodal. Treisman y Davies observaron una detección más eficiente de patrones temporales-espaciales y de dianas semánticas en una presentación entremodal que en una intramodal. Rollins y Hendricks (1980) y Martin (1980) replicaron este resultado cuando la profundidad del procesamiento semántico del estímulo auditivo se variaba sistemáticamente. Isreal replicó el mayor efecto de la interferencia entre versus la intramodal entre una tarea

de seguimiento y otra de tiempo de reacción cuando la modalidad de ambas tareas era manipulada ortogonalmente utilizando estímulos de tiempo de reacción visual o auditivos emparejados con un display de seguimiento visual o auditivo.

Considerando las modalidades de respuesta, los trabajos de Harris y col. (1978), McLeod (1977), Vidulich (1988), Vidulich y Wickens (1981), Wickens (1980, 1988) y Wickens y col. (1983) han mostrado la mayor eficacia de la realización conjunta de la tarea de seguimiento con tareas discretas que utilizan respuestas orales en oposición a las manuales.

Por otro lado, el esquema propuesto por Wickens tiene en cuenta los factores que influyen en el patrón de interferencia entre dos tareas concurrentes y que se relacionan con la competición en el acceso al mecanismo central. Sin embargo, es muy poco explicativo de la forma en la que los elementos estructurales y energéticos se relacionan unos con otros. Este aspecto se encuentra más desarrollado en el modelo de estado energético-cognitivo propuesto por Sanders (1983) y Gopher y Sanders (1984). Este modelo intenta integrar los conceptos energéticos con una descripción estructural del sistema. Su desarrollo derivó de las investigaciones acerca de los efectos de estresores (falta de sueño, tiempo de tarea, drogas) sobre la ejecución en tareas de reacción de elección (Frowein, 1981; Sanders, Wijmen y Van Arkel, 1982). Estas investigaciones mostraron el efecto selectivo de dichos estresores, esto es, afectan a mecanismos específicos; su efecto sobre la ejecución no es general. Estos resultados fueron interpretados dentro del esquema del modelo neurofisiológico del control de atención propuesto por Pribram y McGinness (1975), quienes identificaron tres generadores energéticos globales de la actividad de procesamiento: arousal, activación y esfuerzo. Según este modelo, el arousal se relaciona con la codificación del input, la activación con los procesos de respuesta y el esfuerzo activa los mecanismos centrales de toma de decisiones y elección. Además, se asume que el esfuerzo actúa como un regulador del nivel óptimo de operación de las otras dos fuentes energéticas y como coordinador de sus actividades. El esfuerzo depende de la

evaluación que el sujeto realiza de la situación y de su motivación. Es un modelo similar al anteriormente expuesto de Kahneman (1973) aunque, en el que ahora nos ocupa, los mecanismos de evaluación tienen un enlace independiente con cada uno de los tres mecanismos energéticos y por tanto tienen una influencia separada sobre diferentes aspectos de la ejecución.

2.4.3. Metodología de análisis de la carga propia de los modelos del recurso

La teoría del recurso utiliza como elementos gráficos básicos de análisis de la carga la función ejecución-recursos (*Performance-Resources Function*) y la *Performance Operating Characteristic* (POC). A continuación se describe brevemente cada una de ellas:

2.4.3.1. La función ejecución-recursos (PRF)

La calidad de la ejecución es una función monótona no decreciente de los recursos invertidos en una tarea. Esta proposición se pone de manifiesto de dos formas. Bajo condiciones de tarea simple, si un sujeto invierte un gran esfuerzo en la tarea, la ejecución mejorará o al menos no decrecerá. Aunque esta proposición es intuitiva, ha recibido poca confirmación experimental directa. Bajo condiciones de tarea dual, la relación entre ejecución y recursos es más fácil de medir, pero requiere mayores supuestos sobre los procesos que subyacen. Cuando un sujeto realiza dos tareas a la vez, y se le pide que focalice su atención principalmente sobre una de ellas, la ejecución varía consistentemente de acuerdo a estas instrucciones (Gopher, 1980; Navon y Gopher, 1980; Friedman, Polson, Gaskill y Dafoe, 1982; Vidulich y Wickens, 1981). Bajo estas circunstancias, la teoría del recurso infiere que el sujeto está modulando la provisión de recursos entre las tareas con el fin de obtener los niveles deseados de ejecución.

Norman y Bobrow (1975) introducen el importante concepto de **función ejecución-recursos (PRF)**. Esta hipotética función relaciona la calidad de la ejecución y la cantidad de recursos invertidos en la tarea.

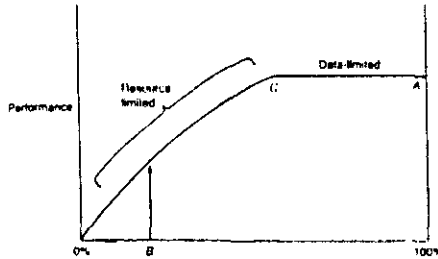


Figura 2.5. Una hipotética función ejecución-recursos. Los puntos A, B, y C se describen en el texto a continuación. Tomado de Wickens (1984).

La ejecución perfecta en la tarea simple tiene lugar cuando se invierten en ella todos los recursos (punto A) y es la mejor ejecución que se puede obtener. A medida que se traspasan recursos de esta tarea a otra concurrente la ejecución decrece (punto B). Cuantos más recursos son reinvertidos en la tarea, la ejecución mejorará hasta alcanzar el punto C, en el cual no es posible un cambio posterior en la ejecución. En la región a la derecha de este punto (C) se dice que la tarea es **limitada de datos** (limitada por la calidad de los datos, no por los recursos puestos en juego). Cuando la ejecución cambia con modificaciones en el número de recursos utilizados, se dice que la tarea es **limitada de recursos** (región a la izquierda del punto C).

2.4.3.2. The Performance Operating Characteristic (POC)

Cuando dos tareas son realizadas a tiempo compartido, y los recursos son asignados diferencialmente entre ellas, la ejecución conjunta de ambas puede ser representada en dos separadas PRF. Además, estos datos pueden ser captados representando un solo punto para cada condición en una *Performance Operating Characteristic* (POC), en la cual la ejecución para cada tarea está representada sobre los dos ejes. Las diferentes formas que puede asumir una POC reflejan las diferencias en la naturaleza de los intercambios entre la ejecución de las dos tareas. Si se aumenta la dificultad de una de dos tareas concurrentes incrementando las demandas de un recurso necesario para la ejecución de ambas tareas, la forma de la POC cambiará, ya que un mayor sacrificio de ejecución será requerido sobre la otra tarea para hacer posible una mejora sobre la tarea ahora más difícil. Si, por otro lado, el incremento en la dificultad se realiza aumentando las demandas en una dimensión que no es relevante para la ejecución conjunta, la ejecución de la tarea en la que se hizo la manipulación puede estar afectada, pero la forma de la POC no debería cambiar.

La idea de la construcción de POC's como una aproximación general al análisis del comportamiento del sistema de procesamiento humano fue introducida por Norman y Bobrow (1975), con referencia a la metáfora del computador, y posteriormente elaborada por Navon y Gopher (1979, 1980), quienes adoptaron conceptos de la teoría microeconómica. Sperling y Doshier (1986) usan el término *Attention Operating Characteristics* (AOC) para dar nombre a estas curvas y discutirlos en el esquema general de la teoría de la detección de señales. Sin tener en cuenta la metáfora de fondo, la técnica empleada por todos los investigadores es la misma: los sujetos son instruidos a realizar dos tareas, con cambios graduales en las prioridades relativas de tareas bajo condiciones de tarea dual (los diferentes niveles de énfasis son indicados verbalmente, y frecuentemente aumentados con información de feedback y atención diferencial).

Las POC tienen cinco características particulares:

1. Ejecución en la tarea simple: Esta característica es mostrada por el punto de intersección de la POC con los dos ejes. Cuando la ejecución para cada tarea simple es superior a la compartida hay, en palabras de Navon y Gopher (1979), un **costo de concurrencia**. Esto puede ser explicado por las demandas de recursos de un "ejecutor a tiempo compartido", que es utilizado únicamente en situaciones de tarea dual. Por tanto, las demandas de este ejecutor y en consecuencia sus efectos sobre la ejecución no se ponen de manifiesto en las situaciones de tarea simple.

Una fuente alternativa del costo de concurrencia resulta de la interferencia periférica. Este término, **interferencia periférica**, se refiere a las situaciones en las cuales la ejecución de la tarea dual se deteriora como consecuencia de las restricciones físicas o fisiológicas del sistema de procesamiento. Interferencia periférica es distinto al concepto de **interferencia estructural**, aquella que hace difícil la ejecución a la vez de dos actos motores independientes. La **interferencia periférica** puede resolverse con la práctica. La interferencia estructural es compatible con los conceptos de la teoría del recurso mientras que la periférica no lo es.

2. Eficiencia a tiempo compartido: Esta característica de las dos tareas está indicada por la distancia media de la curva al origen.

3. Grado de intercambio lineal: Esta relación entre las dos tareas indica la extensión a la que los recursos están divididos o son intercambiables entre las tareas.

4. Sesgo de asignación: Manteniendo en mente que la POC es una serie de puntos, cada uno relativo a un ensayo distinto de la tarea dual, el sesgo de asignación está indicado por la proximidad de un punto dado de la POC a un eje sobre el otro. Por tanto, los puntos situados en la diagonal positiva reflejarán una asignación del recurso del 50% entre ambas tareas, y las PRF de cada tarea serán equivalentes.

5. Combinación de eficiencia y asignación: Existe un grave problema teórico cuando se intenta comparar puntos de la POC que difieren en eficiencia y asignación de recursos y que no están cercanos a la diagonal positiva. Para salvar esta dificultad se han intentado encontrar equivalencias con la teoría de la detección de señales. Sin embargo, se encuentran dos problemas en particular:

1) A diferencia de la ratio de falsas alarmas y aciertos en la teoría de detección de señales, las medidas de ejecución de las dos tareas pueden estar expresadas en unidades totalmente distintas; 2) A menos que se establezca algún supuesto sobre la forma invariante de todas las PRF, no hay una escala que pueda adaptar un cambio de unidad de ejecución a un cambio de unidad de asignación del recurso.

En relación al primer punto, Wickens y sus colaboradores (Wickens, 1980; Wickens, Mountford y Schreiner, 1981; Wickens y col., 1983) han asumido que los cambios iguales en la asignación del recurso entre diferentes tareas son equivalentes a unidades iguales de la variabilidad dentro de dichas tareas. Esta asunción, análoga a la realizada por Fechner (1860) en escalamiento psicofísico, hace que los ejes de la POC sean normalizados por una evaluación de la variabilidad intra sujeto.

El segundo problema, relacionado con las diferencias entre niveles de ejecución es también complejo. En este caso las presunciones sobre la forma de la PRF tienen algunas implicaciones estadísticas. Si se asume una función logarítmica, variaciones iguales en la asignación de recursos producirán menores variaciones en ejecución en los niveles altos de la misma que en los bajos. Por otro lado, si se asume una PRF lineal las diferencias estadísticas serán fuertemente equivalentes para todos los niveles.

La metodología POC supone una serie de supuestos implícitos. El primero y principal es la sensibilidad de la ejecución a los esfuerzos de procesamiento. Se asume que la ejecución en una tarea está monótonamente relacionada con el conjunto de recursos invertidos. Aun más, se asume que este conjunto es la fuente general de variabilidad en la ejecución a tiempo compartido. Otro supuesto es que el humano es

capaz de controlar los recursos y su aportación, al menos a lo largo de la porción general del rango sensible a la respuesta. Un tercer importante supuesto es que la POC refleja la limitación de la ejecución del sistema a capacidad total, y que esta capacidad es fija. Si los sujetos no operan al máximo de capacidad o si la capacidad puede aumentarse (Kahneman, 1973), la interpretación de una POC es imposible.

Dos supuestos adicionales son la independencia de las tareas y la invarianza de los procesos. Una manipulación de las prioridades relativas bajo condiciones de tarea dual no tiene sentido a menos que las tareas componentes mantengan su independencia de recursos. Así mismo, se asume que cuando se alteran los niveles de énfasis, las tareas no cambian su estructura de procesamiento básica. Solamente se asume que puede variar el nivel al cual estas estructuras están en juego.

La adopción del método POC conlleva una consideración sistemática de los efectos de la política de asignación. La metodología POC ha generalizado considerablemente el uso de patrones de interferencia entre tareas realizadas concurrentemente, como método para el estudio de la carga. Cuando la carga se define en términos de la interacción entre operadores y tareas, y debido a que las curvas POC retratan estas interacciones, su valor dentro de este esquema es evidente.

El inconveniente general de esta aproximación es que el diseño de experimentos es considerablemente más complejo. También consume más tiempo en la recogida y análisis de los datos. Otro inconveniente, más importante, resulta de la posible interferencia de la tarea secundaria con la primaria. Por tanto, hay una necesidad de encontrar tareas secundarias que no interfieran o que interfieran lo menos posible con la tarea primaria. Una categoría de estas tareas ha sido propuesta y desarrollada por Donchin y sus colegas (Donchin, 1975; Donchin, Kramer y Wickens, 1982; Donchin, 1984) usando los potenciales evocados como fuente de datos a partir de la cual inferir las variaciones sobre la carga de la tarea primaria.

Una importante característica de la teoría del recurso es su habilidad para tratar los efectos de la práctica y la dificultad de la tarea como manifestaciones diferentes del mismo constructo: la eficacia marginal de la inversión del recurso o ganancia de ejecución conseguida por unidad de recursos invertidos (Navon y Gopher, 1979). En términos de la representación POC, la versión más fácil o más practicada de una tarea genera una curva más alejada del origen que la POC de otra más difícil.

Para evitar confusiones en la interpretación de la POC cuando se utiliza la tarea secundaria como método de investigación, es importante representar los resultados de la tarea dual en un espacio de POC (siempre que se utilice una sola política de asignación), mas que recoger únicamente los decrementos de la ejecución de la tarea secundaria. De esta manera, el investigador puede representar más adecuadamente cómo los sujetos están asignando recursos según las diferentes condiciones.

Al tratar estos aspectos surge de nuevo la analogía con la teoría de la detección de señales, en cuanto a sus índices de eficacia (d') y de sesgo cognitivo (β). En la realización de la tarea dual, los decrementos en la tarea secundaria y en la primaria son recogidos e interpretados análogamente en términos de un índice de eficacia (la distancia desde el origen de la POC) y de un sesgo de asignación (la distancia desde la diagonal positiva).

2.4.4. Recurso único versus múltiple: la influencia de las prioridades

Un examen cuidadoso revela que realmente no existen grandes diferencias entre el modelo de recursos múltiples y el de Kahneman, el cual asume un recurso indiferenciado con competencia para estructuras satélite. Ambos predicen que la eficacia de las tareas duales será menor si ambas comparten estructuras comunes. Al igual que la teoría del recurso múltiple, el modelo de Kahneman puede también explicar la insensibilidad a la dificultad invocando al concepto de "límites de datos". Sin embargo, el modelo de recurso indiferenciado no puede explicar fácilmente la ejecución perfecta en tareas duales. Es posible, en este modelo, asumir que dos tareas

puedan ser realizadas a la vez de forma eficaz, pero no perfectamente, si sus estructuras de codificación y de respuesta son diferentes. Pero si ambas tareas demandan algún grado de procesamiento central de un único recurso (toma de decisiones, memoria u operaciones de traducción) habrá interferencia a menos que las tareas sean fuertemente limitadas de datos. Si este no es el caso, para poder explicar la ejecución perfecta en tareas duales es necesario asumir la existencia de recursos separados o diferentes a nivel central.

Según Broadbent (1982) la tarea de lectura muy rápida de Allport y col. (1972) permite el uso de una secuencia de estímulos previa, de forma que estrategias eficaces de muestreo de los estímulos empleadas por los sujetos, junto con medidas de la ejecución insuficientemente sensibles, oscurecen los decrementos en ejecución de la tarea dual. Aunque estos argumentos pueden ser válidos, lo importante no es tanto si la ejecución es totalmente perfecta o no, sino que el decremento sea tan pequeño como lo es.

Quizás la diferencia más clara entre estos dos modelos es que mientras uno postula la existencia de un único recurso, el otro postula la de varios. Para establecer este último supuesto empíricamente se requieren una de las dos técnicas experimentales siguientes: A- aquella que relaciona los cambios en la POC (*Performance Operating Characteristic*) con cambios en el solapamiento de recursos, y B- el efecto de las manipulaciones de la dificultad sobre la POC. A continuación se describe brevemente cada una de ellas.

A.- POC y cambios en el solapamiento de recursos

Es necesario no sólo identificar una POC suave entre dos tareas, sino también mostrar que la POC es discontinua cuando los recursos no se comparten.

Vidulich y Wickens (1981) compararon la asignación de recursos intramodal con la entremodal, en los estados de codificación y de respuesta. Sus sujetos realizaron la tarea de búsqueda en la memoria de Sternberg (1969) a la vez que otra de

seguimiento. La primera de estas tareas fue presentada en sus versiones auditiva y visual y los sujetos podían responder de forma oral o manual. Además los participantes recibieron dos instrucciones diferentes de asignación de prioridades a una u otra tarea en cada condición.

Como predicen ambos modelos, la eficacia de la tarea dual fue una función inversa del total de recursos solapados entre la tarea de seguimiento visual-manual y las modalidades de input-output de la tarea de búsqueda en la memoria. Sin embargo, aparecieron dos resultados directamente compatibles con la teoría de recursos múltiples: a) en la condición auditiva-oral la ejecución fue prácticamente perfecta; b) el efecto de la manipulación de la asignación de prioridades fue una función monótona creciente del nivel de solapamiento entre modalidades de input-output de ambas tareas. Estos resultados confirman que cuantos más recursos compartan las tareas habrá un mayor grado de capacidad asignable intercambiable, definida operacionalmente como recursos que pueden modular la ejecución según las prioridades de la tarea.

Conclusiones similares fueron encontradas por Friedman y col. (1982). Ellos utilizaron una tarea dual formada de una tarea de memoria y otra de detección, y diferentes asignaciones de prioridades entre ambas tareas según el hemisferio cerebral utilizado. Ellos observaron que cuando se utilizaba el mismo hemisferio para las dos tareas se producía una gran interferencia, y que la manipulación de prioridades tenía un gran efecto sobre la ejecución. El uso de hemisferios diferentes mejoró en gran medida la ejecución y atenuó el efecto de la manipulación de prioridades.

B.- Efectos de la manipulación de la dificultad sobre la POC

Para Navon y Gopher (1979) la interacción entre la dificultad de la tarea y las prioridades proporciona apoyo a la teoría de múltiples recursos. Según esta interacción, las prioridades ejercen una influencia más fuerte sobre la ejecución cuando la demanda de recursos comunes es alta que cuando es baja.

Gopher, Brickner y Navon (1982) utilizaron tres niveles diferentes de énfasis en la prioridad de una tarea de seguimiento y otra de clasificación de dígitos. En esta última la dificultad se modificó en dos dimensiones; una variando sus demandas cognitivo-verbales y otra modificando la complejidad de su respuesta motora. El efecto de las prioridades sólo se puso de manifiesto en la segunda dimensión. Los autores explicaron estos resultados asumiendo que el incremento en la demanda para las respuestas espaciales-motoras asociado con el incremento en la complejidad de la respuesta creó una mayor demanda de las respuestas espaciales y motoras de la tarea de seguimiento, y que el incremento en las demandas de recursos cognitivo-verbales de la tarea de dígitos no impuso demandas sobre los recursos del seguimiento.

En el estudio de Vidulich y Wickens (1981) los sujetos realizaron la tarea dual en dos diferentes niveles de dificultad de seguimiento, determinados por el orden de los controles. Los autores encontraron que el coste de las respuestas manuales de la tarea de memoria fueron mayores en el segundo orden que en el primero, lo que sugiere que el locus de control de orden más alto estuvo en los recursos relacionados con la respuesta. Este resultado permite predecir que las prioridades interactúan con el orden de control cuando hay competencia entre recursos comunes de respuesta (cuando la tarea de memoria requiere respuesta manual) en mayor medida que cuando no la hay.

2.4.5. Recursos versus procesador central

Long (1976) y Tulving y Lindsay (1967) han concluido que en tareas de detección y de reconocimiento el procesamiento es compartido simultáneamente entre las señales auditivas y las visuales más que discretamente distribuido. Esta demostración de "capacidad compartida" está relacionada con la cuestión del procesamiento paralelo versus serial (Taylor, 1976; Townsend, 1974) y establece un punto de convergencia entre el modelo de un procesador central de capacidad limitada (LCCP) y la teoría del recurso (único o múltiple). Claramente, el procesador del modelo de filtro o el del de LCCP puede tener las cualidades de un recurso si el

procesador a) puede distribuirse con suficiente rapidez entre tareas o canales de información, y b) es capaz de ajustar el tiempo de actuación proporcionalmente según las estrategias del operador y las prioridades de la tarea.

En muy bajas frecuencias de muestreo (como las de las estrategias de fijación visual) la última es una estrategia claramente disponible y por tanto, puede ser fácilmente validada por medidas objetivas (Senders, 1964). Si, por el contrario, se postula una distribución de mayor frecuencia, parece imposible distinguir si los recursos de procesamiento o las estructuras son verdaderamente compartidos por las tareas o son modulados por la rápida distribución. Realmente, parece que en algunos niveles de procesamiento, la distribución discreta de la atención es claramente un fenómeno identificable (LaBerge, Van Gilder y Yellott, 1971).

2.4.6. Estructura jerárquica de los recursos

Las implicaciones de la representación tridimensional de la estructura de los recursos son: 1) las tareas que completamente no demandan un solapamiento de ningún recurso serán realizadas perfectamente en condiciones duales, y 2) si dos tareas utilizan parcialmente recursos diferentes su grado de interferencia se verá afectado o no según sea la distancia funcional entre el componente de no solapamiento de estos recursos.

Varias investigaciones (Treisman y Davies, 1973) han mostrado que dos tareas dentro de una única celdilla de la representación (por ejemplo, dos tareas auditivas-verbales) interferirán en mayor grado que dos tareas de celdillas adyacentes (por ejemplo, una tarea auditiva-verbal y otra visual-verbal o una auditivo-verbal y otra auditivo-espacial). Sin embargo, los datos no muestran que dos tareas que demandan celdillas adyacentes serán realizadas al mismo tiempo perfectamente. Estas consideraciones sugieren que los recursos de procesamiento humano pueden ser definidos jerárquicamente. Un ejemplo de este esquema propone la existencia de recursos auditivos y visuales en algún grado diferentes, cada uno exclusivo de su

específica modalidad. Estos no pueden ser transferidos a otra modalidad para facilitar la ejecución, pero pueden ser compartidos por dos tareas de la misma modalidad. Además, existen recursos perceptivo-verbales y perceptivos-espaciales más generales, cada uno de los cuales puede ser compartido entre modalidades pero no entre códigos. Por encima de este nivel en la jerarquía, existe un conjunto de recursos de procesamiento general perceptivo-central disponible tanto para el procesamiento espacial como el verbal de la información auditiva y de la visual, pero que no está disponible para los procesos de respuesta. Finalmente, en un nivel más general existiría un conjunto de recursos indiferenciados disponible para y demandado por todas las tareas, en las modalidades, códigos y estados que se requieran. Se podría asumir que estos recursos más generales representan lo convencionalmente denominado como atención, conciencia, filtro o LCCP. Sin embargo, la existencia de recursos generales no elimina el valor explicativo del concepto de recursos múltiples.

Según Wickens (1986) esta representación jerárquica, aunque permite explicar el incremento de la interferencia como una función del incremento en la proximidad de las tareas en el espacio total de recurso, no es enteramente adecuada. El problema es que esta jerarquía explícitamente propone un orden de dominancia de dimensiones que sitúa a las modalidades por debajo de los códigos y a estos bajo los estados. Según esta representación el efecto de una alteración estructural no será observado dentro del nivel de una estructura compartida superior en la jerarquía. Más específicamente, este esquema predice que el efecto de las modalidades separadas versus compartidas en condiciones de tarea dual sólo será observado si ambas tareas comparten un código común de procesamiento (por ejemplo, ambas son espaciales). Los datos recogidos por Vidulich y Wickens (1981) sugieren claramente que esto no es así. Igualmente, el modelo jerárquico descrito predice que el efecto de códigos separados versus compartidos sólo será observado si se emplea un estado de procesamiento común. La demostración de Brooks (1968) de la interacción entre tareas de memoria inmediata espacial y verbal y modalidades de respuesta, da evidencia en contra de esta interpretación jerárquica. De este modo, aunque puede de hecho existir algún grado de dominancia de orden entre dimensiones, no parece que esta sea unidireccional. Esto es, es probable que códigos diferentes mejoren la eficacia en

tareas duales sobre los compartidos siempre que también se utilicen estados diferentes. Además es probable que algún aspecto de un modelo de "características compartidas" deba ser empleado para predecir la interferencia; siendo una característica definida como un nivel a lo largo de cada una de las dimensiones dicotómicas del espacio de recurso. Así, Derrick (1981) encontró una correlación alta ($r=0.76$) entre la interferencia pronosticada según un modelo de características compartidas, y la interferencia observada a través de 10 condiciones creadas por 4 tareas compartiendo todas las combinaciones posibles en las distintas dimensiones.

2.4.7. Estrategias y carga

Las estrategias que un operador puede seleccionar cuando realiza una tarea son relevantes al tema de la carga ya que el operador puede elegir realizar una tarea de forma cualitativamente distinta a medida que aumenta la dificultad. Este cambio puede realizarse de forma que se mantengan ciertos aspectos de la ejecución mientras que se sacrifican otros (Welford, 1978). La habilidad del operador para utilizar estrategias de este tipo es una consecuencia natural del hecho de que la ejecución puede variar a lo largo de diferentes dimensiones (aciertos y falsas alarmas, rapidez y precisión, asignación de recursos a la tarea A y a la B). Por tanto, el operador es capaz de alterar la ejecución de varias formas según la situación (demandas de recursos) lo requiera. En consecuencia, para asegurar un conocimiento adecuado de los efectos de las variables de carga, todos los aspectos de la ejecución deben ser medidos cuidadosamente.

Un ejemplo hipotético es aquel en el que el operador elige sacrificar la latencia de sus respuestas, a medida que las demandas de recursos incrementan en los niveles bajos, y la precisión cuando las demandas son impuestas en niveles altos. Las estrategias de búsqueda de información del operador y la focalización de la atención se alteran en condiciones de alta carga, focalizándose relativamente más en menos canales de información más alta. La distinción de Rasmussen (1981) entre estrategias

de solución de problemas sintomáticas y topográficas, proporciona un ejemplo más de como las estrategias cambian del primer tipo al segundo cuando la carga incrementa. Como ejemplo final, Sperandio (1978) investigó la carga experimentada por controladores aéreos a medida que aumentaba el número de aviones en su pantalla. El encontró que los controladores empleaban diferentes estrategias de agrupación de los aviones por sus atributos cuando la carga aumentaba de niveles bajos (menos de 6 aviones) a moderados (de 6 a 10) o a altos (más de 10 aviones).

La discusión previa enfatiza las dificultades que la flexibilidad de la conducta humana impone a la precisión de la medida. Otra variable interviniente debe ser considerada frecuentemente entre las demandas de la tarea y la ejecución. En un sistema inflexible, no estratégico, el incremento en las demandas de la tarea invariablemente sacrificará a la ejecución o a la capacidad residual de forma constante. En un sistema que se adapta estratégicamente, como es el operador humano, el incremento en las demandas puede alterar (quizás de forma más eficiente) las estrategias. Estas estrategias consumen menos recursos para obtener el mismo nivel de ejecución en la dimensión de mayor importancia o de mayor utilidad para el operador (mientras sacrifica la ejecución en otras dimensiones). Los modelos de carga deben tomar en cuenta esta flexibilidad.

2.4.7.1. Relaciones entre recursos y estrategias

La relación entre los recursos y las estrategias adoptadas por los sujetos al llevar a cabo tareas duales puede ser vista en dos niveles: intra e inter tareas.

A un nivel intratarea, se pueden emplear diferentes estrategias de ejecución que pueden incrementar o disminuir las demandas de recursos de las tareas componentes. Por ejemplo, la codificación y la lectura en voz alta de material verbal pueden diferir en cuanto a la profundidad de procesamiento (Craik y Lockhart, 1972), y esto presumiblemente altera el énfasis sobre el código fonético frente al semántico (Posner, 1978). Esto modificaría la interferencia relativa entre tareas que difieren en

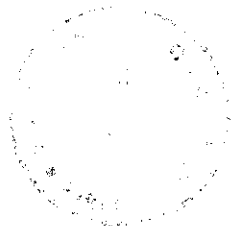
su dependencia de recursos verbales frente a auditivos (Martín, 1980). Con tareas de seguimiento, Wickens, Gill, Kramer, Ross y Donchin (1981) han comprobado como el uso de estrategias puede modificar los niveles de error.

A un nivel entre tareas, se pueden emplear las estrategias para adoptar distintas políticas de asignación de los recursos entre las tareas. Por ejemplo, Schneider y Fisk (1982), Schneider y Shiffrin (1977) y Schneider y col. (1985) han demostrado que cuando los sujetos son instruidos para enfatizar la tarea limitada de recursos de una tarea dual, la eficacia se aproxima al nivel máximo.

Otra demostración de la importancia de la estrategia de asignación en tareas duales fue establecida por el estudio de Brickner recogido en Gopher (1980). En esta investigación, dos grupos de sujetos practicaron una tarea dual bajo dos condiciones: asignación de prioridades fija o variable. Se observó que la ejecución de los sujetos entrenados fue mejor. Posiblemente, las habilidades en asignación diferencial de recursos que los sujetos adquirieron durante el entrenamiento les fueron muy útiles para ajustar de forma óptima los recursos proporcionados a tareas de demandas variables.

Tsang (1983) demostró la importancia de entrenar a los sujetos para asignar recursos apropiadamente entre dos tareas en un ambiente dinámico, en el cual la dificultad (demandas) de la tarea primaria está continuamente variando. En circunstancias normales, los sujetos manejan la dificultad manteniendo la ejecución de la tarea primaria a niveles óptimos por medio de la prestación y devolución de recursos de la tarea secundaria a la primaria según se necesiten (Wickens, Tsang y Benel, 1979). Sin embargo, cuando el sujeto se aúna explícitamente a las estrategias de asignación óptimas, la ejecución es aún mejor.

**CAPITULO 3. LA MEDIDA DE LA CARGA: PREDICCIÓN Y
TECNICAS DE EVALUACION.**



3.1. INTRODUCCION

Las diferentes técnicas de evaluación de la carga propuestas varían en un número de dimensiones que pueden servir como criterios para la selección de un procedimiento concreto, adecuado para una determinada aplicación. Un procedimiento puede ser adecuado para un cierto propósito, o para una determinada aplicación, pero no serlo para otro. Varios autores (Chiles, 1982; Eggemeier, 1984; Gartner y Murphy, 1976; Sheridan y Stassen 1979; Shingledecker, 1983, 1987; Wickens, 1984b; Wierwille y Williges, 1978) han discutido sobre la mayoría de las dimensiones de las técnicas de evaluación y predicción de la carga que pueden afectar a su aplicabilidad. A continuación, se describen estos criterios.

3.1.1. Criterios para la selección de los índices de carga

Se ha propuesto un número de criterios que idealmente deberían ser alcanzados por cualquier técnica de evaluación de la carga. Desde luego, es cierto que algunos de estos criterios pueden intercambiarse unos por otros (Wickens, 1979, 1981), y, de este modo, raramente encontraremos una técnica que los posea todos. Estos criterios son los siguientes:

1. **Sensibilidad:** Se refiere a la capacidad de una técnica para detectar cambios en los niveles de dificultad de la tarea o en sus demandas.
2. **Poder de diagnóstico:** Un índice debería no sólo identificar cuando la carga varía, sino también la causa de esta variación. En términos de la teoría de recursos múltiples, debería indicar cual de las capacidades o recursos varían por los cambios en las demandas del sistema. Esta información hace posible la implementación de mejores soluciones.

3. **Selectividad:** El índice debería ser selectivamente sensible sólo a las diferencias en las demandas de capacidad y no reflejar los cambios en factores como la carga física o el estrés emocional que pueden no estar relacionados con la carga mental o capacidad de procesamiento de la información.
4. **Intrusión de la tarea primaria:** El índice no debería interferir con, contaminar o perturbar la ejecución de la tarea primaria cuya carga esta siendo evaluada.
5. **Fiabilidad:** Como con cualquier medida de la conducta, un índice de carga debería ser fiable. Sin embargo, si la carga es evaluada en diversos momentos (p.e., si es necesario estudiar los cambios de carga que se producen durante el curso de una tarea u operación), es importante que el índice ofrezca una estimación fiable de la carga lo suficientemente rápida de manera que los cambios transitorios puedan ser evaluados.
6. **Requisitos de implementación:** Incluye aspectos como la instrumentación y *software* necesarios para la recogida y análisis de los datos, y el grado de entrenamiento del operador requerido para poder obtener resultados válidos.
7. **Aceptación por el operador:** Hace referencia a la percepción que tienen los sujetos de la validez y utilidad del procedimiento.

Se ha propuesto una variedad de técnicas de predicción y evaluación de la carga, algunas tienen muchos de los criterios descritos anteriormente, pero pocas los satisfacen todos. Por ejemplo, Wierwille y Williges (1978) identificaron 28 técnicas específicas utilizadas para medir la carga. A pesar de esta diversidad, la mayoría de los procedimientos de evaluación empíricos pueden ser clasificados dentro de cinco categorías generales. La primera es la técnica de predicción de la carga basada en el análisis de línea de tiempo. Las restantes cuatro categorías de evaluación (Wierwille y Williges, 1978) son las medidas de tarea primaria, las de tarea secundaria, las fisiológicas y las técnicas subjetivas.

3.2. ANALISIS DE LINEA TEMPORAL

Esta técnica está diseñada principalmente para la predicción de la carga, y proporciona al diseñador de sistemas un perfil de la carga que soportan los operadores durante una tarea u operación típica (p.e. aterrizaje de un avión, puesta en marcha de una planta generadora de energía, etc.). En una simplificada pero muy útil versión, asume que la carga es proporcional a la razón entre el tiempo ocupado realizando la tarea y el tiempo disponible para llevarla a cabo. Si el operador está ocupado con una tarea el 100% del tiempo, entonces la carga es del 100% durante ese intervalo. Así, por ejemplo, la carga de una actividad podría ser computada dibujando líneas que representen cada operación o subtarea, de longitudes proporcionales a su duración. Se sumaría la longitud total de las líneas, y la cantidad resultante se dividiría por el tiempo total (Parks, 1979). De esta forma, se puede comparar la carga obtenida para diferentes miembros del equipo (p.e., piloto, copiloto, e ingeniero de vuelo), y reasignar las tareas entre los miembros si es que existiera un gran desequilibrio. Aún más, permite identificar y representar momentos de carga punta o de sobrecarga, en los que la carga es mayor del 100%.

3.2.1. Limitaciones

El análisis de línea de tiempo, aunque es una valorable herramienta predictiva, tiene algunas importantes limitaciones. Normalmente asume un modelo simplificado de "canal único" de la atención humana, en el cual cualquier actividad observable demanda una completa atención (independientemente de sus demandas de recursos o de los recursos particulares de dicha tarea), mientras que a las actividades no observables (toma de decisiones, solución de problemas) se les asigna el valor cero, ya que no están asociadas con actos motores. Según esta técnica, cuando se realizan dos tareas a tiempo compartido se trabaja a un 200% de carga. Sin embargo, en ciertos casos se pueden realizar dos tareas concurrentemente de forma muy eficaz si ambas demandan pocos recursos o si demandan recursos de diferente naturaleza.

Otra limitación de esta técnica es que es aplicable en un esquema relativamente "forzado" en el cual no se pueden intercambiar las actividades en ningún momento. Por otro lado, como Gotsdanker y Senders (1980) han puesto de manifiesto, en una variedad de situaciones (típicamente p.e., del transporte aéreo) el operador tiene un considerable grado de flexibilidad para realizar las tareas requeridas en momentos alternativos. El operador es por tanto capaz de minimizar las demandas sobrecargantes siguiendo un esquema eficaz. Cuanto más posible es el autoesquema (más flexibilidad tiene el operador para programar sus tareas), mayores dificultades se encuentran para aplicar la técnica de análisis de línea temporal, para predecir momentos de carga extrema o punta.

3.3. MEDIDAS DE TAREA PRIMARIA

Estas medidas describen la ejecución de la tarea en la que se está midiendo la carga. Normalmente, se espera que el incremento en la dificultad de la tarea incremente la demanda de recursos o carga. En un primer momento, los parámetros de dificultad de la tarea son expresados por medidas como el intervalo entre estímulos, el número de fuentes de información, la complejidad de control, o la compatibilidad, que deberían ofrecer buenas medidas predictivas de la carga de la tarea. Sin embargo, hay importantes inconvenientes en el uso de los parámetros de tarea primaria para predecir y evaluar carga.

3.3.1. Limitaciones

En primer lugar, frecuentemente es difícil cros-calibrar estas distintas medidas entre las tareas. Por ejemplo: ¿Cuántas teclas pueden ser añadidas a un teclado en una tarea de transcripción compleja a máquina?, o ¿Cuánta inestabilidad de control puede añadirse a una tarea de seguimiento para producir un incremento de carga que sea equivalente a incrementar la información transmitida de 2 a 3 bits en una tarea de tiempo de reacción de elección o a incrementar la carga sobre la memoria de trabajo de 5 a 7 chunks?.

En segundo lugar, es imposible decir que la dificultad de la tarea (manipulada por el experimentador) produce carga (experimentada por el sujeto) a menos que la ejecución de la tarea primaria sea controlada y medida cuidadosamente. En términos de una PRF el incremento en la dificultad de la tarea claramente incrementa la carga (recursos demandados), si en todos los casos el operador esta asignando suficientes recursos para obtener la ejecución máxima. A medida que aumenta la dificultad, el conjunto de recursos incrementa proporcionalmente. Sin embargo, si el operador adopta una política de asignación de recursos habrá progresivos decrementos en ejecución a medida que aumenta la dificultad. Por tanto, no parece apropiado concluir que la carga ha variado en este último caso, ya que la misma cantidad de recursos es invertida en todas las condiciones.

En resumen, se debe ser cuidadoso de asumir a priori que una manipulación de las demandas de la tarea ha incrementado la carga. Esta afirmación solamente puede hacerse si se sabe con certeza que el sujeto está poniendo en juego la cantidad suficiente de recursos para producir la mejor ejecución posible en cada condición.

Un medio para determinar el nivel de capacidad residual de una tarea particular realizada a un determinado nivel es la evaluación del margen de carga de la tarea primaria. Para obtener el margen de carga se siguen los siguientes pasos: 1) Se especifica un nivel del criterio al cual debe ser realizada la tarea. En contextos aplicados el criterio es normalmente establecido por el ingeniero de sistemas (p.e., la desviación máxima permitible de un patrón de vuelo próximo al aterrizaje de un avión, o la rapidez de escritura de un mecanógrafo). 2) Se elige un parámetro de tarea primaria que afecte a las demandas de recursos de un tipo particular. 3) Finalmente, este parámetro es manipulado hasta que se alcance el nivel de ejecución dado por el criterio. La magnitud de la manipulación de este parámetro es el margen de carga. Esto proporciona un índice de la cantidad de demandas adicionales, a partir de las condiciones iniciales, que puede darse en los recursos en cuestión antes de que la ejecución sea insatisfactoria.

En el contexto de la teoría de múltiples recursos, el margen de carga es un vector cuyas dimensiones reflejan los distintos tipos de recursos. Por ejemplo, una tarea que en gran medida requiere recursos de procesamiento perceptivos/centrales (p.e., detección de señales o comprensión verbal) se verá más afectada por el incremento en sus demandas de codificación que por incrementos en la complejidad de la respuesta. Por otro lado, una tarea con mayores demandas motoras, como es la de seguimiento, será relativamente más sensible a un incremento en la dificultad de las características de su respuesta que a un decremento en la calidad del display.

3.4. LA TÉCNICA DE TAREA SECUNDARIA

La imposición de una tarea secundaria como una medida de los recursos residuales o capacidad no utilizada en la tarea primaria (Ogden, Levine y Rolfe, 1973) es una técnica que en algunos aspectos está muy relacionada con el margen de carga de la tarea primaria. *En el margen de carga, la capacidad de reserva es "absorbida"* por el incremento de la dificultad de la actividad original. En la técnica de tarea secundaria, los recursos residuales son absorbidos por la introducción de una nueva actividad, la tarea secundaria. De esta forma, la ejecución de la tarea secundaria es idealmente inversamente proporcional a las demandas de recursos de la tarea primaria. Así, las tareas secundarias pueden reflejar diferencias en las demandas de recursos, automatización, o práctica que no están reflejados en la ejecución de la tarea primaria. Como ejemplos del uso de esta técnica, Garvey y Taylor (1959) encontraron que las diferencias en controles de seguimiento eran reflejadas sólo con la adición de una medida de tarea secundaria. Bahrick y Shelly (1958) encontraron que la tarea secundaria era sensible a las diferencias en automatización. La ejecución de sus sujetos en una tarea de tiempo de reacción serial no difería entre una secuencia de estímulos predecible y otra aleatoria. Sin embargo, la ejecución de la tarea secundaria discriminó entre las dos: con la práctica la secuencia repetida requería menos recursos. Los artículos de Rolfe (1973) y Ogden y col. (1979) proporcionan otros numerosos ejemplos del uso de técnicas de tarea secundaria.

Una variante de la técnica de tarea secundaria es el uso de una tarea cargante (Ogden y col, 1979; Rolfe, 1973). Cuando se utiliza una técnica de tarea secundaria el investigador está interesado en la variación en el decremento de la tarea secundaria (a partir de una condición de control de tarea secundaria en solitario) para inferir diferencias en las demandas de recursos de la tarea primaria. Así, la tarea primaria es tanto la tarea de interés como la tarea cuya prioridad es enfatizada. En la técnica de tarea cargante, por otro lado, se proporcionan diferentes instrucciones de asignación (prioridades). El sujeto tiene que poner en juego todos los recursos necesarios en la tarea cargante, y el grado de intrusión de esta tarea sobre la ejecución de la tarea "primaria" es examinado para comparar diferencias entre tareas primarias. Dentro del contexto de la discusión previa sobre la teoría del recurso, es claro que las técnicas de tarea secundaria y de tarea cargante solamente dirigen el foco de interés experimental a diferentes áreas de la POC. Cuando se emplea la técnica de la tarea secundaria, la tarea primaria es enfatizada y la varianza de la ejecución se pone de manifiesto en la tarea secundaria. Cuando se utiliza la técnica de tarea cargante la varianza de la ejecución se muestra en la tarea primaria. A partir de esta representación (POC) podemos ver que el método de tarea secundaria reflejará mayores diferencias que el de tarea cargante.

Una multitud de tareas secundarias han sido propuestas y empleadas muchas veces para evaluar la capacidad residual de las tareas primarias. A continuación se describen algunas de las más utilizadas (ver Ogden y col, 1979 y Wierwille y Williges, 1980 para una revisión más exhaustiva).

1- **Rhythmic-tapping** o **tarea de producción de intervalos**: Michon (1966) ha propuesto una tarea secundaria en la cual los sujetos tienen que mantener una velocidad constante de pulsaciones en un teclado. Michon asume que las demandas de la tarea primaria afectarán a la capacidad de selección de la respuesta para iniciar el movimiento (pulsación) en los intervalos de tiempo precisos. La carga de la tarea primaria se vera reflejada por el incremento en la variabilidad del intervalo entre pulsaciones. La principal limitación de esta técnica es su obstrusividad. Una mano del

operador debe estar continuamente disponible para pulsar las teclas y por tanto no puede ser empleada para realizar ninguna actividad de la tarea primaria.

2- Tarea de Posner y colaboradores ("*Probe reaction time*"): Posner y sus colegas han empleado esta técnica con considerable éxito al explorar las demandas de recursos asociadas con un número de tareas perceptivo-motora y cognitivas (Ells, 1973; Posner y Boies, 1971; Posner y Keele, 1969). Se presenta un estímulo impredecible y se requiere una rápida respuesta. A medida que la tarea primaria demanda más recursos, el procesamiento de la tarea secundaria es más lento. Desafortunadamente, debido a sus escasas demandas de procesamiento central, esta técnica parece ser en gran medida "perceptivo-motora" y por tanto particularmente sensible a la competición de modalidades de input y de respuesta. En consecuencia, puede proporcionar resultados diferentes en función de si se utilizan *probes* intra o entre modalidad. (McLeod, 1978; Wickens y col., 1983).

3- Tarea de búsqueda en la memoria de Sternberg: Al igual que en la anterior, se mide el tiempo de reacción a un estímulo inesperado, pero en este caso la respuesta es una decisión si-no, o un miembro de un anteriormente memorizado "*positive set*" de caracteres mantenidos en la memoria de trabajo. A medida que se incrementa el tamaño de este conjunto de caracteres (p.e., de 2 a 4 caracteres), el tiempo de reacción aumenta, pues un número mayor de items deben ser buscados en la memoria de trabajo. Debido a que el tiempo de reacción es normalmente representado gráficamente como una función de la carga de la memoria en el eje de abscisas, el incremento en el tiempo de reacción viene dado por la pendiente de la gráfica.

Cuando se mide la carga, el tiempo de reacción es medido bajo alta y baja carga de memoria, tanto para la tarea sola como para cuando es realizada simultáneamente con otra. La comparación de las dos funciones permite al experimentador hacer inferencias sobre la naturaleza de las demandas de recursos de la tarea primaria. Si la pendiente de la gráfica es más pronunciada en la condición dual que en la simple, entonces se concluye que la tarea primaria disminuye la razón

de búsqueda en la memoria. Debido a que la búsqueda en la memoria es una actividad de procesamiento central, esto indicaría que la tarea primaria tiene una carga de procesamiento central. Si sólo el "intercepto" es alcanzado por la tarea primaria pero la pendiente no cambia, entonces se asume que la fuente general de demandas de recursos de la tarea primaria es perceptiva-motora. Desde luego, puede haber casos en los que ambos, intercepto y pendiente, pueden verse afectados.

La técnica puede ser igualmente empleada con una versión fácil y difícil de la tarea primaria para evaluar la naturaleza de las demandas de recursos impuestas al incrementar la dificultad. Esta capacidad para discriminar los efectos de la carga proporciona a la técnica diagnosticidad, mientras que sacrifica su sensibilidad. La técnica ha sido bien validada en laboratorio y situaciones de simulación (Crosby y Parkinson, 1979; Schifflett, 1980; Spicuzza, Pincus y O'Donnell, 1974; Wickens, Hyman, Dellinger, Taylor y Meador, 1985).

4- **Estimación y producción de tiempos:** Intuitivamente, se hace evidente que el paso subjetivo del tiempo está de algún modo afectado por lo ocupado que esté el sujeto. Este fenómeno ha sido trasladado a dos diferentes técnicas de medida de carga: estimación y producción de tiempos. De hecho, dos modelos diferentes subyacen a cada técnica. En la tarea de producción el sujeto tiene que producir intervalos constantes de tiempo (p.e., levantar un dedo cada 10 segundos). Hart (1975) asume que para producir estos intervalos se utiliza un contador mental interno. A medida que aumenta la carga de la tarea primaria, el contador se verá más afectado, y los intervalos serán sistemáticamente sobreesimados.

En la tarea de estimación retrospectiva el sujeto tiene que estimar, después de haber realizado la tarea, la cantidad de tiempo que ha pasado. Hicks, Miller y Gaies (1977) observaron que las duraciones de tiempo son infraestimadas en condiciones de alta carga. Así, ellos encontraron que la ejecución de una tarea de tiempo de reacción más difícil lleva a los sujetos a infraestimar el tiempo que ellos tardaron en realizar la tarea. En un segundo estudio, ellos observaron que el tiempo era infraestimado cuando los sujetos realizaban una tarea de memoria en comparación a una condición

control durante la cual no hacían nada. Los efectos obtenidos por Hicks y col. son explicados mediante un modelo de contador activo o por el razonamiento intuitivo de que el tiempo pasa más deprisa cuando se está entrenado, y más despacio cuando se está aburrido. Por otro lado, usando la técnica de estimación retrospectiva con tareas de aviación, Hart (1975) observó efectos opuestos: mayor dificultad de la tarea primaria causaba estimaciones más largas del tiempo. Sus datos fueron explicados por un modelo de "nivel de adaptación" en el cual se presupone que los sujetos basan sus evaluaciones del tiempo en un nivel "típico" de actividad. Si un periodo de tiempo está vacío, ha ocurrido menos actividad de la normal y el intervalo será infraestimado. Si el periodo está al completo (alta carga) la duración será sobreestimada (para poder hacer tantas cosas necesito un tiempo de por lo menos X segundos).

El problema general de la estimación de tiempos como una medida de carga es la inconsistencia de los posibles procesos y resultados. Su principal beneficio es su bajo grado de interferencia. Las producciones de tiempo necesitan ser realizadas sólo infrecuentemente durante la tarea primaria, y las estimaciones retrospectivas deben ser dadas una vez que la tarea primaria ha sido completada.

5- Generación de números aleatorios: Baddeley (1966) ha investigado una tarea secundaria en la cual los sujetos tienen que generar una serie de dígitos aleatorios. Al calibrar su tarea, Baddeley observó que el grado de aleatoriedad (medido según la teoría de la información) declina monótonamente a medida que incrementa la carga de la tarea primaria. A medida que la carga aumenta, los sujetos generan secuencias repetidas como 123, 123, 123. En términos de la teoría de la información, a mayor carga, menos información es suministrada. Esta, como la tarea de producción de intervalos de Michon, es también algo intrusiva ya que se debe generar una serie continua de respuestas.

6- La tarea de seguimiento de inestabilidad crítica: Jex (1967, 1979) ha desarrollado una tarea de seguimiento que es el análogo computerizado de la tarea de balancear un anillo al final de un dedo. El anillo puede ser acortado o alargado, disminuyendo o incrementando respectivamente la estabilidad de la tarea, y alterando

así la frecuencia con la cual es requerido el control de estabilización. Jex encontró que la ejecución en esta tarea es más efectiva para medir diferencias en la carga que otras tareas de seguimiento. Aún más, es más útil como tarea cargante ya que un nivel dado del parámetro de dificultad λ (análogo por computador del recíproco de la longitud del anillo) esencialmente fuerza al sujeto a asignar una cantidad mínima de recursos a la tarea para no perder el control completamente. A medida que λ aumenta (la longitud del anillo decrece), este requisito mínimo aumenta. Al igual que la tarea de Sternberg y la de producción de intervalos, la tarea crítica ha sido comercializada (Jex, 1979). Como la generación de números aleatorios, tiene la desventaja de que requiere una generación continua de respuestas y por tanto es muy intrusiva.

3.4.1. Ventajas e inconvenientes

La técnica de tarea secundaria posee dos muy distintos beneficios. En primer lugar, tiene un alto grado de validez aparente, ya que precisamente mide la cantidad de atención residual de que dispondrá un operador en el caso de que ocurra un fallo inesperado. Desde luego, el margen de carga de la tarea primaria también posee la misma validez aparente. Sin embargo, las medidas de tarea primaria del margen de carga pueden ser más difíciles de comparar entre diferentes tareas ya que el investigador está limitado a manipular una dimensión de dificultad que es parte de la tarea primaria. La misma dimensión de dificultad puede no estar presente en dos tareas primarias las cuales son comparadas. Por contra, una tarea secundaria dada puede, en principio, ser aplicada a cualquier tarea primaria. Sin embargo, este supuesto tiene algunas dificultades teóricas.

Un costo asociado con esta técnica es que, como el margen de carga, la técnica de tarea secundaria, como cantidad vectorial, debe tomar también en cuenta el hecho de que hay diferentes clases de recursos. Las diferencias en carga que resultan de la manipulación de una variable de la tarea primaria pueden ser muy infraestimadas si los recursos demandados por la tarea primaria no concuerdan con los demandados por

la tarea secundaria. Wickens y Kessel (1980), por ejemplo, encontraron que la tarea crítica de seguimiento de respuesta cargante desarrollada por Jex (1967) no era apropiada como tarea cargante cuando era empleada con una tarea de monitorización perceptiva. Otro ejemplo de este desparejamiento ocurriría si una tarea de comprensión auditiva de palabras o una de aritmética mental (auditiva, verbal, demandas perceptivo/centrales) fuera utilizada para evaluar la carga atribuible a las manipulaciones de la carga de respuesta de seguimiento (visual, espacial, demandas de respuesta). Estas dos tareas demandan recursos muy diferentes y por tanto, las demandas de la tarea de seguimiento serían, en gran medida, infraestimadas.

Un segundo problema que se encuentra frecuentemente con la técnica de la tarea secundaria es que puede interferir o disturbar la ejecución de la tarea primaria. Por un lado, esto puede ser no conveniente o peligroso si la tarea primaria es una como la de vuelo: una diversificación de los recursos en el peor momento podría llevar a un grave accidente. Por otro lado, la intrusión de la tarea primaria podría presentar problemas de interpretación. Cuando dos tareas primarias son comparadas por una tarea secundaria y hay una correlación negativa entre los decrementos de la tarea primaria y la secundaria entre las dos comparaciones, puede ser difícil interpretar cuál de las tareas primarias realmente demanda más recursos.

Es interesante subrayar que una de las soluciones ofrecidas a estos problemas es la elección de tareas secundarias que sean altamente distintas de la tarea primaria (Rolfé, 1973). Este remedio puede ser empleado solamente con un costo potencial -una reducida sensibilidad para los atributos específicos de los recursos de la carga de la tarea primaria. De hecho, la mejor respuesta a todos estos problemas está en las *instrucciones*. Hay que *enfaticar consistentemente que la tarea secundaria es secundaria* y que los decrementos de la tarea primaria deberían ser iguales en todas las tareas primarias y lo más pequeños posible.

Kahneman (1973) ha propuesto que la técnica de tarea secundaria ideal es aquella que emplea una batería de medidas de tarea secundaria sensibles a los diferentes recursos del sistema. Brown (1968) propone que la tarea de generación de

números de Baddeley está asociada con carga perceptiva/cognitiva, y la tarea de producción de intervalos de Michon (o la tarea crítica de Jex) con carga de la respuesta. En los casos en que es claro que un nivel de una dimensión no contribuye a la ejecución de la tarea primaria, la dimensionalidad de la batería puede ser reducida eliminando este tipo de tareas secundarias. Por ejemplo, una tarea de procesamiento verbal sin componentes espaciales no necesita ser evaluada usando una tarea secundaria espacial. Sin embargo, en los casos en los que se realiza una actividad que potencialmente requiere recursos de todos los tipos, una medida de carga segura debería suponer una batería que también incorporase estos tipos de recursos (Wickens, 1986a).

Las diversas medidas de carga de la tarea secundaria pueden, desde luego, ser evaluadas en función de su grado de intrusión. Este factor está negativamente correlacionado con la aceptación del operador. Métodos como la tarea de seguimiento crítica, la tarea de producción de intervalos, o la generación de números aleatorios, que requieren una respuesta continua, son altamente intrusivos. La estimación de tiempos, por otro lado, es ideal en este sentido ya que requiere pocas respuestas. Las tareas de tiempo de respuesta de Posner o la de Sternberg son intermedias si la frecuencia de estímulos es relativamente baja (un estímulo cada 4 a 7 segundos). El problema que surge al reducir el número de estímulos es que cuando se hace la respuesta menos frecuente para evitar la intrusión, los datos se vuelven por definición menos fiables.

3.5. MEDIDAS FISIOLÓGICAS

Una solución al problema de la intrusividad encontrado en la tarea secundaria es registrar, no intrusivamente, las manifestaciones de carga o incrementos en la movilización de recursos a través de medidas fisiológicas, apropiadamente elegidas, de la actividad del sistema nervioso central o el autónomo.

A partir del punto de partida de la teoría de recursos múltiples, las medidas fisiológicas son generalmente menos precisas que las tareas secundarias. La tarea secundaria puede estar asociada razonablemente bien con las demandas impuestas de los diferentes recursos. Sin embargo, todavía no hay suficientes datos para establecer conclusiones definitivas sobre si los cambios en un índice fisiológico particular reflejan cambios en las demandas de ciertos recursos específicos, en cuyo caso la medida es diagnóstica, o cambios en cualquiera o en todos los recursos, en cuyo caso su diagnosticidad es sacrificada a expensas de su sensibilidad. Tres medidas prototípicas serán descritas de forma general y contrastadas en términos de dos criterios, *diagnosticidad* y *sensibilidad*.

1.- **Potenciales evocados:** Cuando se utilizan potenciales evocados se asumen muchos de los supuestos de la técnica de tarea secundaria. Estudios realizados con los potenciales evocados como tarea secundaria (Isreal, Chesney, Wickens y Donchin, 1980; Isreal, Wickens, Chesney y Donchin, 1980; Kramer, Wickens y Donchin, 1983; Natani y Gomer, 1981), encontraron que la introducción de una tarea primaria concurrente de naturaleza perceptivo/cognitiva, típica de la tarea de un controlador aéreo, atenúa la amplitud P300 (el componente de los potenciales evocados más frecuentemente utilizado). El incremento en la dificultad de la tarea por requerir más elementos del display a ser monitorizados reduce aún más la amplitud P300. La medida de potenciales evocados es algo diagnóstica, en el sentido que refleja carga cognitivo/perceptiva, pero es relativamente insensible a las variaciones de la carga de respuesta. Así, su *diagnosticidad* es obtenida a expensas de la *sensibilidad*. La medida de potenciales evocados tiene dos particulares ventajas.



En contraste a las demás medidas fisiológicas descritas a continuación, proporciona una medida graduada directa de la actividad cognitiva más que una medida indirecta de la actividad autónoma. En contraste con la mayoría de las tareas secundarias, por otro lado, no requiere que el sujeto dé respuestas, y por tanto es menos intrusiva.

2.- Diámetro pupilar: Varios investigadores han observado que el diámetro de la pupila correlaciona de forma muy precisa con las demandas de recursos de un gran número de diversas actividades cognitivas (Beatty, 1982) Estas incluyen la aritmética mental (Kahneman, Beatty y Pollack, 1967), carga de la memoria de trabajo (Beatty y Kahneman, 1966; Peavler, 1974), y solución de problemas lógicos (Bradshaw, 1968; Beatty, 1982). Esta diversidad sugiere que la medida del diámetro pupilar puede ser altamente sensible, aunque a costa de no ser diagnóstica. Refleja las demandas impuestas en cualquier lugar del sistema. Su desventaja es que los cambios relevantes de la pupila son del orden de décimas de un milímetro. Esto significa que la medición precisa requiere considerables equipos de medida. Además, los cambios en iluminación ambiental deben ser controlados ya que estos afectan a la pupila. Debido a su asociación con el sistema nervioso autónomo, la medida también es sensible a los cambios emocionales.

3.- Variabilidad en las pulsaciones del corazón: Un número de investigadores han examinado diferentes medidas asociadas con la variabilidad o regularidad del pulso como una medida de la carga mental. Generalmente se encuentra que la variabilidad decrece a medida que la carga aumenta (Mulder y Mulder, 1981). Cuando esta variabilidad está específicamente asociada con las periodicidades resultantes de la respiración, la medida es denominada "*sinus arritmia*" (Kalsbeek y Sykes, 1967; Mulder y Mulder, 1981). Como el diámetro pupilar, la medida de sinus aritmia es sensible a un número de diferentes manipulaciones de la dificultad y parece ser más sensible que diagnóstica. Wickens y Derrick (1981) investigaron esta medida con 4 tareas diferentes realizadas en distintas combinaciones siguiendo el esquema del modelo de múltiples recursos. Ellos concluyeron que la medida de la variabilidad refleja el total de las demandas impuestas en todos los recursos del sistema de

procesamiento mas que la cantidad de recursos en competición (y por tanto los decrementos en tarea dual) entre tareas.

Las medidas descritas anteriormente son las más utilizadas, sin embargo también podemos encontrar algunos estudios que utilizan el nivel de conductancia de la piel para determinar el esfuerzo realizado por el operador (Koelega, Brinkman, Hendriks y Verbaten, 1989), o los movimientos oculares (May, Kennedy, Williams, Dunlap y Brannan, 1990).

3.5.1. Ventajas y limitaciones

Los índices fisiológicos tiene dos grandes ventajas: 1) proporcionan un registro relativamente continuo de los datos sobre el tiempo, 2) no son intrusivas de la ejecución de la tarea primaria. Por otro lado, frecuentemente requieren el uso de electrodos (potenciales evocados y pulso) o la imposición de restricciones de tipo físico (medida pupilar), y no son por tanto realmente no intrusivas en un sentido físico. Estos requisitos influyen en su aceptación. Tienen un mayor costo potencial: las diferencias de carga medidas por medios fisiológicos deben ser utilizadas para inferir qué fallos en la ejecución se producirán, o para inferir cómo el operador se sentirá con la tarea. Las medidas de tarea primaria y secundaria evalúan lo primero directamente, mientras que las medidas subjetivas evalúan lo segundo.

3.6. MEDIDAS SUBJETIVAS

Las medidas subjetivas han sido muy utilizadas para evaluar la carga de los operadores (Moray, 1982; Williges y Wierwille, 1979; Koelega, Brinkman, Hendriks y Verbaten, 1989). Entre las razones para este uso frecuente destacan sus ventajas prácticas (p.e. fácil implementación, no intrusividad) y la existencia de datos actuales que avalan su capacidad para suministrar medidas sensibles de la carga. Se asume que un mayor gasto de capacidad está asociado con los sentimientos subjetivos de esfuerzo

que pueden ser adecuadamente registrados por los sujetos. La aceptación de este supuesto ha llevado a un gran número de investigadores (Gartner y Murphy, 1976; Johanssen y col., 1979; Sheridan, 1980) a sugerir que las medidas subjetivas pueden dar unos válidos y sensibles indicadores de la carga.

Las estimaciones subjetivas de la dificultad de la tarea representan quizás la medida más aceptable de carga desde el punto de vista del usuario de sistemas actual, el cual se siente confortable al establecer u ordenar los sentimientos subjetivos de esfuerzo o demandas atencionales encontradas al realizar una tarea o conjunto de tareas (Eggemeier, 1981; Moray, 1982; Reid, Shingledecker y Eggemeier, 1981). Puede haber circunstancias en las que un diseñador de sistemas prefiera utilizar datos relativos a como se siente un operador con una tarea que a como esta es ejecutada. Es importante, sin embargo, saber con que precisión un operador puede evaluar las demandas impuestas a sus limitados recursos, qué dimensiones subyacen a estas estimaciones, y cómo son escaladas.

A continuación se describen brevemente las escalas y procedimientos de evaluación de la carga subjetiva más importantes.

1.- **Escala de Cooper-Harper:** Quizás la medida subjetiva de carga más antigua y mejor validada es la Escala de Características de Manejo de Aviones de Cooper-Harper. Incluye numerosas referencias a las demandas de la tarea y a la compensación que realiza el piloto. El uso de esta escala como un índice de carga implica asumir que las habilidades de vuelo y la carga del operador están directamente relacionadas (Moray, 1982; Williges y Wierwille, 1979). Es un procedimiento de árbol de decisión que presenta las cualidades de vuelo en una escala de 10 puntos (Cooper y Harper, 1969).

Los datos disponibles avalan la relación entre la escala de Cooper-Harper y la carga. Por ejemplo, Hess (1977) demostró una relación monotóna entre las puntuaciones obtenidas con la escala y el parámetro "fracción de atención" del modelo de control óptimo de Kleinman, Baron y Levinson (1970), el cual ha sido propuesto

como un índice de carga. Esta conclusión también ha sido encontrada por Moray (1982), Jex y Clement (1979) y McDonnell (1968) utilizando tareas de seguimiento. McDonnell (1968) encontró una clara evidencia de una alta correlación entre las puntuaciones obtenidas con la escala de Cooper y el grado de inestabilidad de la tarea secundaria. Jex y Clement (1979) encontraron que esta medida correlaciona altamente ($r=0.96$) con una medida de la capacidad residual evaluada por una tarea de seguimiento. De forma similar, Moray (1982) encontró una $r=0.99$.

Además, la escala parece ser una medida sensible, pero no diagnóstica, de una gran variedad de variables de la tarea de vuelo, como complejidad perceptiva del display, estabilidad del vehículo, etc. (Crabtree, 1975; Krebs y Wingert, 1976; Lebacqz y Aiken, 1975; Schultz, Newell y Whitbeck, 1970).

Se puede concluir que en el dominio restringido de las tareas de seguimiento o de control manual, la escala de Cooper-Harper proporciona una medida fiable y aceptable. Sin embargo, la escala es totalmente específica de la tarea de vuelo y por tanto no es aplicable a tareas de otra naturaleza.

Varios investigadores (North y col., 1979; Wierwille y Casali, 1983; Wierwille, Rahimi y Casali, 1985; Wolfe, 1978) han propuesto algunas versiones modificadas de la escala original de Cooper-Harper. Wolfe (1978) y North y colaboradores (1979) utilizaron una escala muy similar a la de Cooper-Harper en cuanto a su formato pero en la que se sustituyeron las referencias sobre las características de vuelo por descripciones de la carga y el esfuerzo del piloto. Los datos disponibles avalan la sensibilidad de la escala, pero no parece ser diagnóstica de las fuentes de variación de la carga.

Una modificación de la escala de Cooper-Harper más aplicable a un rango general de tareas, es la propuesta por Wierwille y Casali (1983). En esta modificación, las referencias a características de manejo del avión se han sustituido por términos sobre el esfuerzo y la carga de las funciones de procesamiento de la información realizadas por un amplio rango de operadores de sistemas. Esta escala ha sido

evaluada mediante varios experimentos de simulador de vuelo (Casali y Wierwille, 1982, 1983; Rahimi y Wierwille, 1982; Skipper, Rieger y Wierwille, 1986; Wierwille, Rahimi y Casali, 1985), en cada uno de los cuales se manipularon diferentes tipos de carga (perceptiva, de procesamiento central y de comunicaciones). La carga perceptiva (Casali y Wierwille, 1982) fue variada manipulando el grado de presentación y el número de condiciones de alarma a ser detectadas por los pilotos en el panel de vuelo del simulador. Los niveles de procesamiento central (p.e. toma de decisiones, solución de problemas), fueron manipulados (Rahimi y Wierwille, 1982) variando el número y complejidad de las operaciones aritméticas y geométricas requeridas para resolver una serie de problemas de navegación. La carga de las comunicaciones (Casali y Wierwille, 1983) fue variada cambiando el grado de presentación de señales de llamada y la similaridad entre las señales extrañas y aquellas a las que el piloto debía responder. En los tres casos, las puntuaciones obtenidas con esta escala demostraron tener una relación monótona con los niveles de carga. La escala parece ser altamente sensible; sin embargo, proporciona una medida de carga global más que altamente diagnóstica. Como otras escalas, no requiere mucha práctica ni mucha instrumentación para ser aplicada convenientemente. Guías para su uso pueden ser encontradas en Wierwille y Casali (1983).

2.- **Escalas de la Universidad de Estocolmo:** Muy pocas escalas han sido utilizadas para medir la carga o factores asociados, como la dificultad percibida. Dos de estas escalas han sido evaluadas en la Universidad de Estocolmo, una relacionada con la dificultad percibida y otra con el esfuerzo percibido.

* **Escala de dificultad percibida:** Es una escala de nueve puntos utilizada para evaluar la dificultad percibida de items de tests de inteligencia que implicaban razonamiento, habilidad espacial y comprensión verbal (Bratfisch, 1972; Bratfisch, Borg y Dornic, 1972; Hallsten y Borg, 1975). Es una escala simétrica, con niveles verbales asociados a cada una de las nueve categorías de dificultad. La escala produjo puntuaciones de dificultad que estuvieron altamente correlacionados con índices objetivos de dificultad en cada tarea. Las puntuaciones de dificultad estuvieron altamente correlacionadas ($r=-0.81$) con la frecuencia de solución de cada item, lo que

se consideró un índice objetivo de dificultad. Hay datos (Dornic y Andersson, 1980) que indican que, en algunos casos, las puntuaciones de la dificultad percibida difieren de las de esfuerzo percibido en tareas de procesamiento de la información, por lo que hay que ser cuidadoso al interpretar las puntuaciones de dificultad percibida como indicadores directos del esfuerzo o carga del operador. El poder de diagnóstico de esta escala no pudo ser evaluado. Requiere mínima instrumentación y no es necesario tener mucha práctica para aplicarla adecuadamente.

*** Escala de esfuerzo percibido:** Otra escala de la Universidad de Estocolmo deriva juicios del esfuerzo asociado con la realización de una tarea (Dornic, 1980a, 1980b; Dornic y Andersson, 1980). Es una escala gráfica anclada en los extremos por los números 0 y 10 y también tiene niveles verbales. Aunque la escala no ha sido muy utilizada, los datos disponibles sobre su sensibilidad han sido favorables. Sin embargo, se necesitan más datos para poder establecer conclusiones sobre las características (sensibilidad, poder de diagnóstico, etc.) de esta escala. Requiere poca instrumentación y consume menos tiempo que el método de comparación de pares (Dornic y Andersson, 1980).

3.- **Escala dimensional de Sheridan:** (Sheridan y Simpson, 1979) Sheridan (1980) ha propuesto que tres dimensiones definen la experiencia subjetiva de la carga mental. Estas son relativas a la proporción de tiempo ocupado o carga de procesamiento de la información, el esfuerzo mental ejercido o la complejidad de la tarea, y el "estrés emocional" de la tarea. Reid y col. (1981) encontraron que los sujetos son capaces de ordenar las tareas a lo largo de estas dimensiones, y parece encontrarse un buen grado de acuerdo entre operadores en estas ordenaciones. Lo que no está claro, sin embargo, es si estas dimensiones son verdaderamente independientes, si reflejan toda la varianza de la carga subjetiva, ni como se relacionan con las dimensiones de recursos de procesamiento que subyacen a la ejecución de la tarea.

4.- *Mission Operability Assessment Technique*: En esta técnica, una serie de factores, incluyendo la carga del piloto y la efectividad técnica del sistema (grado al cual el sistema ayuda al operador a realizar la tarea con éxito) son combinados dentro de un concepto global denominado "operatividad del sistema" (Donnell y col., 1981; Helm y Donnell, 1979). Consiste en dos escalas ordinales, cada una de 4 puntos, una relativa a la carga del operador y otra sobre la efectividad técnica, que dan lugar a una matriz de 16 elementos sobre la operatividad del sistema. La escala es muy específica de tareas de vuelo, sin embargo se necesitarían muy pocas modificaciones para que aumentara su aplicación a otro tipo de tareas (O'Donnell y Eggemeier, 1986). La escala no representa una medida directa de la carga; aunque la carga es un factor primordial dentro del concepto de operatividad del sistema, las relaciones específicas entre carga y operatividad no han sido especificadas.

5.- *Subjective Workload Assessment Technique (SWAT)*: (Reid, Eggemeier y Shingledecker, 1982; Reid, Shingledecker y Eggemeier, 1981; Reid, Shingledecker, Nygren y Eggemeier, 1981). Asume que la carga está principalmente compuesta por tres dimensiones: carga debida al tiempo, carga debida al esfuerzo mental y carga debida al estrés. Estas tres dimensiones son adaptaciones de factores propuestos por Sheridan y Simpson (1979) y otros autores (Jahns, 1973; Johannsen y col., 1979; Kahneman, 1973; Moray, 1982), como principales fuentes de carga. En esta escala, cada dimensión está representada por una escala individual de tres puntos con descripciones verbales.

La técnica SAWT ha demostrado ser sensible a las variaciones de la carga y de la dificultad de una gran variedad de diferentes tareas (Eggemeier, Crabtree, Zingg, Reid y Shingledecker, 1982; Reid, Shingledecker y Eggemeier, 1981; Notestine, 1983, 1984; Skelly, Reid y Wilson, 1983; Shingledecker, Crabtree, Simons, Courtright y O'Donnell, 1980). Esta gran sensibilidad hace que la técnica no sea diagnóstica, en el sentido de distinguir entre la carga de procesamiento central, la perceptiva y la de respuesta. Es una técnica de papel y lápiz, sencilla y que no requiere mucho tiempo de aplicación.

6.- *NASA Task Load Index (TLX)*: Este instrumento de evaluación ha sido descrito en detalle por Hart y Staveland (1988). El instrumento distingue las seis dimensiones de carga subjetiva siguientes: demanda mental, demanda física, demanda temporal, ejecución, esfuerzo, y nivel de frustración. En general, este instrumento supone dos pasos secuenciales, uno de los cuales es completado antes de comenzar la sesión experimental, y otro inmediatamente después de que se haya realizado la tarea.

En el primer paso, el sujeto compara cada una de las seis dimensiones de carga con las demás, en forma de comparaciones de pares, para determinar qué dimensión de cada par es percibida como mayor fuente de carga. En función del número de veces que haya sido seleccionada una dimensión en las 15 comparaciones de pares, se atribuye un peso a cada una de ellas. Este peso puede, por tanto, variar entre 0 (la dimensión no ha sido ninguna vez elegida como mayor fuente de carga) y 5 (la dimensión ha sido elegida en todos los pares en los aparecía).

En el segundo paso, el sujeto tiene que estimar, en una escala de 0 a 100 subdividida en intervalos de 5 unidades, el papel que ha jugado cada dimensión en la tarea que acaba de realizar. Normalmente estas estimaciones se repiten para cada ensayo.

Con el peso y las estimaciones de cada dimensión se computa un índice global de carga. El valor de este índice viene dado por la suma de las estimaciones de cada dimensión, ponderadas por su peso inicial, dividida por 15 (el número de pares total).

Recientemente, esta técnica ha sido utilizada por investigadores como Vidulich y Tsang (1986) o Hancock (1989).

3.6.1. Ventajas y limitaciones

Los beneficios de las técnicas subjetivas son aparentes. No son intrusivas y son relativamente fáciles de obtener. Su principal costo es la incertidumbre con la que un estamento verbal del operador verdaderamente refleja la disponibilidad o las demandas de recursos de procesamiento.

A pesar de las ventajas de las técnicas subjetivas, hay un número de importantes restricciones en su interpretación. Existen varias guías para su uso que deberían ser consideradas en cualquier aplicación (p.e., Gartner y Murphy, 1976; Sanders, 1979; Sheridan y Simpson, 1979; Williges y Wierwille, 1979). Estas limitaciones generalmente se refieren a la influencia potencial sobre las estimaciones subjetivas de, por un lado, factores que pueden influir en el grado de carga experimentado por el operador (p.e., confusión entre la carga mental y la física), o, por otro lado, limitaciones metodológicas que pueden influir en los niveles de carga registrados (p.e., retrasos en la recogida de datos).

Una limitación potencial en la interpretación de las medidas subjetivas es la posibilidad de que el operador confunda la carga mental con la física. Varios autores (Johannsen y col., 1979; Moray, 1982) han sugerido que los sentimientos de carga pueden estar relacionados con la activación fisiológica, lo que haría posible que se diera una confusión entre ellas. Si se está interesado en una medida global de carga, este no es un problema serio. Sin embargo, si interesa más una distinción entre carga mental y carga física, esta posibilidad de confusión debería ser considerada al interpretar resultados.

Un segundo problema posible es la incapacidad del operador para distinguir las demandas externas, o dificultad de la tarea, del esfuerzo actual o carga experimentada al enfrentarse con esas demandas (Gartner y Murphy, 1976). Esta confusión puede producir estimaciones sesgadas de la carga actual, pudiendo ser sobre o infra estimada ya que el operador siente que la tarea "debería" requerir más o menos trabajo del que el actualmente experimenta (Dornic y Andersson, 1980).

Un tercer factor que podría limitar la utilidad general de las técnicas subjetivas de análisis de la carga hace referencia a la naturaleza de las relaciones entre el gasto actual de capacidad y el esfuerzo experimentado por el operador. El supuesto de que *el incremento en el gasto de capacidad estará asociado con los sentimientos subjetivos de esfuerzo* forma la base teórica de la sensibilidad de las medidas subjetivas. Sin embargo, es muy probable que no toda la actividad de procesamiento realizada por el operador esté disponible para una introspección consciente (Gopher y Donchin, 1986).

Una cuarta restricción en la interpretación de las medidas subjetivas de carga ha sido puesta de manifiesto por varios autores (Vidulich y Wickens, 1986; Wickens y Derrick, 1981; Wickens y Yeh, 1982, 1983), basados en alguna *disociación entre las estimaciones subjetivas de la carga y la ejecución de la tarea primaria en una serie de tareas de procesamiento y de control motor*. El patrón de esta disociación sugiere que las medidas subjetivas están muy afectadas por factores como el número de tareas o de elementos de la tarea a ser realizados, con relativamente poca atención a si las tareas demandan los mismos o diferentes recursos de procesamiento.

En cuanto a las limitaciones de tipo metodológico, una guía muy importante en el uso de las técnicas subjetivas, es determinar específicamente qué tareas o elementos van a ser evaluados por el operador. Sanders (1979) mantiene que los juicios que deben hacer los sujetos sobre la carga deben ser específicos (cuestiones relativas a la influencia de determinadas variables específicas sobre la carga) más que de tipo global.

Otra limitación de las técnicas subjetivas es su dependencia de la memoria a corto plazo del operador que está completando la escala. Cuanto mayor es el periodo de tiempo entre la realización de la tarea y la estimación de la carga, mayor probabilidad habrá de obtener estimaciones distorsionadas. Sin embargo, existen muy pocos datos al respecto (Eggemeier y col., 1983; Notesune, 1983) y además estos no parecen ser muy concluyentes. De todas formas, y hasta disponer de más datos, la mejor guía para minimizar la pérdida de información en la memoria a corto plazo, es que el sujeto estime la carga percibida inmediatamente después de la tarea.

3.7. RELACION ENTRE LAS MEDIDAS DE CARGA

Si todas las medidas de carga demostraran alta correlación con las demás y la varianza residual fuera debida al error aleatorio no sería necesario realizar más investigación para su validación. Se podría aplicar cualquiera de ellas, la que pareciera más fiable y menos costosa para el problema concreto que se esté tratando. Generalmente, se encontrarán altas correlaciones entre las medidas si estas son evaluadas entre tareas de estructura similar y con grandes variaciones de la dificultad.

Jex y Clement (1979) encontraron una alta correlación entre las medidas subjetivas y las de tarea secundaria de la dificultad del control de vuelo. Sin embargo, las correlaciones pueden no ser altas, e incluso ser negativas cuando se comparan tareas muy diferentes. Un ejemplo es el experimento dirigido por Heron (1980) en el cual una innovación diseñada como ayuda a una tarea fue preferida por los usuarios al prototipo original, pero generó una peor ejecución. Disociaciones similares han sido observadas por Childress, Hart y Bortolussi (1982) midiendo la carga de los pilotos asociada con las innovaciones del display de la cabina de vuelo. Por tanto, es importante determinar que características de una tarea influyen más fuertemente en una determinada medida de carga mientras que no afectan a otra medida. Por ejemplo, Yeh y Wickens (1988), Wickens y Derrick (1981) y Wickens y Yeh (1983) concluyeron que las medidas subjetivas eran relativamente más sensibles al número de actividades que comparten los mismos recursos que debe realizar un operador, mientras que las medidas de la ejecución de la tarea primaria reflejaban en mayor grado la dificultad de una tarea simple. La ejecución de la tarea primaria es también relativamente más sensible a la competición de recursos entre tareas que las medidas subjetivas.

Cuando aparecen estas disociaciones entre medidas, la cuestión de cual es la mejor depende claramente del uso que se vaya a dar a la información recogida. Si se utiliza la carga para predecir los márgenes de ejecución o "atención residual" disponible para enfrentarse con los fallos en ambientes operativos críticos, lo más adecuado sería utilizar las medidas de tarea primaria o secundaria como guía para la

elección entre sistemas, a pesar de que el sistema elegido pueda presentar mayores puntuaciones subjetivas o fisiológicas de dificultad. Por otro lado, si la cuestión es la facilidad y grado de uso por el consumidor, el establecimiento de esquemas de trabajo-descanso, o la satisfacción en el puesto, y las variaciones en ejecución son relativamente menos críticas, entonces se debería dar más importancia a las medidas subjetivas (y quizás también a las fisiológicas).

La selección de una u otra técnica de medida de la carga debe estar guiada por los objetivos de cada aplicación y por el estudio y análisis de las características de cada una de ellas.

En cuanto a su sensibilidad, las medidas de tarea primaria parecen ser insensibles a las variaciones de carga cuando los niveles de carga son entre bajos y moderados y la ejecución del operador es adecuada. En este caso, el incremento en el nivel de carga no varía el nivel de ejecución ya que el operador dispone de suficiente capacidad o recursos para compensar los aumentos de carga (la ejecución será siempre adecuada). En este caso, las medidas subjetivas, fisiológicas o las de tarea secundaria serán más sensibles. En particular, la técnica de tarea secundaria está especialmente diseñada para evaluar la capacidad de procesamiento de reserva que queda cuando se han asignado los recursos suficientes a la tarea primaria.

En el caso de situaciones con niveles más altos de carga, a medida que la carga aumente disminuirá la ejecución. En este caso, el procedimiento más adecuado es el de tarea primaria. Aunque también serían adecuados los demás procedimientos, ellos sólo serán requeridos cuando el método de tarea primaria no sea lo suficientemente sensible como para detectar pequeñas variaciones de la carga. En situaciones de carga extrema la ejecución será catastróficamente baja y, aunque se podría utilizar cualquier técnica, la evaluación de la carga en estas situaciones sería más difícil utilizando procedimientos de medida de la ejecución.

En cuanto al poder de diagnóstico, las medidas subjetivas, generalmente son poco diagnósticas debido a la incapacidad del operador para discriminar recursos

individuales. Igualmente, las medidas de tarea primaria muestran un bajo poder de diagnóstico en la mayoría de los casos, ya que normalmente no es posible identificar la fuente específica de un fallo en la ejecución. Las medidas de tarea secundaria son consideradas como altamente diagnósticas y proporcionan un índice de la carga impuesta sobre recursos específicos. Las medidas fisiológicas pueden ser globales (diámetro pupilar) o altamente diagnósticas (potenciales evocados).

No hay datos sistemáticos sobre el grado de intrusión de cada una de las técnicas. Aunque ha habido algunos estudios al respecto (Casali y Wierwille, 1982, 1983, 1984; Rahimi y Wierwille, 1982; Wierwille y Connor, 1983; Wierwille, Rahimi y Casali, 1985) que permiten alguna comparación entre distintas técnicas, no hay datos concluyentes. De todos modos, el procedimiento más intrusivo parece ser el de tarea secundaria. Las técnicas subjetivas parecen ser las menos intrusivas, ya que las escalas se rellenan una vez realizada la tarea.

Por otro lado, las técnicas subjetivas presentan generalmente los menores requisitos de implementación ya que normalmente son pruebas de papel y lápiz. Las técnicas de tarea primaria también pueden ser frecuentemente utilizadas con una dificultad de implementación mínima. Los aparatos necesarios y los requisitos para el análisis de los datos son normalmente más elevados cuando se utilizan las medidas fisiológicas y los procedimientos de tarea secundaria. Además, la mayoría de las técnicas de tarea secundaria requieren un periodo de entrenamiento del operador.

Con muy pocas excepciones (Hallsten y Borg, 1975; Katz, 1980), hay muy pocos datos relativos al grado de aceptación de cada una de las técnicas. El grado de aceptación depende en gran medida de la población a la que pertenecen los sujetos. Cuanto más familiarizados estén los participantes con este procedimiento o con la tarea, mayor aceptación tendrá. En este sentido es muy importante explicar adecuadamente a los sujetos los propósitos y procedimientos de evaluación, lo cual es especialmente importante cuando se utilizan medidas fisiológicas.

**CAPITULO 4. LA CARGA MENTAL EN TAREAS DE
DETECCION Y DIAGNOSTICO DE FALLOS**

4.1. INTRODUCCION

Ha habido un gran interés por el análisis y estudio de las variables relacionadas con la detección y el diagnóstico de fallos en multitud de sistemas y operaciones de diferentes tipos. Las razones principales de este gran interés se relacionan, por un lado con los grandes avances y cambios que ha sufrido la mayor parte de los puestos de trabajo actuales, que han cambiando en cuanto a los recursos que demandan, dejando atrás las tareas de tipo manual y siendo cada vez más actividades fundamentalmente de tipo cognitivo; y, por otro lado, por necesidades prácticas del mundo real. De esta forma, la mayoría de los estudios de campo realizados sobre las tareas u operaciones de mantenimiento (Christensen y Howard, 1981), concluyen que la mayor parte del tiempo de mantenimiento (entre el 65 y el 75%) se dedica al diagnóstico de fallos, y que además, la mayor parte del tiempo dedicado al diagnóstico no es utilizado adecuadamente (errores que llevan al operador a pensar que el fallo está en un componente o pieza que después resulta no estar defectuoso). Por ejemplo, Gardénier (1981), estudiando los procesos de detección y diagnóstico en la navegación marítima, pudo concluir que en la mayoría de los accidentes (sobre el 80%) la causa está en un fallo humano. Los estudios realizados en este campo (Paramore y col., 1979; Gray, 1978) demuestran que en la mayoría de los casos de accidentes, el personal responsable estaba bien preparado y entrenado, era competente y no estaba bajo los efectos de drogas ni alcohol, ni estrés emocional ni estrés ambiental. Además, rara vez ha habido problemas en la actuación de control. Una vez que se ha tomado una decisión, las operaciones se realizan rápida y adecuadamente.

Los problemas más frecuentes son, por tanto, debidos a errores en la detección y el diagnóstico del fallo. Muchas veces estos errores se deben a la confusión de información vista o a mirar un display que no es el adecuado, o a errores de detección, los cuales dan lugar a diagnósticos incorrectos.

Los errores al detectar o diagnosticar una situación anormal en un sistema hombre-máquina multivariado pueden ocurrir por muchas razones. Por ejemplo, puede

ocurrir que la variable relevante no sea mostrada al operador, como fue el caso del valor electromagnético en el accidente de Three Mile Island (IEEE Spectrum, Special Issue on Three Mile Island, 1979). O también puede ser que las expectativas de los operadores les lleven a la no interpretación de la información mostrada, como sucedió en el accidente aéreo de Tenerife (Human Factors Report on the Tenerife Accident, 1979). Puede ser que el operador ignore información que está disponible, concentrando su atención en unas variables y excluyendo otras, como sucedió en el accidente de la Eastern Air Lines 401 en Everglades (Weiner, 1977). O puede ser que el operador humano esté aturdido por demasiada información, como ocurrió de nuevo en Three Mile Island, donde un operador testificó que podría haber apagado todas las luces de alarma ya que no proporcionaban ninguna información útil para el diagnóstico.

Se asume que la intervención humana apropiada en el manejo de fallos o anomalías del sistema implica tres tareas o etapas diferentes:

- . Detección del fallo.
- . Diagnóstico.
- . Acción de remedio.

La intervención inadecuada puede deberse al error en cualquiera de estas tres etapas, y fundamentalmente será el resultado de la descompensación entre los recursos del operador y las demandas de la tarea (carga mental).

A continuación se expondrán los factores que la mayor parte de las investigaciones resaltan como los que tienen una mayor influencia en la detección y el diagnóstico de fallos. Algunos de estos factores se deben principalmente a características de la tarea, mientras que otros harán referencia fundamentalmente a las capacidades y recursos o características de los sujetos. Debido a que la mayoría de estos factores afectan a ambos procesos, y a que normalmente ambos van unidos (en las operaciones de los sistemas reales, los operadores no sólo deben detectar un fallo sino también emitir un diagnóstico del mismo), no se distinguirá entre factores que afectan a la detección y factores que afectan al diagnóstico, si bien resulta evidente

que algunos de ellos afectan más a la detección y otros al diagnóstico. No se expondrán los factores que pueden influir en la realización de las acciones de remedio, ya que esta no es la finalidad de esta investigación.

4.2. FACTORES QUE AFECTAN A LA DETECCION Y DIAGNOSTICO DE FALLOS

Estudios experimentales han demostrado que los sujetos que tienen que solucionar un problema no siguen estrictamente reglas, como por ejemplo la de reducción de la incertidumbre (Rigney, 1968; Rouse, 1978). Los sujetos del estudio de Rigney (1968) operaron con una eficacia máxima de un tercio, y en otro experimento (Bryan y col, 1956), sólo la mitad de los test electrónicos realizados por los sujetos tenían sentido. Esta relativa ineficacia no debería ser atribuida a torpeza o estupidez, sino más bien a que la base de información de los técnicos puede ser incompleta o incorrecta.

Otro factor implicado en la ineficacia podría ser la tendencia a "hacer algo"; Rasmussen y Jensen (1974) encontraron que los técnicos parecían estar poco interesados en seguir una búsqueda lógica del problema, sin embargo estaban muy preocupados por la corrección del problema particular que estaban tratando (Bond, 1981).

Einhorn (1970) resalta la importancia de la tarea, la personalidad del sujeto y las condiciones ambientales, como determinantes de la estrategia de decisión seleccionada.

A continuación se describe la importancia y efecto que sobre la detección y el diagnóstico de fallos en el sistema, pueden tener ciertos factores. El conocimiento de estos factores nos ayudará a determinar aquellas situaciones en las que se producirán mayores o menores niveles de carga.

4.2.1. Características de los displays

Es evidente la influencia de este factor en la detección y el diagnóstico de fallos. El tipo de display, la saliencia de la información, el uso de determinadas luces, colores, la localización de la información, etc., son características que pueden facilitar o dificultar la detección de una situación anómala, haciendo más o menos perceptibles ciertas informaciones. Es decir, la ejecución puede verse afectada por las características visuales del problema. Por ejemplo, Brooke y Duncan (1981), utilizando redes funcionales para representar el sistema, encontraron que se cometían menos errores de diagnóstico y se necesitaba menos tiempo para emitirlo cuando las conexiones entre los componentes de la red eran líneas rectas y la información fluía de izquierda a derecha. Otro autor, Earing (1977) encontró efectos de la saliencia de las variables que indicaban el estado del sistema.

Hay suficiente evidencia que sugiere que la habilidad del operador para monitorizar y diagnosticar fallos del sistema eficazmente, depende de la forma en la que la se muestra la información (Barnett y Wickens, 1988; Boulette, Coury y Bezar, 1987; Carswell y Wickens, 1987; Coury, Boulette, Zubritzky y Fisher, 1986; Coury y Pietras, 1989; Woods, Wise y Hanes, 1981). Por ejemplo, Coury y Pietras (1989) encontraron que la información necesaria para supervisar el funcionamiento normal de una planta de procesamiento de fluidos simulada, era significativamente diferente a la información requerida para controlar una planta en condiciones de fallo. Además, la forma en la que la información y los datos eran presentados a los operadores tuvo un efecto significativo sobre su habilidad para controlar el sistema y dar solución al problema. De particular interés fue el hecho de que los operadores alteraron significativamente sus estrategias de obtención de información cuando los datos y la información eran mostrados de forma subóptima.

Las formas eficientes de mostrar información han sido un gran tema de interés de la Ergonomía desde sus comienzos (Wickens y Kramer, 1985). Debido a los grandes y rápidos avances en tecnología de computadores, las capacidades gráficas

incrementaron, y se han introducido nuevos tipos de problemas de investigación. La creciente complejidad de los sistemas dinámicos actuales requiere la presentación de una gran cantidad de información al operador humano. Al mismo tiempo, los rápidos avances en computador y tecnología de displays han incrementado la capacidad para presentar información compleja, multielemento, en un único display, y la libertad para seleccionar el aspecto y el modo de presentación.

Hay dos grandes cuestiones en el diseño de displays: a) su habilidad para representar propiamente el mundo que se desea representar, y b) las propiedades físicas del propio display. Las guías para el diseño de displays todavía se focalizan en las cualidades psicofísicas del mismo, con el objetivo de mejorar su legibilidad (p.e. Helander, 1987). Sin embargo, hay un reconocimiento creciente de la necesidad de considerar además el procesamiento perceptivo y cognitivo humano (Foley y Moray, 1987; Wickens, 1987). La literatura de la Ergonomía Cognitiva de los últimos años muestra un creciente interés en aspectos representacionales y de procesamiento de la información, lo que une a esta disciplina con la Psicología Cognitiva.

Si la función de un display es comunicar información y asegurar el procesamiento eficiente de la misma, el diseñador debe representar la información contenida en el display de forma que proporcione la información ha ser representada y las operaciones ha ser realizadas sobre ella.

Podemos aprender mucho sobre cómo organizar, acceder y manipular la información mostrada de la investigación actual en Psicología Cognitiva, incluyendo el reconocimiento de patrones, la búsqueda visual, mapas cognitivos (Chase, 1986) y la organización perceptiva. Particularmente relevantes son los estudios teóricos y empíricos que se están realizando sobre agrupamiento, interacciones dimensionales, procesamiento analítico y holístico, relaciones perceptivas entre aspectos globales y locales de patrones visuales, procesamiento de arriba a abajo o top-down (Treisman, 1986) y el papel del análisis de la frecuencia espacial en la percepción de formas y objetos (Ginsburg, 1986).

4.2.1.1. Compatibilidad entre el procesamiento central y el estímulo

La importancia de la compatibilidad entre el formato de la información mostrada y los códigos de procesamiento central, particularmente en la memoria de trabajo, es bien reconocida (Wickens, 1987).

A medida que los displays visuales son más frecuentemente utilizados para representar sistemas complejos, hay un mayor interés en la compatibilidad entre la información mostrada y el modelo mental que los operadores humanos tienen de los sistemas. Los displays de sistemas complejos pueden mejorar la ejecución humana si ellos se ajustan a los modelos mentales humanos del sistema, o si ayudan a construir modelos mentales correctos (Eberts y Schneider, 1985; Eberts, 1986; Coury y col., 1986; Goettl, Wickens y Kramer, 1991; Sanderson, 1986; Woods, 1986).

La información mostrada puede situarse a lo largo de un continuo, definiendo la extensión a la cual esta información es inherentemente espacial-analógica (p.e., información relativa a localizaciones, transformaciones análogas o movimientos continuos) o lingüística-simbólica y verbal (p.e., instrucciones, códigos alfanuméricos, direcciones u operaciones lógicas). Los dos extremos de este continuo "espacial-verbal" también parecen definir dos diferentes sistemas de memoria de trabajo que pueden ser utilizados para realizar las transformaciones mentales necesarias para llevar a cabo las tareas (Baddeley y Hitch, 1974). Estos sistemas de memoria pueden ser denominados "espacial" y "verbal". Además, los cuatro formatos de display principales (habla, sonidos, imágenes, dibujos analógicos) pueden ser definidos en términos de la modalidad sensorial (auditiva vs. visual) y el código de procesamiento (verbal vs. espacial-analógico).

Existe una serie de guías para elegir el formato de display más adecuado a las demandas de una tarea. Estas guías especifican enlaces de alta compatibilidad entre formatos del estímulo y las operaciones de procesamiento central (Wickens, Sandry y Vidulich, 1983; Wickens, Vidulich y Sandry-García, 1984). Esto es conocido como compatibilidad estímulo-procesamiento central. Específicamente, las tareas con

requisitos espacial-analógicos (p.e., navegación, seguimiento, o evaluación de localizaciones relativas) se benefician más de formatos visuales, particularmente visual-espaciales. Por contra, las tareas que utilizan extensamente la memoria de trabajo verbal o fonética (p.e., cuando dígitos y palabras deben ser repetidos o recitados) se benefician más de inputs de tipo auditivo-verbal (p.e., habla). Debido a que el canal del habla es necesariamente serial y transitorio, puede ser importante añadir a los mensajes información visual redundante, particularmente si el mensaje es largo. El formato espacial-auditivo (p.e., intensidad de tonos, localización espacial aparente) no parece ser óptimo para el procesamiento de cualquier tipo de información, exceptuando las alarmas simples, pistas redundantes o cuando otros formatos están fuertemente sobrecargados (Thompson, 1981; Vinge y Pitkin, 1972; Wickens y col., 1984).

La codificación redundante en diferentes formatos es de gran importancia. Si la información es presentada para su uso inmediato (Garner y Fefoldy, 1970) o se trata de instrucciones sobre procedimientos (Booher, 1975), el uso redundante de diferentes formatos es beneficioso. Por ejemplo, la presentación redundante de información en modalidades auditiva y visual permitirá evitar el ruido del ambiente de procesamiento (p.e., distracción verbal, ruido de fondo, etc.) que puede influir en un formato más que en otro. El uso redundante de instrucciones espaciales (pictóricas) e impresas permitirá su utilización adecuada por grupos de sujetos con diferentes habilidades (p.e., usuarios con alta habilidad espacial vs. sujetos con alta habilidad verbal). Los costos de construir formatos redundantes en términos de displays extra serán usualmente compensados por el aumento de la fiabilidad humana que produce la redundancia.

Un segundo aspecto de la compatibilidad estímulo-procesamiento hace referencia al modelo interno. Los operadores tienen modelos internos que describen sus concepciones del sistema. Estos modelos internos frecuentemente se corresponden con la manera en que estos sistemas se comportan en el mundo real (Gentner y Stevens, 1983; McClosky, 1983). La información debería ser mostrada de forma que se corresponda con estos modelos internos (Roscoe, 1968). La generación de estos modelos internos está determinada por las experiencias previas de los operadores con

el sistema. El papel del conocimiento anterior o expectativas también ha sido estudiado. La mayoría de los modelos actuales de percepción de objetos y eventos lo ven como un proceso interactivo entre el procesamiento bottom-up (dirigido por los datos) y el procesamiento top-down (dirigido conceptualmente) (Rumelhart, 1977; Treisman y Schmidt, 1982). Este aspecto de los modelos mentales del operador, y su relevancia para la detección y el diagnóstico de fallos, serán discutidos posteriormente (ver apartado 4.2.2).

4.2.1.2. Principio de Compatibilidad de Proximidad.

Displays Integrales vs. Separados

Cómo se combinan las dimensiones físicas para formar dimensiones perceptivas ha sido extensamente estudiado por Garner y sus colegas (Garner, 1974, 1978). Garner distingue entre dimensiones separables, integrales y configurales. Estímulos que varían en una sola dimensión son percibidos como una entidad unitaria, mientras que los que varían en dimensiones diferentes son percibidos en función de sus distintas dimensiones o atributos. El sistema perceptivo humano tiene una capacidad limitada para procesar información de diferentes estímulos a la vez, sin embargo, puede procesar en paralelo diversas dimensiones de un solo objeto (Kahneman y Henik, 1981). Por tanto, cuando es necesario observar e interpretar diversas variables dentro de un solo display, estas serán más rápida y eficazmente interpretadas si son representadas en un formato integrado. Las dimensiones integrales facilitan la ejecución cuando están perfectamente correlacionadas, y en aquellos casos en los que la atención selectiva a cada dimensión es imposible. Las dimensiones separables permiten la atención selectiva a cada dimensión, pero no facilitan la ejecución cuando son redundantes.

En un esfuerzo para proporcionar un esquema organizador de la investigación sobre los aspectos cognitivos de la interpretación de datos, Wickens y sus colaboradores (Barnett y Wickens, 1988; Boles y Wickens, 1987; Carswell y Wickens, 1987; Casey y Wickens, 1986; Fracker y Wickens, 1989; Wickens, 1986b; Wickens

y Boles, 1983; Wickens y Andre, 1990), han propuesto y probado el **principio de compatibilidad de proximidad**. Según este principio, cuando la información tiene que ser integrada debería estar próxima en el display. Es decir, en aquellas tareas en las que se requiere una "alta proximidad mental" (p.e., integración de la información), la ejecución mejorará cuando las variables estén próximas, mientras que en las tareas que requieren "baja proximidad mental" (p.e., procesamiento independiente de dos o más variables, o la focalización de la atención en una mientras se ignoran otras), la ejecución se beneficiará de variables o displays más separados.

En principio, la proximidad puede ser definida por la cercanía en el espacio o en el tiempo, pero también en términos de otros atributos relativos a la forma del estímulo, *objectness*, integrabilidad, u otros principios de agrupamiento de la Gestalt (Barnett, Goetd, Kramer y Wickens, 1986). Aunque Wickens y Boles (1983) propusieron que la proximidad podía ser definida en términos de la proximidad espacial, Wickens y Andre (1990) concluyeron que la proximidad espacial no parece conformar con este principio. En su estudio (Wickens y Andre, 1990), los autores hacen referencia a dos posibles formas de definir la proximidad de las variables: una por atributos físicos (cercanía espacial y similitud del color) y otra por *objectness* (p.e., dos dimensiones combinadas en un solo objeto están más próximas que dos dimensiones de diferentes objetos). Los resultados de este estudio avalan el principio de compatibilidad de proximidad cuando esta está definida por el color. La separación espacial tuvo un efecto muy limitado sobre la atención y la integración de la información. Por tanto, parece que el espacio, por sí solo, no es una dimensión relevante en el display. Por otro lado, el grado de *objectness* resultó ser una buena medida de la proximidad, y facilitó, como había sido previsto por el principio de proximidad, la integración, mejorando la ejecución.

Las discrepancias entre estos resultados y los de investigaciones previas (Andre y Wickens, 1988, 1989) sobre este principio, pueden ser atribuidas al diseño experimental utilizado ya que, en los anteriores experimentos, la proximidad espacial y de color fueron manipuladas conjuntamente.

Una solución muy utilizada para resolver el problema de presentar información multidimensional de sistemas complejos ha sido utilizar displays integrales (*object-like*). Los formatos de display integrales usan varias dimensiones de un único objeto para retratar el estado del sistema (p.e. polígonos, caras esquemáticas, etc). Los formatos de display separables usan displays univariados separados, bien en un formato digital tradicional (alfanumérico) o bien utilizando la misma dimensión de varios objetos para mostrar información multivariada (p.e. diagramas de barras). Muchos estudios han encontrado que los displays integrales son superiores a los separables cuando las variables de los datos están altamente correlacionadas y/o cuando la integración de los datos a partir de varias fuentes es requerida antes de tomar una decisión (Goldsmith y Schvaneveldt, 1984; Carswell y Wickens, 1988; Beringer, 1985; Beringer y Chrisman, 1987; Boulette y col., 1987).

En este sentido Cooper y colaboradores (Cooper, 1976, 1982; Cooper y Regan, 1982) llegaron a la conclusión de que los sujetos utilizan diferentes estrategias (analíticas y globales) en la comparación de representaciones visuales. Cuando los estímulos son multidimensionales los sujetos son consistentes con el modelo analítico, mientras que cuando se trata de representaciones holísticas (p.e., fotografías humanas), la mayor parte de los sujetos utilizan una estrategia de tipo global. Además, los sujetos globales son más flexibles en el uso de estrategias alternativas que los analíticos. Cooper (1982) consideró que estas distinciones se podían deber a diferencias en el modo de representación y codificación de patrones visuales en la memoria. Por tanto, la forma (integral o separada) en la que se muestra la información determina en muchos casos la estrategia (global o analítica) que van a utilizar los sujetos.

La superioridad del display integral (*object-like*) puede ser atribuida a dos propiedades de la percepción humana: a) El sistema perceptivo humano tiene una capacidad limitada para procesar una única dimensión con objetos múltiples al mismo tiempo, mientras que es capaz de procesar en paralelo varias dimensiones de un único objeto (Kahneman y Treisman, 1984); b) Las características globales u holísticas pueden ser procesadas más rápidamente que las características locales (Navon, 1977, 1981; Pomerantz, 1981). En los displays integrales, los cambios en el estado del

sistema son representados mediante cambios en la forma global del objeto, de forma que pueden ser procesados globalmente (Munson y Horst, 1986). En términos de la teoría de múltiples recursos (Wickens, 1984), este tipo de displays produce un código espacial que permite el procesamiento integral de los datos del sistema.

Sin embargo, la ventaja de un display integral puede anularse bajo ciertas condiciones. Uno de los factores que puede tener un profundo efecto sobre el procesamiento de datos correlacionados, multidimensional, es la incertidumbre. Por ejemplo, Coury y col. (1986a; 1989) demostraron que cuando el estado del sistema era cierto, el operador era capaz de clasificar más rápidamente los displays integrales que los separables, pero cuando el estado del sistema era incierto, los displays separables eran superiores a los integrales.

Otro importante determinante de la efectividad de los displays integrales son los requisitos atencionales de la tarea. Por ejemplo, Casey (1987) comparó displays integrales y separables en la detección y el diagnóstico de fallos de sistemas cuyas variables estaban relacionadas por correlación o por causalidad. En ambos tipos de sistemas los displays separables fueron superiores. La tarea de diagnóstico requería focalizarse en los componentes del display en orden a identificar la causa del fallo. Así, mientras el procesamiento holístico proporcionado por el display integral es útil para el análisis del estado general, va en detrimento de la ejecución cuando las tareas requieren atención selectiva, siempre que los componentes del sistema estén fuertemente interrelacionados. Esta idea ha sido puesta de manifiesto muy recientemente en el estudio realizado por Goettl, Wickens y Kramer (1991), en el que se comparan dos formas de representación gráfica de datos, diagrama de puntos (información separada) y gráfico de barras (display integrado) (Coury, Boulette y Smith, 1989). Los autores concluyen que la forma más eficaz para representar datos gráficos depende de la tarea que se esté realizando.

Se ha dedicado un gran esfuerzo de investigación al estudio de la relación entre el tipo de display (integral vs. separable) y dos factores: la estructura de la información ha ser presentada (correlacionada vs no correlacionada) y la naturaleza

de la tarea (integral vs no integral). Las dimensiones físicas utilizadas para representar la información han recibido mucha menor atención. Aunque se admita que no hay una relación uno-a-uno entre las dimensiones físicas y las perceptivas, sería importante considerar cómo las dimensiones físicas interactúan para formar las dimensiones perceptivas en función de su integrabilidad, separabilidad y configurabilidad. Un paso en esta dirección ha sido dado recientemente por Barnett y Wickens (1988).

Los displays *object-like* han mostrado ser de gran utilidad para reducir el desorden o desorganización de los displays de las cabinas de aviones (Taylor, 1987) y de las salas de control de plantas nucleares (Woods, Wise y Hanes, 1981). La cuestión clave para establecer el éxito de estos displays descansa en su apropiada configuración para generar características emergentes que son críticas para el operador (Wickens y Andre, 1990). En este sentido Sanderson, Flach, Buttigieg y Casey (1989) han mostrado como un display en forma de diagrama de barras puede ser superior al de un objeto en una tarea de integración, si el gráfico de barras tiene características fuertemente emergentes.

4.2.1.3. Displays Gráficos vs. Alfanuméricos

Una gran línea de investigación se focaliza en cuestiones representacionales. Cuestiones típicas de esta línea de investigación son la selección y evaluación de simbología de displays, las ventajas de un tipo de formato sobre otro, y los beneficios de formatos puros versus mixtos en displays complejos.

Se comparan los formatos analógicos con los digitales. En este contexto los formatos analógicos se refieren y se definen en función de lo espacial, de la representación continua. En muchos casos, formatos analógicos y representaciones gráficas se refieren a lo mismo, son intercambiables. Los formatos analógicos tradicionales incluyen diagramas de barras, clusters de puntos, y diales. Los formatos digitales incluyen codificación alfanumérica como dígitos, letras y palabras. La relativa eficacia de estos formatos ha sido investigada con una gran variedad de tareas. Por

ejemplo, Boles y Wickens (1987) compararon formatos analógico (diagramas de barras), digital y verbal en una tarea de juicio numérico y encontraron que los indicadores analógicos fueron respondidos más rápidamente que los digitales o los verbales. Schwartz y Howell (1985) compararon la ejecución en una tarea de seguimiento bajo condiciones en las que la información de la posición fue presentada gráfica y digitalmente. La ejecución mejoró con los displays gráficos, particularmente bajo condiciones de cambio rápido. Bauer y Eddy (1986) estudiaron la representación de sintaxis de lenguaje de comandos. Ellos compararon el uso de metacaracteres especiales y de representaciones gráficas para representar relaciones gramaticales, y encontraron que la representación gráfica era superior tanto durante el aprendizaje como en una tarea de referencia. Además, Eberts (1987) concluyó que los gráficos disminuyen la carga.

Boles y Wickens (1987) encontraron que las tareas que requerían la integración de elementos del display se beneficiaban de displays de formato puro, mientras que las tareas duales lo hacían de displays de formato mixto.

Los símbolos gráficos, especialmente pictogramas, pueden ser preferidos a los alfanuméricos debido a la semejanza entre la forma del símbolo y la del objeto que representa. Esto puede ser especialmente beneficioso bajo demandas muy duras de memoria. Sin embargo, la similitud intra-grupo entre pictogramas puede incrementar el tiempo de búsqueda y de identificación. Por ejemplo, Remington y Williams (1986) usaron una tarea visual con un sólo target para evaluar un conjunto de símbolos de CRT en una situación de display de un helicóptero. Ellos encontraron que los símbolos numéricos eran superiores a los gráficos. Los autores atribuyen este hecho a la familiaridad y discriminabilidad de los símbolos numéricos, por un lado, y al alto grado de similitud intragrupo entre los símbolos gráficos, por el otro. Más recientemente, Workman y Fisher (1987) propusieron una nueva métrica de la similitud basada en el grado de solapamiento entre "dibujos borrosos" de los símbolos. Las puntuaciones en similitud derivados de esta métrica pueden ser utilizadas para seleccionar el subconjunto más discriminable de un conjunto de símbolos con significado.

Coury y Pietras (1989) compararon los efectos de tres tipos de display: display digital, display gráfico, y display múltiple (con información digital y gráfica), sobre la detección de fallos en un sistema dinámico simulado. Ellos encontraron que la información digital, por sí sola, era insuficiente para la tarea de detección de fallos. Con el display digital los operadores tardaron más tiempo en detectar el fallo que con el display gráfico. Pero la conclusión más importante fue la gran superioridad del display múltiple. Esta superioridad del display múltiple fue atribuida, en parte, a la presentación redundante de información y datos del sistema, y a la evidencia confirmatoria suministrada por dos diferentes representaciones del estado del sistema. Además, aunque en su investigaciones no se consideró el efecto de la presión temporal, los resultados sugieren que el display múltiple proporcionaría la mejor ejecución de detección de fallos en tareas con limitación temporal.

Por otro lado, Coury, Boulette y Smith (1989), Buttigieg y col. (1988) y Coury y Purcell (1988) encuentran que los gráficos de barras pueden poseer propiedades integrales y separables. Sus resultados sugieren que el display de gráfico de barras puede ser un formato muy versátil y es potencialmente la mejor elección para mostrar datos correlacionados, multidimensionales, cuando el rango de incertidumbre es alto.

Una línea relacionada de investigación se focaliza en la representación de gráficos generados por computador (Hochberg, 1986). El diseñador de estas representaciones debe considerar tanto la estructura física relevante (p.e. varias perspectivas) como los procesos perceptivos relevantes (Powrie, 1989). Un intento para desarrollar displays más eficaces para representar información tridimensional sobre pantallas bidimensionales fue el display en perspectiva para el control del tráfico aéreo (Ellis y col., 1987; McGreevy y Ellis, 1986). El uso del display en perspectiva se consideró más "natural" que el convencional (plano), y mejoró el tiempo de decisión (Ellis y col., 1987).

4.2.1.4. Displays Predictivos

Los humanos no son buenos prediciendo estados futuros a partir de la información actual. Las tareas que requieren que el operador realice predicciones del estado futuro del sistema, serán realizadas de forma inadecuada, y normalmente el operador se tendrá que enfrentar con una fuerte carga. En consecuencia, no es sorprendente que la mayoría de las investigaciones encuentren que los displays predictivos mejoran la ejecución de control y planificación (West y Clark, 1974; Whitfield, Ball y Ord, 1980; Smith y Crabtree, 1975; Jensen, 1981). Este tipo de displays liberan a la memoria de trabajo de la complejidad computacional de imaginar el estado futuro de un sistema a partir del estado actual. También liberan a la memoria de la necesidad de aplicar algoritmos para predecir cómo responderá el sistema a sus inputs presentes y esperados.

Tipicamente, los displays predictivos trabajan de acuerdo a uno de tres principios. Si le es posible al sistema conocer actualmente los inputs a los cuales debe responder el humano, de forma anticipada, entonces estos inputs pueden ser mostrados al operador. En aquellas situaciones en las que no es posible conocer de antemano el estado futuro del sistema, debido a la influencia de múltiples variables ambientales, este puede ser inferido o estimado mediante la simulación de la dinámica del sistema (Sheridan, 1981). También se pueden utilizar técnicas predictivas de análisis, como por ejemplo la regresión (Jensen, 1981).

4.2.1.4. Utilización del color. Efectos de la redundancia

La posibilidad de implementar gráficos de color en los displays ha fomentado el interés por la efectividad del color como un medio de codificación de la información. El interés principal ha sido investigar si el efecto redundante del color mejoraba la ejecución (Christ, 1975, 1983; Christ y Corso, 1983; Thackray y Touchstone, 1991). La codificación redundante del color hace referencia a los casos

en que las variables o estímulos son codificadas por su color y su forma, pero pueden ser totalmente identificadas solamente por el color o sólo por la forma (p.e., cuando todos los círculos son rojos y todas las figuras rojas son círculos).

En general, no se ha encontrado que la condición redundante sea más eficaz que la no redundante (Christ, 1975; Jubis y Turner, 1988; Thackray y Touchstone, 1991).

También se ha propuesto como alternativa una condición de color parcialmente redundante (p.e., los estímulos difieren tanto en color como en su forma, pero sólo la forma es totalmente definitoria de los mismos). Najjar, Patterson y Corso (1982) compararon los efectos del color parcialmente redundante con el color redundante, y no encontraron diferencias sobre la ejecución. Jubis (1990) comparó cuatro condiciones de codificación (forma, color redundante, color parcialmente redundante, y sólo color) en dos tareas, una de identificación y otra de búsqueda, bajo diferentes niveles de carga de inspección (número de estímulos a ser identificados) y de densidad del display (número de estímulos presentados). En general, los resultados de este experimento sugieren que las condiciones de color y color redundante producen una mejor ejecución que la codificación por la forma en tareas de búsqueda, pero no en tareas de identificación (Davidoff, 1987; Luder y Barber, 1984; Zwaga y Duijnhouwer, 1984). Además, este efecto fue más pronunciado cuando aumentaba la densidad de los displays (Carter, 1982; Christ, 1975; Jubis y Turner, 1988; Luder y Barber, 1984; Noble y Sanders, 1980; Treisman y Gelade, 1980).

Sin embargo, y a pesar de estos resultados, muy pocas ventajas objetivas del color han sido encontradas en la investigación experimental en situaciones de tarea más complejas que la búsqueda visual (Hale y Billmeyer, 1988).

4.2.2. Modelo Interno

En los últimos años, un gran área de investigación empírica, cada vez más creciente, se ha focalizado en la manera en que los modelos mentales de los usuarios/operadores de sistemas simples y complejos influye en sus interacciones con dichos sistemas (Adelson, 1981; Eberts y Schneider, 1985; Halasz y Moran, 1983; Kay y Black, 1985; Mayer, 1975; Norman, 1983b).

Los modelos mentales del sistema facilitan el aprendizaje, la retención de procedimientos y la intervención (Halasz y Moran, 1983; McKeithen y Reitman, 1981; Jagcinski y Miller, 1978). En el proceso de control es tan importante la información mostrada al operador como el conocimiento que este tiene del sistema (Curry, 1981; Ferguson, 1980). Además, la detección de anomalías y fallos en un sistema complejo, realista, requiere que el operador utilice una gran variedad de "pistas" (olores, luces, tonos, etc.) y una representación interna o modelo muy compleja del sistema (Curry, 1981).

El término "modelo mental" puede tomar un significado diferente para cada autor (Eberts, 1987). De forma simple, el concepto de modelo mental o modelo interno hace referencia a la concepción que tiene el operador del funcionamiento del sistema (Wickens y Kramer, 1985). Una definición más completa, y que en general recoge la concepción de autores como Norman (1983, 1986) y Young (1983) podría ser la siguiente: un modelo mental es una representación, formada por el operador, de un sistema y/o tarea, basada en la experiencia previa y en las observaciones presentes, la cual proporciona la mayoría de su entendimiento o comprensión posterior del sistema, y, por tanto, dicta el nivel de ejecución en la tarea.

Rouse y Morris (1986) definieron los modelos mentales como:

"los mecanismos por los cuales los humanos son capaces de generar descripciones de la finalidad y la forma del sistema, explicaciones del funcionamiento y de los estados del sistema observados, y predicciones de los estados futuros del mismo" (p. 351).

Para Rasmussen (1979, 1986) los modelos mentales son representaciones internas de los factores ambientales que determinan las interrelaciones entre los datos observables.

A diferentes niveles de abstracción, Rasmussen (1979, 1986) consideró que diferentes modelos mentales podían ser construidos a partir de la forma física, la función física, la estructura funcional, la función abstracta y el significado funcional del sistema. Desde el mismo área de trabajo que Rasmussen, Rouse y Morris (1986) opinan que se pueden desarrollar diferentes modelos mentales en función de la propuesta del sistema, su función, su estado, o su forma. Sin embargo, Carroll y Olson (1987), a partir de su trabajo sobre la interacción hombre-computador, consideran que sólo es útil para la ejecución la información sobre la función y el estado del sistema, por lo que no toman en cuenta la información relativa a la propuesta o a la forma del sistema.

En el campo de la Ergonomía, los modelos mentales son usualmente concebidos como una forma de representación analógica. Rouse y Morris (1986) opinan que es razonable sugerir que los modelos mentales son frecuentemente pictóricos o similares a una imagen. Otro autor, Lindgaard (1987) piensa que el rango de las representaciones internas parte de dibujos mentales de objetos concretos y llega hasta concepciones e inferencias abstractas. Rasmussen (1986) concibe las representaciones del sistema, en su forma interna como un modelo mental, como visualizaciones del sistema físico o como modelos más abstractos de procesamiento de datos. Los ergónomos tienden, dada la naturaleza visualizable de muchos de sus

sistemas de interés, a ver los modelos mentales en términos de representaciones analógicas y como formas gráficas mas que simbólicas de las mismas (para una discusión sobre la concepción del modelo mental en Ergonomía y en la Psicología Cognitiva ver Wilson y Rutherford, 1989).

El concepto de modelo interno ha sido tomado en cuenta en numerosas perspectivas (Gentner y Stevens, 1983), y es la base de los modelos de control manual óptimo (Pew y Baron, 1978), de control de navegación (Veldhuyzen y Stassen, 1977), y de control de procesos (Rasmussen, 1983, 1986), entre otros campos.

El concepto de modelo interno puede guiar el diseño de programas de entrenamiento del operador y de mejora de displays. Los programas industriales y el diseño de displays deberían ser realizados de forma que fueran compatibles con el modelo interno del operador, de manera que la respuesta del sistema sea la esperada a partir de dicho modelo (Roscoe, 1968; Riley, 1986; Greeno, 1978; Greeno y Simon, 1984).

El concepto de modelo mental también ha sido utilizado para describir cómo los operadores controlan sistemas dinámicos (Jagcinski y Miller, 1978) y cómo detectan fallos en sistemas dinámicos (Gai y Curry, 1976; Wickens y Kessel, 1981). El modelo de Wickens y Kessel (1981) asume que el modelo interno del sistema que mantiene mentalmente el operador consiste de un conjunto de resultados esperados del sistema para conocer los inputs del sistema, dado que el sistema está operando con normalidad. La monitorización efectiva implicaría la constante comparación entre los resultados obtenidos y los esperados en función de los inputs observados. Si existen discrepancias entre los resultados obtenidos y los esperados, estas se almacenan y son acumuladas. Si esta acumulación de discrepancias excede algún criterio interno, se ha detectado un fallo. Según esta concepción, la latencia de la detección está influida por factores como los siguientes:

(1) El establecimiento del criterio. Si el criterio es bajo, la detección será rápida, pero se producirán más falsas alarmas. Así, la variación en el

establecimiento del criterio dicta qué aspectos de la eficacia de la ejecución (rapidez-precisión) deben ser evaluados.

(2) La fidelidad del modelo interno. Esto hace referencia al rango de posibles resultados esperados para un determinado input observado. Claramente, las diferencias entre los resultados esperados y los observados serán acumuladas más rápidamente si este rango se ve reducido por una mayor fidelidad del modelo.

(3) El número de canales de input disponibles a partir de las variables o displays del sistema observadas frente a la estimación interna del estado actual del sistema. Cuantos más canales estén disponibles (más información sea mostrada), más facilitada se verá la detección.

(4) Los recursos de procesamiento o atención asignada a los procesos de acumulación y detección. Algunos de estos factores serán expuestos con más detalle en los apartados posteriores.

Brehmer (1987) ha atribuido el creciente interés en los modelos mentales al cambio en la moderna tecnología, y Hollnagel (1988) ha arguido que este interés ha surgido de la necesidad de tener una metáfora apropiada que permita describir al operador moderno. La aceptación de la existencia y valor del concepto de modelos mentales puede encontrarse en gran parte de la literatura relevante (p.e., Brehmer, 1987; de Kleer y Brown, 1981; Moray, 1987; Hanisch, Kramer y Hulin, 1991). Muchas de las posiciones de los investigadores pueden ser resumidas diciendo que en la operación de sistemas,

"en el dominio basado en conocimiento (*knowledge-based*), la planificación de la interacción con el ambiente depende del conocimiento estructural, y de las representaciones mentales de las configuraciones estructurales de elementos y sus relaciones funcionales [modelos mentales]" (Rasmussen, 1986, p. 118).

4.2.3. Estrategias de muestreo y búsqueda de información

En general, la tarea de diagnóstico consiste en una búsqueda para identificar un cambio en el funcionamiento normal de un sistema. Varias estrategias diferentes son posibles para esta búsqueda y el dominio en el cual se realiza depende de la situación específica y de los objetivos inmediatos del operador (Rasmussen, 1978, 1981; Morris y Rouse, 1985). La estrategia óptima depende de la situación concreta.

Básicamente, la búsqueda de información puede realizarse de dos diferentes formas: a) búsqueda sintomática (a partir de los síntomas) y b) búsqueda topográfica (a partir de la localización de cada componente en la estructura del sistema) (Rasmussen, 1981). La diferencia principal entre estas dos clases de búsqueda está en el diferente uso de la información observada. Cada observación implica la identificación de una fuente de información y la lectura del contenido del mensaje. En la búsqueda sintomática, las referencias a la identidad del estado del sistema se obtienen a partir del mensaje leído; en la topográfica, esta referencia se obtiene a partir de la localización de la fuente de información, mientras que los mensajes están sujetos a juicios de bueno/malo que son utilizados para el control táctico de la búsqueda.

A continuación se describen brevemente ambos tipos de estrategias.

4.2.3.1. Búsqueda topográfica

Esta búsqueda se realiza a partir de una adaptación bueno/malo del sistema por la cual la extensión de campo potencialmente "malo" se va gradualmente reduciendo hasta que la localización del cambio está determinada con suficiente certeza para permitir la selección de la acción apropiada (se va reduciendo el campo de atención).

El dominio en el que esta búsqueda se realiza puede variar. La búsqueda puede ser realizada directamente en el sistema físico, pero en la mayoría de los casos, esta es una operación mental a un nivel de abstracción que depende de la finalidad e intención inmediata del operador, de la forma del modelo disponible y de las circunstancias particulares.

Los elementos principales de la estrategia son: el **modelo** del sistema utilizado para estructurar la búsqueda, la clase de **datos** utilizados para representar el estado del sistema, y las **reglas** tácticas utilizadas para controlar la secuencia de búsqueda.

La búsqueda depende de una representación gráfica o mapa del sistema que informa sobre la localización de las fuentes potenciales de observación para las cuales la *información de referencia* (la relativa al funcionamiento normal o adecuado del sistema) está disponible. El mapa es un modelo que permite identificar estas fuentes de observación potencial en términos de la topología del sistema físico, de su estructura interna anatómica o funcional, o de sus propuestas externas o resultados esperados del sistema. La secuencia de búsqueda está basada en un conjunto de reglas o heurísticos que sirven para limitar el campo de atención. Si los resultados del sistema están relacionados con distintas partes o subsistemas internos, el juicio bueno/malo de estos resultados permitiría identificar eficazmente el campo interno en el que se debe realizar la búsqueda. Además, si se ha identificado una relación input/output no adecuada, debería encontrarse la ruta causal relacionada con dicha función input/output. Por tanto, las decisiones de búsqueda táctica estarán basadas exclusivamente en la información obtenida a partir del juicio de tipo bueno/malo de una sola observación. La información relativa a la posible naturaleza del fallo será utilizada por otras estrategias de búsqueda (que serán discutidas posteriormente) que serán frecuentemente utilizadas para guiar la búsqueda topográfica.

En general, se pueden distinguir dos formas diferentes de localización del cambio en el estado normal del sistema. Una en función de las partes o subsistemas en las que se encuentra el cambio, y otra en función de un patrón o vía causal. En ambos casos, la búsqueda puede estar basada en juicios de tipo bueno/malo, bien de

la magnitud de las variables de estado, o bien a cerca de sus mutuas relaciones. Cuando la búsqueda está basada en datos de referencia (valores de las variables cuando el sistema funciona con normalidad) esta se realiza como una secuencia de juicios bueno/malo de estas variables individualmente. Esta secuencia puede llevarse a cabo trazando la anomalía a lo largo de un patrón causal, o buscando un estado anormal. En ambos casos, el sistema debe corresponderse con el modelo de referencia disponible. Sin embargo, se podría obtener una búsqueda más eficaz si el sistema puede ser dividido por una secuencia de estados de prueba que afectan a diferentes partes del sistema, establecidas por unas combinaciones cuidadosamente seleccionadas, y para las que se pueden preparar modelos de referencia. La administración y evaluación de esta secuencia de pruebas depende de los argumentos lógicos, combinatorios, obtenidos de una efectiva memoria a corto plazo.

La estrategia de búsqueda basada en juicios simples sobre la magnitud de cada variable de estado individualmente, puede ser efectiva, sin embargo, las decisiones tácticas dependerán de supuestos sobre la dirección de la causalidad. Por esto, las estrategias basadas en juicios sobre las relaciones, que representan las propiedades del sistema, son superiores en eficacia.

Las reglas tácticas controlan el uso de un determinado mapa del sistema. Este mapa puede mostrar la estructura del sistema a distintos niveles de abstracción (Rasmussen, 1980) y el nivel apropiado para la búsqueda depende de los objetivos y las intenciones del operador que tiene que emitir el diagnóstico. Sin embargo, el nivel de abstracción más efectivo puede también cambiar durante la búsqueda, p.e. comenzando haciendo juicios sobre el conjunto del sistema, después pasando a juicios sobre el funcionamiento de cada parte componente, y acabando la búsqueda atendiendo a los elementos de que se compone cada parte del sistema (Lind, 1981).

La información disponible en las observaciones es utilizada de forma antieconómica por la búsqueda topográfica, ya que este uso depende solamente de juicios de tipo bueno/malo. Además esta estrategia no toma en cuenta el conocimiento que puede proporcionar la experiencia del operador con otros fallos anteriores. Sin

embargo, la búsqueda topográfica tiene la ventaja de depender de un modelo sobre el funcionamiento normal del sistema, el cual puede ser desarrollado durante la fase de diseño del sistema, o puede ser obtenido a partir de la observación del sistema durante una situación de operación normal. Esto permite verificar la consistencia y el grado de corrección de la estrategia: ya que no depende de modelos sobre el mal funcionamiento, estará menos influido por las anomalías múltiples o desconocidas que las estrategias basadas en síntomas.

4.2.3.2. Búsqueda sintomática

Las estrategias de búsqueda sintomática están basadas en la información contenida en las observaciones para identificar el estado del sistema, en vez de en la localización de la fuente de información (búsqueda topográfica). Las decisiones de búsqueda derivan de las relaciones internas de los datos y no de la estructura topológica de las propiedades del sistema.

En principio, la búsqueda se realiza a través de uno de los grupos de datos anormales, síntomas, para encontrar aquel conjunto que se adapta mejor al patrón de conducta actual del sistema. Los patrones de referencia pueden ser recogidos empíricamente a partir de incidentes del sistema o derivados del análisis o la *simulación de las respuestas del sistema ante determinadas anomalías*. Además, los patrones de referencia pueden ser generados durante el proceso de búsqueda, si el operador tiene disponible un modelo funcional que pueda ser modificado para que tome en cuenta la hipótesis actual sobre el fallo del sistema.

Dependiendo de la estructura y memoria del operador, la búsqueda puede ser paralela, mediante el **reconocimiento de patrones**, o secuencial, mediante **tablas de decisión**.

El reconocimiento de patrones juega un papel fundamental en el diagnóstico humano; puede identificar eficazmente estados familiares del sistema directamente, pero también puede ser utilizado frecuentemente, por ejemplo, durante la búsqueda topográfica para guiar las decisiones tácticas. Los reconocimientos suelen estar basados en síntomas de referencia generales o difusos en términos de patrones de fallo genéricos referidos a determinados tipos de funciones o partes físicas, como características de ruido, inestabilidad o formas de no linealidad.

La búsqueda por tablas de decisión depende de un conjunto de reglas tácticas que guían la búsqueda y que pueden estar basadas en la probabilidad de ocurrencia, la estructura jerárquica de los atributos, o las relaciones funcionales, almacenadas como árboles de fallo, etc. Los operadores humanos podrían usar tablas de decisión para la verificación de reconocimientos más ambiguos. Las tablas de decisión han sido utilizadas en la monitorización de plantas por Berenblut y col. (1977) y Lihou (1981).

Si una búsqueda está basada en patrones de referencia generados durante la tarea (*on-line*) por la modificación de un modelo funcional en correspondencia con un cierto problema, la estrategia se denomina búsqueda por hipótesis y prueba. La eficacia de esta búsqueda depende de las tácticas de generación de hipótesis. Generalmente, en el diagnóstico humano, las hipótesis resultan de búsquedas topográficas inciertas o de reconocimientos difusos. Hay evidencia que sugiere que las personas pueden tener dificultad para generar hipótesis precisas y adecuadas. Por ejemplo, Mehle (1980) y Fisher (1980) encontraron que los sujetos generaban listas de hipótesis incompletas, y además, los sujetos decían que sus listas eran más completas de lo que en realidad eran (lo sobreestimaban).

La búsqueda por síntomas tiene ventajas desde el punto de vista de la economía de la información, pero tiene la desventaja de necesitar un patrón de referencia del estado anormal concreto, siendo necesario considerar múltiples fallos y anomalías posibles del sistema.

La carga mental, al utilizar un modelo del estado anormal del sistema para generar un conjunto de síntomas para probar una hipótesis, es tan alta que muy frecuentemente el operador humano preferirá probar la hipótesis haciendo correcciones en el sistema de acuerdo a dicha hipótesis, y así probar si el patrón de respuesta del sistema cambia a su estado normal. Esta estrategia es generalmente efectiva en la solución de problemas menores, sin embargo, no será efectiva cuando los problemas o fallos del sistema son mayores, en cuyo caso debería considerarse la ayuda por computador al operador durante la prueba de hipótesis.

Las diferencias intrínsecas entre las diversas estrategias de diagnóstico y el efecto de los cambios de estrategia utilizadas por los sujetos, ha sido un tema típico en los experimentos de diagnóstico. Por ejemplo, Shepherd y col. (1977) estudiaron los efectos de diferentes formas de entrenamiento para el diagnóstico y encontraron que los operadores entrenados por las reglas obtenidas de operadores expertos fueron superiores en el diagnóstico de fallos con los que no se habían encontrado antes, comparados con los operadores que habían sido entrenados en teoría de la planta. Estos, en cambio, fueron superiores a los operadores entrenados practicando diagnósticos. Estas diferencias pueden ser fácilmente explicadas ya que los diferentes métodos de entrenamiento implican el uso de diferentes estrategias.

La búsqueda topográfica en estructuras de flujo abstractas es muy similar a la búsqueda basada en reglas en las redes libres de contexto descritas por Rouse y col. (1981). Como Rouse (1981) nota, el modelo basado en reglas describe razonablemente bien las estrategias libres de contexto, pero no lo hace tan bien cuando las tareas son específicas de contexto. Esto probablemente se deba a que el contexto inicie un cambio en la estrategia hacia una búsqueda por síntomas, la cual depende de las experiencias previas individuales de los sujetos.

En las salas de control futuras, la tarea de diagnóstico será realizada por una compleja interacción entre los computadores y los operadores. La solución al problema de asignación de las funciones de procesamiento de datos al computador y al operador debe ser resuelto por el análisis cuidadoso de la carga mental impuesta a los

operadores. En el caso en el que el sistema no funciona adecuadamente, el trabajo del operador implica un ajuste multidimensional entre los recursos y las demandas. La carga mental del operador dependerá en gran medida de las posibilidades que tenga el operador para resolver los conflictos entre demandas y recursos, mediante la selección de una estrategia apropiada a la tarea. La tabla 4.1 muestra los factores que influyen en la elección de una determinada estrategia. Las estrategias son frecuentemente complementarias en diversos aspectos.

El criterio de ejecución es claramente un factor complejo que no depende únicamente de aspectos como la carga mental o la cantidad de información mostrada. El criterio subyacente a la elección de una estrategia en una determinada situación será subjetivo, ya que los recursos y sus valores asignados variarán para cada individuo. Además, factores emocionales como la curiosidad, el gusto por el riesgo, etc., pueden influir en esta elección. En personas muy capacitadas, con altas habilidades, existe una relación entre las diferentes situaciones de tarea y la estrategia utilizada en cada una de ellas (Rasmussen y Jensen, 1974; Rasmussen, 1981).

FACTORES	ESTRATEGIAS	Búsqueda Topográfica	Reconocimiento	Tabla de Decisión	Hipótesis y prueba
Tiempo consumido		--	baja	--	--
Número de observaciones		alta	baja	--	baja
Dependencia de la percepción del paciente		--	alta	--	--
Carga en la memoria a corto plazo		baja	baja	alta	alta
Complejidad de los procesos cognitivos		baja	baja	--	alta
Complejidad del medio funcional		baja	--	--	alta
Aplicabilidad general de las reglas tácticas		alta	--	--	baja
Dependencia experiencia de mal funcionamiento		baja	alta	--	baja
Dependencia preanálisis de mal funcionamiento		--	--	alta	--

Tabla 4.1. Dependencia de diferentes factores en la elección de una u otra estrategia de diagnóstico (Rasmussen, 1981).

El criterio de ejecución que guía las elecciones humanas no es observable, tiene que ser inferido a partir del análisis de la ejecución. Rasmussen (1981) menciona algunos de los principios encontrados en estudios sobre estrategias (Rasmussen y Jensen, 1974) y en informes sobre errores ocurridos en plantas (Rasmussen, 1979). La regla principal que subyace a las estrategias parece ser elegir la vía de menor resistencia. En lugar de hacer planes generales de búsqueda, la tendencia de los sujetos es tomar rápidas e impulsivas decisiones basadas en la información observada en cada momento. Es decir, no se tiende a la recapitulación de los resultados previos ni a la consideración de argumentos causales o funcionalmente más complejos. Al mismo tiempo, hay un "punto de no retorno" en la atención aplicada en el instante de la decisión. Es decir, la información puede ser observada de nuevo después de haber tomado una decisión sin que esto tenga ningún efecto, aunque esta información *contradiga claramente a la decisión tomada. Estas tendencias suponen estrategias que demandan inferencias a partir de varias observaciones muy poco fiables si los datos están disponibles de forma secuencial, ya que frecuentemente darán referencias de diagnóstico individualmente. Este mismo efecto se observa en informes de accidentes ocurridos. En situaciones complejas anormales, un conjunto de indicaciones anormales tiende a ser interpretado como una coincidencia de ocurrencias familiares, más que como un patrón relacionado con un estado complejo del sistema o de la planta desconocido (Rasmussen, 1979).*

El criterio que subyace a la elección de la "vía de menor resistencia" es minimizar la carga de la memoria a corto plazo. Para evitar la carga impuesta sobre la memoria a corto plazo, se pueden dar las siguientes recomendaciones para el diseño del sistema: presentación simultánea de la información que debería ser utilizada junta, relevar al operador de tareas secundarias como la memorización de la información anatómica o funcional del sistema, conversión de los datos al dominio de la búsqueda, generación de datos de referencia, etc.

La tarea de diagnóstico debería ser realizada por ambos, computador y operador. El operador debe aceptar la ayuda que el computador le proporciona, para lo cual debe ser capaz de entender las estrategias que este sigue. Además el

computador debe ser capaz de adaptarse a la estrategia de búsqueda preferida por el operador y de presentarle la información al nivel de abstracción que este solicite (Goodstein y Rasmussen, 1980; Goodstein, 1981; Rasmussen, 1981).

Por otro lado, hay evidencia de que las personas tienden a utilizar la información positiva (aquella relativa a lo que falla), pero no toman en cuenta la información negativa (información sobre lo que no ha fallado), cuando trata de encontrar un fallo en el sistema (Rouse, 1978).

Los métodos de búsqueda de información, el criterio de selección de las fuentes, y el valor de la información varían en función del tiempo (Rouse y Rouse, 1984), a medida que se realiza la tarea.

En la medida en que la frecuencia de muestreo requerida de una fuente o fuentes aumenta, así incrementa la probabilidad de que la atención se vea sobrecargada, y cuanto más sobrecargada esté la atención menos posible será que se observe una variable cuyo valor es no normal. Moray propone como un modelo plausible para el proceso de adquisición de datos, la teoría de la decisión secuencial (Kvalseth, 1980; Moray, 1980). Según este modelo se asumiría que el operador establece criterios en función de probabilidades y utilidades a priori para cada observación, y continua viendo una información hasta que haya acumulado suficiente evidencia para satisfacer el criterio, y tomar una decisión como por ejemplo si el sistema funciona adecuadamente o no. Esta teoría también ha sido utilizada por Gai y Curry (1978) en combinación con la teoría de la decisión óptima para modelar la detección de fallos con considerable éxito.

Hay abundante evidencia de que la atención puede afectar a la fuerza de las señales que entran en el sistema nervioso y al criterio de respuesta del observador (Moray y col., 1976). Es decir, las expectativas a priori de las personas determinan o tienen un gran efecto sobre la información a la que se le presta atención, y esto determina su respuesta. La evidencia que es esperada (preconcepciones) por el operador es rápidamente aceptada como evidencia conclusiva, y aquella que no es

esperada tiende a no ser tomada en cuenta. La resistencia de los observadores humanos a aceptar evidencia contraria a sus hipótesis preconcebidas está relacionada con los sesgos de respuesta (Taylor, 1975; Gai y Curry, 1978; Moray, 1980). Parece que los humanos quieren confirmar, más que explorar, la precisión de sus hipótesis.

4.2.4. Complejidad del Sistema

Los efectos de la complejidad del sistema sobre el nivel de ejecución en tareas de detección y diagnóstico de fallos ha sido estudiado por diversos autores, encontrando la mayoría de ellos efectos muy claros de este factor (p.e., Rouse, 1978, 1979; Wohl, 1981, 1982; Brooke y Duncan, 1981).

Por ejemplo, Rouse (1981) utilizó una serie de tareas en las que la estructura del sistema estaba representada por medio de redes lógicas, en las que el sujeto debía encontrar un fallo. La complejidad de cada red era principalmente función del número de unidades que la formaban (25 y 49), y del tipo y número de operadores lógicos (Y y/o O) que entraban en juego. Se realizaron diversos experimentos con estas tareas, que permitieron concluir que la ejecución era peor a medida que aumentaba el tamaño del problema, es decir aumentaba el número de unidades que formaban la red (Brooke y Duncan, 1981). Los sujetos parecían mostrar una incapacidad para utilizar la información relativa a lo que no fallaba que se hacía más aparente a medida que la complejidad del sistema era incrementada. Este efecto se puso de manifiesto en el aumento de pruebas del sistema que realizaron los sujetos y en el tiempo que estos necesitaron para emitir el diagnóstico. Una de las conclusiones generales que se encontró en la investigación, fue que el tiempo requerido para resolver el problema de diagnóstico del fallo, en términos de la complejidad del problema, está relacionado con el nivel de entendimiento que tiene el sujeto del sistema, además de con las propiedades intrínsecas del problema.

Efectos similares de la complejidad del sistema han sido encontrados por Brooke y Duncan (1981), McDonald y col. (1983) o, menos recientemente, Goldbeck, Bernstein, Hillix y Marx (1957). McDonald y col. (1983) encontraron incrementos en el tiempo requerido para realizar la tarea y en el número de pruebas realizadas cuando aumentaba la complejidad del sistema.

Además, la complejidad del sistema influye en el tiempo necesario para realizar el diagnóstico (Wohl, 1981). En su investigación, Wohl (1981), utilizando como sistemas diversos circuitos electrónicos (muy similares a las redes del estudio de Rouse (1981)), propuso como hipótesis inicial que la forma de la distribución de los tiempos medios de diagnóstico estaban determinados por los siguientes aspectos (prácticamente todos ellos relativos a la complejidad del sistema):

- a. El número de líneas, entradas o salidas, directamente asociadas con el componente defectuoso.
- b. La accesibilidad de los puntos de unión eléctrica como puntos de prueba o comprobación.
- c. El número total de componentes directamente conectados a todos los puntos de prueba que tuvieran lecturas o valores fuera del rango de tolerancia (p.e., el conjunto de componentes sospechosos).
- d. La capacidad de inferencia lógica del técnico.

Además, otra de sus hipótesis iniciales fue que el número de pasos de búsqueda de información requerido para emitir el diagnóstico, estaba determinado por:

- a. Densidad de líneas de unión (DLU): La frecuencia de componentes con un determinado número de entradas o salidas (p.e., número de componentes con N entradas).

- b. *Densidad de puntos de unión de las líneas (DPUL): La frecuencia de puntos de unión con un determinado número de componentes implicado (p.e., el número de puntos de unión que van a dar a M componentes).*

Con estas hipótesis y otras en mente, se estableció como un índice de la *complejidad del circuito (Ω)*, la siguiente ecuación:

$$I = N M$$

donde N es la media de DLU, y M es la media de DPUL.

Este índice simplemente representa el número esperado de componentes conectados con un componente determinado. Además se establece el índice lambda, definido como el valor esperado del grado de diagnóstico ($\lambda = 1/T$; donde T es el tiempo medio de prueba para cada componente).

La reducción del índice de complejidad puede tener un fuerte impacto en el tiempo de ejecución. Los valores de estos dos índices pueden establecerse a la hora de elegir entre diseños alternativos, y además, el nivel y tipo de entrenamiento tienen efectos diferenciales sobre estos índices.

4.2.5. Nivel de Automatización del Sistema

Debido a los grandes y rápidos avances tecnológicos que han hecho posible que los sistemas tengan altos niveles de automatización, muchos autores se han planteado cual debe ser el papel del operador humano en los sistemas hombre-máquina actuales. En este sentido, diversos autores (p.e., Ephrath y Young, 1981; Kessel y Wickens, 1978; Walden y Rouse, 1978; Wickens y Kessel, 1979, 1981) han investigado las diferencias en rapidez y precisión de la detección y el diagnóstico de fallos en función del tipo de control del sistema (manual vs. automático) y de la carga (medida por el procedimiento de tarea secundaria).

Por ejemplo, Ephrath y Young (1981) encuentran que los mayores niveles de carga asociados con el control manual producen una peor ejecución de diagnóstico cuando el control es de tipo manual que cuando es automático. De igual forma, Wickens y Kessel (1979), utilizando niveles de carga intermedios, encuentran que la carga que sufre el operador afecta de forma adversa a la calidad de la detección de fallos. Ellos encuentran, en principio, una superioridad global de la ejecución de detección de fallos cuando el sistema es manual, sin embargo, encuentran que mientras la precisión de la detección es mejor en condiciones de control automático, la rapidez de la detección es mejor en los sistemas manuales.

Estos resultados parecen indicar que el nivel de carga asociado con una situación puede dictar el modo de participación del operador (manual-automático) preferido. En tareas de detección de fallos que suponen bajos niveles de carga la ejecución será superior cuando el sistema es controlado manualmente, mientras que cuando los niveles de carga son altos la ejecución es mejor cuando el sistema es automático.

Además, Wickens y Kessel (1979), encontraron que durante las fases de entrenamiento de los operadores, estos aprendían más y mejor como detectar y diagnosticar fallos cuando el entrenamiento se realizaba con un sistema manual que con su versión automatizada. Realizaron un nuevo experimento en el que se puso de manifiesto que los sujetos que operaban con un sistema controlado automáticamente, pero que habían sido entrenados en condiciones de control manual, mostraban una ejecución mejor que los sujetos de los restantes grupos (con entrenamiento en sistema automático y realización en sistema automático (AU-AU), con entrenamiento en sistema manual y realización en sistema manual (MA-MA) y con entrenamiento en sistema de control automático y ejecución en sistema manual (AU-MA)). Solamente en el grupo MA-AU se observó esta transferencia positiva, que además resultó ser esencialmente perceptiva y atribuible a los requisitos que el control manual efectivo impone al operador cuando tiene que extraer información relevante de displays visuales.

4.2.6. Relaciones de Correlación y Causalidad entre las Variables de Estado del Sistema

En general, un sistema complejo no está uniformemente interconectado. Según Moray (1981), el sistema total puede ser visto como un conjunto de subsistemas. En el nivel más simple, cada variable individual está muy relacionada con un pequeño conjunto de otras variables, y muy poco relacionada con las demás. A partir de las correlaciones entre variables y de sus conexiones causales, el sistema puede ser visto como un conjunto de variables "atómicas", que formarían parte a su vez de "moléculas", dentro de las cuales las intercorrelaciones son muy altas.

Esta estructura sugiere ciertas tácticas y estrategias óptimas que se pueden utilizar para examinar el sistema total. Podemos partir de dos situaciones:

a) Asumimos que el operador ha realizado una serie de observaciones que le hacen creer que el sistema está funcionando con normalidad, y que simplemente el sujeto ha completado una observación. ¿Qué variables serán las que debe observar a continuación?. En este caso, el operador no debería observar una variable altamente correlacionada con la anterior, sino que debería observar una variable perteneciente a otra "molécula". Esto se debe a que la alta correlación ya le da considerable información basada en la observación que acaba de realizar. El puede predecir el valor de las variables altamente correlacionadas con una probabilidad a priori a partir de la observación que ya ha realizado, y observar estas variables no reducirá su incertidumbre. Por tanto, la observación de una variable muy relacionada con la que se acaba de observar es antieconómico. Por tanto, cuando el funcionamiento del sistema parece normal, las observaciones sucesivas *deberían realizarse para las variables muy poco correlacionadas con la anterior.*

b) Si, por otro lado, la última observación realizada por el operador le indica alguna anomalía, este debería examinar los valores de aquellas variables correlacionadas con la anterior, con el fin de diagnosticar la causa de dicha anomalía. Si ninguna de las posteriores observaciones indica una anomalía, entonces se deduce que esa, la primera variable en que se encontró un fallo, debe ser la causa del mismo. En el caso contrario se deben seguir examinando variables altamente correlacionadas.

Esta regla última tendrá el efecto de que el operador preste atención solamente a un pequeño conjunto de variables, las variables relevantes, hasta que emita un diagnóstico.

4.2.7. Características de la Tarea

Las fuentes de información seleccionadas por las personas están muy relacionadas con el tipo de tarea (Bassili y Regan, 1977). Es decir, dado un determinado problema, el esquema general de la tarea puede afectar a la elección de determinados métodos de búsqueda de información.

En sus experimentos, Rouse (1981) utilizó dos tipos de tareas, unas libres de contexto (tareas 1 y 2) y otra específica de contexto (tarea 3). Las dos primeras fueron libres de contexto ya que no estaban asociadas a ningún sistema o pieza de equipo en particular, y además los sujetos no se enfrentaban en ningún momento con el mismo problema dos veces, de forma que no podían desarrollar habilidades particulares para un problema. La tercera tarea era específica del contexto, ya que se trataba de un sistema de circuito electrónico real, en el que aparecían etiquetas en cada unidad que hacían referencia al nombre de esa pieza del equipo (turbina, compresor, sistema de lubricación, etc.).

Como resultado general se encontró que la ejecución en tareas libres de contexto está altamente correlacionada con la ejecución en tareas familiares específicas del contexto.

Además, las estrategias de búsqueda de información utilizadas fueron diferentes. En las tareas libres de contexto, los sujetos utilizaron principalmente una estrategia topográfica, mientras que cuando se trataba de una tarea de una tarea específica de contexto, la estrategia utilizada fue la sintomática.

4.2.8. Nivel y Tipo de Entrenamiento del Operador

En prácticamente todas las investigaciones se pone de manifiesto el efecto que el entrenamiento tiene sobre el nivel de ejecución. La incertidumbre y, por tanto, la necesidad de observar una mayor cantidad de información o de observar la misma información varias veces, disminuye con el entrenamiento (Bainbridge, 1978).

Además, investigaciones como la de Vessey (1985) encuentran diferencias en el tiempo necesitado para realizar la tarea por expertos y novatos (los sujetos expertos son más rápidos que los novatos). Sin embargo, los expertos tienden a invertir más tiempo que los novatos en aspectos de planificación y evaluación (Gugerty y Olson, 1986). Los expertos son más capaces de formarse una idea global del sistema, pero esto les supone invertir más cantidad de tiempo en desarrollar su comprensión y representación, y además, consideran más globalmente las interacciones entre los componentes del sistema (Adelson y col., 1984). Los sujetos expertos forman un modelo conceptual detallado del sistema y tienden a incorporar aspectos abstractos más que objetos concretos específicos en su representación del problema (Larkin, 1983). Sus modelos añaden múltiples niveles y son lo suficientemente ricos para soportar simulaciones mentales (Jeffries y col., 1981; Adelson y Soloway, 1985). Utilizando una tarea de discriminación de señales y de toma de decisiones, Bisseret (1981) encontró diferencias en ejecución en función del nivel y tipo de experiencia.

En los experimentos ya mencionados realizados por Rouse y sus colaboradores (Rouse, 1981), se encontró una transferencia positiva entre el entrenamiento en tareas libres de contexto y la ejecución en una tarea similar pero específica de contexto. Esta transferencia positiva del entrenamiento fue interpretada como un reordenamiento de prioridades dentro de un conjunto de reglas de solución del problema. Este reordenamiento de prioridades capacita a los sujetos para utilizar en mayor grado su conocimiento de la estructura del problema, y así llevar a cabo pruebas del funcionamiento de cada componente más eficaces para reducir la incertidumbre. Resultados similares fueron encontrados por Hunt y Rouse (1981).

Aunque la experiencia se ha mostrado como un importante factor que afecta a la detección y el diagnóstico de fallos del sistema, las diferencias en ejecución debidas a este factor son menos aparentes cuando se proporcionan guías para la solución del problema. En este sentido, Potter y Thomas (1976) encontraron que los técnicos de mantenimiento con una experiencia de seis meses o menor solucionaron menos problemas que el personal experimentado. Sin embargo, estas diferencias no fueron observadas cuando los problemas estaban acompañados de procedimientos detallados de solución del problema. Resultados similares fueron encontrados en el estudio de Elliot y Joyce (1971), en el cual estudiantes de bachillerato, utilizando una guía de solución del problema, fueron capaces de identificar fallos en equipos electrónicos tan eficazmente como los técnicos de las Fuerzas Aereas utilizando manuales tradicionales.

Según Zakay y Wooler (1984), la transferencia del entrenamiento en toma de decisiones desde condiciones normales hacia situaciones de presión temporal será muy pobre.

Por otro lado, el nivel de entrenamiento también afecta al modelo interno o representación que el sujeto forma del problema (Chi y col, 1981; Adelson, 1981, 1984; Weiser y Shertz, 1983; Larkin y col, 1980; Koubek y Salvendy, 1991). La diferencia cualitativa en la representación del problema entre expertos y novatos es la responsable de las diferencias en la ejecución de ambos grupos. Los modelos mentales

de los sujetos inexpertos o novatos son incompletos e imprecisos, y están formados por malas interpretaciones del sistema o sus componentes (Hanisch, Kramer y Hulin, 1991).

En un experimento muy reciente realizado por Koubek y Salvendy (1991), con el fin de encontrar diferencias en la ejecución cognitiva de dos grupos de sujetos, expertos y superexpertos, no se han encontrado diferencias en el tiempo necesitado por ambos grupos para resolver la tarea. Sin embargo, otras investigaciones han encontrado diferencias en los tiempos de tarea entre sujetos novatos y expertos (Vessey, 1985). Esto parece indicar que las diferencias en el tiempo necesitado para solucionar el problema se ponen de manifiesto entre grupos extremos, pero no es así cuando los grupos están más cercanos en sus habilidades (Henneman y Rouse, 1984).

Parece que las diferencias entre los niveles más altos de experiencia están más directamente relacionadas con el desarrollo de la representación inicial del problema. Sin embargo, y como uno de los más relevantes resultados de su investigación, si se encontraron diferencias en las estrategias de búsqueda de información utilizadas por los expertos y los superexpertos. Los sujetos superexpertos buscaron en primer lugar la información de tipo más general, mientras que los expertos buscaban información más específica relacionada con la tarea.

4.2.9. Tiempo Disponible

Diversas investigaciones sobre toma de decisiones han puesto de manifiesto que la presión temporal es un factor que influye en la efectividad de la decisión (Wright, 1974; Ben Zur y Breznitz, 1981; Einhorn, 1970; Zakay y Wooler, 1984).

El diagnóstico correcto de los problemas que tienen lugar en sistemas complejos requiere un considerable esfuerzo cognitivo y bastante cantidad de tiempo. Si el tiempo disponible es muy limitado, la probabilidad de que se produzca un error de diagnóstico aumentará (Miller y Swain, 1987).

Estudiando el tiempo de reparación para diversos sistemas, Wohl (1981) encontró que, por un lado, (1) la probabilidad de localizar el componente que falla en un sistema, aumenta con el tiempo, principalmente porque a medida que aumenta el tiempo disponible para emitir un diagnóstico el conocimiento del operador aumenta y se produce una rápida reducción de la ambigüedad; e, inversamente, (2) la probabilidad de localizar el componente que falla decrece con el tiempo disponible, posiblemente debido a la mayor dificultad para interpretar los síntomas y los valores de las variables de estado.

4.2.10. Diferencias Individuales

En la mayoría de las investigaciones sobre detección y diagnóstico de fallos y solución de problemas realizadas con multitud de sistemas diferentes, se ha encontrado un factor de diferencias individuales de gran importancia. Por ejemplo, investigaciones sobre solución de problemas en diversos sistemas realizadas en los años cincuenta (p.e., Saupe, 1954; Saltz y Moore, 1953; Moore, Saltz y Hoehn, 1955; Glaser y Phillips, 1954; Highland, Newman y Waller, 1956), comprobaron como los sujetos que podríamos llamar "malos solucionadores de problemas" (aquellos cuya ejecución es peor) realizaban chequeos del sistema que eran irrelevantes para el problema, y además omitían chequeos relevantes. Además los sujetos que encontraban y solucionaban los problemas de forma más eficaz, conocían más y mejor el funcionamiento del sistema, y utilizaban este conocimiento más adecuadamente. Más actualmente, Baldwin (1978), entre muchos otros, ha encontrado resultados similares.

La conclusión general de todos estos estudios es que los sujetos que mostraron una peor ejecución en el diagnóstico del problema del sistema, se caracterizaron por el uso inadecuado e incompleto de la información, la ineficaz generación y prueba de hipótesis sobre el posible problema, y por el uso de estrategias poco flexibles.

En general, se distinguen tres fuentes principales de diferencias individuales. Estas hacen referencia a las habilidades y aptitudes de los sujetos, a sus estilos cognitivos y a factores de personalidad.

4.2.10.1. Habilidades y Aptitudes

La conclusión general de la mayoría de los estudios realizados sobre este aspecto podría ser que de todas las habilidades, las que tienen una mayor relación con la ejecución en diagnóstico de problemas son aquellas relativas al conocimiento del trabajo específico. De esta forma, las medidas de aptitud general no son muy predictivas de la ejecución en solución de problemas (Highland y col., 1956, Demaree, Crowder y Morrison, 1955; Gallwey, 1982).

Además, es difícil encontrar diferencias en ejecución entre el grupo de sujetos con altas puntuaciones en tests o pruebas de habilidad y aptitud, y el grupo con puntuaciones medias. Por ejemplo, Henneman y Rouse (1984) encontraron que la ejecución en solución de problemas podría estar relacionada con aspectos de la habilidad de los sujetos solamente si algún nivel mínimo de esta no está presente; una vez superado este nivel mínimo, otros factores, como los estilos cognitivos, tienen más importancia.

Por otro lado, existe un extenso debate en cuanto a la existencia de una aptitud general de tiempo compartido. Esta aptitud hace referencia a la capacidad para realizar diversas tareas al mismo tiempo. Las investigaciones dirigidas a buscar esta aptitud han sido realizadas utilizando tareas duales de naturaleza diversa y con diferentes niveles de dificultad. La existencia de este factor general de recursos atencionales ha sido defendida principalmente por Hunt y colaboradores (Hunt, 1980; Hunt y Lansman, 1982; Hunt y Farr, 1984; Lansman y Hunt, 1983; Lansman y col, 1983; Sverko y col., 1983). Sin embargo, otros autores han criticado duramente la existencia de esta aptitud y la metodología utilizada en estas investigaciones (Ackerman,

Schneider y Wickens, 1984; Fogarty, 1987). Los resultados disponibles hasta el momento no permiten concluir que exista una aptitud general para el tiempo compartido. Esta capacidad para controlar la atención parece deberse tanto a un efecto específico de la modalidad del input como a una especie de aptitud general (Hunt, 1987). Como concluyen Ackerman y colaboradores (1984), las pruebas y estudios de la habilidad de tiempo compartido deben estar basadas en modelos explícitos y técnicas confirmatorias (Mulaik, 1972).

4.2.10.2. Estilos Cognitivos

Se ha encontrado que un número de medidas de estilos cognitivos correlacionaban con la ejecución en tareas de solución de problemas. Dos dimensiones de estilos cognitivos fueron las más discutidas en los estudios revisados. Estas dos dimensiones son dependencia-independencia de campo y flexibilidad-impulsividad.

Rouse y sus colaboradores (Henneman y Rouse, 1984; Hunt, Henneman y Rouse, 1981; Rouse y Hunt, 1984; Rouse y Rouse, 1979, 1982) analizaron los resultados de una investigación en la que los sujetos resolvían un número de tareas de solución de problemas, tanto en situaciones de simulación como en el equipo real. Los estilos cognitivos de los sujetos fueron medidos por dos tests: Emparejamiento de Figuras Familiares y Test de Figuras Ocultas. Los resultados encontrados fueron los siguientes:

- a) Los sujetos dependientes de campo fueron, inicialmente, más lentos en la solución de los problemas simulados, pero se tomaron más rápidos con la práctica, hasta un punto en el que no había diferencias entre los dependientes y los independientes de campo en cuanto al tiempo requerido para solucionar el problema.
- b) Los sujetos categorizados como impulsivos cometieron más errores que los reflexivos. Los impulsivos no mejoraron con la práctica.

c) Las medidas de habilidad (test ACT) no estuvieron significativamente correlacionadas con las medidas de ejecución, pero se encontraron coeficientes de regresión significativos para las medidas de habilidad y de estilos cognitivos cuando ambas eran combinadas en una ecuación de regresión para predecir la ejecución.

En esta misma línea, Seppala y Salvendy (1985) encontraron una relación altamente significativa entre el tiempo de ejecución y la puntuación en el Test de Figuras Ocultas, la cual aumenta cuando la búsqueda de información se realizaba mediante una estructura de menú jerárquica.

Así mismo, los resultados de Gallwey (1982) pusieron de manifiesto que el Test de figuras Ocultas era el mejor predictor de la ejecución en una tarea de inspección visual y detección de fallos de diversos tipos.

Variedad de estudios han intentado identificar las relaciones entre los estilos cognitivos y la conducta de búsqueda de información. Por ejemplo, los sujetos con altas puntuaciones en "apertura a la información" tienden a juzgar una gran proporción de la información recordada como relevante (Davidson, 1977). Las personas "internas" buscan más información que las "externas" (Prociuk y Breen, 1977). Los sujetos "tolerantes a la ambigüedad" buscan información relativa a sucesos futuros, mientras que los "intolerantes a la ambigüedad" prefieren información sobre la situación actual (Dermer, 1973). Los "sistemáticos" y los "intuitivos" difieren en su evaluación de la información (McKenney y Keen, 1974). Por último, los sujetos muy "creativos" tienden a hacer uso de una mayor variedad de fuentes de información que los menos "creativos" (Kasperson, 1978).

Otros estudios, aunque no tan directamente relacionados con la ejecución en el diagnóstico y detección de fallos o solución de problemas, han mostrado el efecto de este tipo de variables. Por ejemplo, los estudios realizados por Federico y sus colaboradores (Federico, 1982, 1983; Federico y Landis, 1980) encontraron efectos de los estilos cognitivos, las habilidades y las aptitudes de los alumnos de una escuela

de Electrónica y Electricidad en sus calificaciones y años necesitados para terminar la carrera.

Fuera del campo de la Ergonomía, diversas investigaciones han encontrado relaciones entre algunos estilos cognitivos y la toma de decisiones (Hunt y col., 1989; Kotteman y Remus, 1990; Bass, 1983; Taylor, 1984; Volonio, 1987). Por ejemplo, Hunt y col., (1989) encontraron una congruencia entre el estilo cognitivo analítico-inuitivo y la estrategia de decisión preferida por los sujetos. Sin embargo, Kotteman y Remus (1990) encontraron que el estilo cognitivo afecta a la conducta del decisor durante las fases de aprendizaje, pero no en su conducta posterior.

Sin embargo, otras investigaciones han encontrado efectos muy poco significativos o incluso nulos de estas variables sobre la ejecución en tareas de diagnóstico de fallos. Por ejemplo, Rose, Fingerman, Wheaton, Eisner y Kramer (1974) encontraron que la flexibilidad de clausura (como habilidad para localizar figuras ocultas) junto con el scanning espacial (como rapidez de la exploración visual) explicaban una modesta pero significativa parte de la varianza en la ejecución de una tarea simple de solución de problemas. Además, cuando los sujetos son entrenados en la mejora de estas habilidades, no se encuentran diferencias en la ejecución (Levine y col., 1980; Brooke y col., 1983).

4.2.10.3. Variables de Personalidad

Diversas investigaciones sobre detección y diagnóstico de fallos incluyen como posibles variables predictoras de la ejecución, aspectos de la personalidad de los sujetos.

La extroversión destaca como la variable más tenida en cuenta en estos estudios (Colquhoun, 1960; Wilkinson, 1961; Davies y Hockey, 1966). Sin embargo, y debido a las discrepancias encontradas entre los resultados obtenidos en estas

investigaciones, no hay posibilidad de extraer conclusiones definitivas sobre el poder predictivo de estas variables.

Por ejemplo, Gallwey (1982) utilizando un tarea de fallos múltiples, midió la personalidad de los sujetos mediante el cuestionario EPI (Eysenck y Eysenck, 1968), y encontró diferencias significativas en el tiempo medio requerido para realizar la tarea, en la probabilidad del error de clasificación, y en la probabilidad global de éxito, entre los extrovertidos y los introvertidos. Sin embargo, el autor reconoce que es difícil encontrar una explicación de este efecto, y que es necesaria mayor investigación sobre la influencia de este tipo de variables.

Sin embargo, y utilizando el mismo cuestionario (EPI) Seppala y Salvendy (1985) encontraron que las variables de personalidad como el neuroticismo o la extroversión contribuyen muy poco (no fueron significativas en la ecuación de regresión) a la predicción del tiempo de ejecución.

CAPTULO 5. CONCLUSIONES

La primera conclusión que se puede obtener, y que cabe destacar, es que la denominada Ergonomía Cognitiva surge como una rama de la Psicología que pretende aplicar los conocimientos y hallazgos de la Psicología Cognitiva al mundo del trabajo, con el fin de diseñar displays, procedimientos, métodos de entrenamiento, etc.. adecuados a las limitadas capacidades humanas, con el principal objetivo de mejorar la eficacia y la fiabilidad de los operadores.

Uno de los más importantes temas de investigación actual en Ergonomía Cognitiva se centra en el estudio de la carga mental asociada a tareas complejas o múltiples. En este sentido, son principalmente relevantes los estudios sobre memoria (fundamentalmente de memoria de trabajo) y los modelos de atención.

No existe una definición precisa ni única del concepto de carga de trabajo mental. Actualmente parece existir cierto acuerdo en que se trata de un concepto multidimensional que surge de la interacción persona-tarea, y que, por tanto, depende de las características de ambos (Gopher y Donchin, 1986; Navon, 1984; O'Donnell y Eggemeier, 1986). Aunque se han distinguido diversas fuentes o dimensiones de carga, en la actualidad parece que el rango se reduce fundamentalmente a tres: mental/sensorial, temporal y emocional (también denominada estrés).

La principal fuente de desarrollo de la investigación sobre la carga de trabajo mental ha surgido de los modelos atencionales. La Teoría de la Información de Shannon y Weaver (1949) fue el origen de la mayor parte de los modelos de carga que se han ido desarrollando. Los primeros fueron los modelos de filtro propuestos por Broadbent (1958) y Welford (1967), los cuales han sido modificados y mejorados por el grupo de Treisman (Treisman y Gelade, 1980). En general, según estos modelos la carga impuesta por una o varias tareas es función del grado al cual las demandas de la tarea o tareas exceden la capacidad de un filtro que se encarga fundamentalmente de la selección de la información de entrada que, por sus características, merece ser analizada o procesada al completo.

Posteriormente el interés se centró en el estudio de la estructura interna del procesador central y de sus componentes utilizando el concepto de estado de procesamiento (McClelland, 1979; Sanders, 1980; Sternberg, 1969). Para estos modelos la interferencia entre dos tareas se debe a que ambas demandan estructuras o mecanismos de procesamiento comunes. La complejidad de estos modelos hacía muy difícil la medida de la carga. Para solucionar este problema surgieron los modelos llamados del recurso, más interesados en la capacidad del sistema de procesamiento que en su estructura.

Uno de estos modelos es el propuesto por Kahneman (1973) por el cual se postula la existencia de un único e indiferenciado conjunto de recursos disponible para todas las tareas y actividades mentales. Esta teoría ha recibido fuertes críticas por no tener en cuenta el grado de interferencia de las tareas y por no ser capaz de explicar los ejemplos de insensibilidad a la dificultad encontrados en diversas investigaciones (Navon y Gopher, 1979, 1980; Norman y Bobrow, 1975). Estos problemas llevaron a Wickens (1980) a proponer un modelo de múltiples recursos, por el cual la carga tiene lugar cuando las demandas de la tarea exceden los limitados recursos de las personas, o cuando en condiciones de tarea dual ambas demandan recursos comunes. En este modelo se distinguen tres clases de recursos en función del proceso en que intervengan: procesos de codificación, procesos centrales y procesos de selección y producción de respuestas.

Los modelos del recurso han sido los que más esfuerzo han dedicado al estudio de la carga de trabajo mental y al desarrollo de procedimientos para su evaluación y medida. Además estos modelos son los que mejor parecen explicar las diferencias encontradas en los niveles de ejecución, tanto en condiciones de tarea simple como en las de tarea dual, debidas a las diferencias en la capacidad de los individuos y a las modificaciones de las demandas de la o las tareas.

Debido al gran auge que están teniendo en la actualidad los modelos conexionistas o PDP (McClelland y Rumelhart, 1985, 1986; Rumelhart y McClelland, 1986) algunos autores (Gopher y Kimchi, 1989) los incluyen como una rigurosa

aproximación computacional al estudio de la carga. Sin embargo, la aplicación de estos modelos es todavía muy escasa. Fundamentalmente destaca la aplicación realizada al estudio de la adquisición de habilidades por Schneider y Detweiler (1988).

En cuanto a la medida de la carga de trabajo mental, los procedimientos propuestos por los distintos autores y paradigmas no son excluyentes, esto es, pueden ser utilizados conjuntamente. El principal criterio para la elección de uno u otro procedimiento depende fundamentalmente de los objetivos de la investigación y de las características de cada uno de ellos. Por ejemplo, cuando se desea detectar momentos puntuales de carga extrema se aplicaría fundamentalmente la técnica de análisis de línea temporal.

Se han distinguido diversos procedimientos de medida que han sido evaluados en función de una serie de criterios: sensibilidad, poder de diagnóstico, selectividad, intrusividad, fiabilidad, requisitos de implementación, y aceptación del usuario. El grado al cual los diversos procedimientos poseen estas características es variable.

En cuanto a su sensibilidad, los procedimientos de tarea secundaria, los fisiológicos y las técnicas subjetivas pueden utilizarse para evaluar la carga potencial de distintas opciones de diseño, tareas o condiciones de operación. Sin embargo, el procedimiento de tarea primaria se utiliza para determinar si la ejecución del operador será aceptable con un diseño, tarea o condición particular.

En relación al poder de diagnóstico, las medidas de tarea primaria y las subjetivas no se consideran diagnósticas, representan una medida global de la carga. El procedimiento de tarea secundaria y los potenciales evocados son los más diagnósticos. Sin embargo, algunas medidas fisiológicas, como por ejemplo el diámetro pupilar, parecen más sensibles que diagnósticas.

Las técnicas de tarea secundaria son las más intrusivas, mientras que las de tarea primaria no representan problemas de intrusividad. Las técnicas subjetivas no presentan problemas de intrusividad, ya que suelen utilizarse una vez realizada la

tarea. Por otro lado, aunque la intrusión de las técnicas fisiológicas no parece ser muy problemática, hay algunos datos que indican que pueden dar lugar a alguna interferencia.

Por último, las técnicas subjetivas son las que tienen una mayor aceptación por parte de los operadores y las que necesitan menos requisitos de implementación. Aunque no hay datos sistemáticos, parece que las técnicas fisiológicas son las menos aceptadas por el operador, y junto con los métodos de tarea secundaria, son las que implican mayores requisitos de implementación.

Cuando se comparan los resultados obtenidos utilizando distintos procedimientos a partir de tareas de naturaleza similar, se llegará a conclusiones similares sobre los niveles de carga. Sin embargo, si esta comparación se realiza utilizando tareas muy diferentes, en cada una de las cuales se evalúa la carga siguiendo un procedimiento distinto, se pueden encontrar importantes disociaciones en los resultados (Wickens y Derrick, 1981; Wickens y Yeh, 1983; Yeh y Wickens, 1988).

Por otro lado, autores como Wickens y Kessel (1981) han dedicado sus esfuerzos a investigar la carga mental que imponen las tareas de detección y diagnóstico de fallos. En este sentido, es preciso considerar las características tanto de la tarea como del humano, teniendo en cuenta, además, la multidimensionalidad de los recursos de procesamiento humanos, aspecto subyacente al concepto de carga (Wickens, 1979). Desde este punto de vista, se pueden extraer las siguientes conclusiones, relevantes al estudio y análisis de la carga mental impuesta por las tareas de detección y diagnóstico de fallos:

1. En primer lugar, es importante advertir cómo las características de los displays y de los sistemas de alarma influyen en las demandas de la tarea. Por ejemplo, a medida que las características de la información a monitorizar sean más salientes, menor carga será impuesta sobre el operador.

Por otro lado, Sheridan (1981) señala que cuando se disparan las alarmas (auditivas y/o visuales) el operador sufre un gran estrés. Sin embargo, Thompson (1981), resalta que el operador estará más cargado cuando no hay sistemas de alarma que cuando las hay.

En aquellas tareas que demandan recursos de tipo espacial, la carga será menor cuando la información se presenta visualmente, mientras que si las tareas demandan recursos de tipo verbal, la carga será menor si se utilizan inputs de tipo auditivo-verbal.

La mayor parte de los estudios han encontrado que los displays integrales son superiores a los separables cuando las variables de los datos están altamente correlacionadas y/o cuando se requiere la integración de los datos a partir de varias fuentes. En términos de la teoría de múltiples recursos (Wickens, 1984), este tipo de displays produce un código espacial que permite el procesamiento integral de los datos del sistema. Sin embargo, la ventaja de un display integral puede anularse bajo condiciones de incertidumbre. Parece que cuando el estado del sistema es incierto los displays separables producen menores niveles de carga. Otro importante determinante de la efectividad de los displays integrales se relaciona con los requisitos atencionales de la tarea (Casey, 1987). Parece que cuando las variables de estado del sistema están altamente relacionadas, y la tarea de diagnóstico requiere que la atención se focalice en cada uno de los componentes del display, los displays separables son preferibles. Así, mientras el procesamiento holístico proporcionado por el display integral es útil para el análisis del estado general, producirá menores niveles de ejecución cuando las tareas requieren atención selectiva, siempre que los componentes del sistema estén fuertemente interrelacionados (Goettl, Wickens y Kramer, 1991).

Eberts (1987) concluyó que los gráficos disminuyen la carga. Sin embargo, más recientemente, otros autores (Goettl, Wickens y Kramer, 1991) han encontrado que la forma más eficaz para representar datos gráficos depende de la tarea que se esté realizando. Bajo demandas muy duras de memoria, parece preferirse la representación

de tipo gráfica. Sin embargo, la similaridad intra-grupo entre gráficos puede incrementar el tiempo de búsqueda y de identificación.

Según el estudio de Coury y Pietras (1989) la información digital, por sí sola, es insuficiente para la tarea de detección de fallos. En su estudio resultó que el display de tipo múltiple (alfanumérico y gráfico) era superior a los displays de tipo simple (alfanumérico o gráfico), fundamentalmente debido a la presentación redundante de información y a la evidencia confirmatoria suministrada por dos diferentes representaciones del estado del sistema.

Boles y Wickens (1987) encontraron que las tareas que requerían la integración de elementos del display se beneficiaban de displays de formato puro, mientras que las tareas duales lo hacían de displays de formato mixto.

Por otro lado, Coury, Boulette y Smith (1989), Buttigieg y col. (1988) y Coury y Purcell (1988) encuentran que los gráficos de barras pueden poseer propiedades integrales y separables. Sus resultados sugieren que el display de gráfico de barras puede ser un formato muy versátil y es potencialmente la mejor elección para mostrar datos correlacionados, multidimensionales, cuando el rango de incertidumbre es alto.

Por otro lado, la mayoría de las investigaciones encuentran que los displays predictivos mejoran la ejecución de control y planificación, puesto que reducen las demandas sobre la memoria de trabajo.

2. El tipo de control que se ejerza sobre el sistema es otra variable que parece intervenir en la naturaleza de los recursos que demandan este tipo de tareas. Generalmente, en los sistemas que implican un control automático (la mayoría de los sistemas complejos actuales), la existencia de un malfuncionamiento debe ser inducida por el operador humano a partir de los cambios que se observen entre los inputs ambientales y las respuestas del sistema. Utilizando el procedimiento de tarea secundaria, Wickens y Kessel (1981) encontraron grandes diferencias en la ejecución de detección de fallos entre los sujetos que operaban en un sistema manual (MA) y

los que lo hacían en un sistema de control automático (AU). Su conclusión general fue que la automatización no necesariamente elimina o reduce la carga del operador humano, sino que simplemente cambia la localización de las demandas de procesamiento. Mientras que la interferencia con otras respuestas manuales podría atenuarse con la automatización, la interferencia con las tareas cognitivas podría aumentar. En tareas de detección de fallos que suponen bajos niveles de carga la ejecución será superior cuando el sistema es controlado manualmente, mientras que cuando los niveles de carga son altos la ejecución mejorará cuando el sistema es automático.

3. La carga impuesta por una tarea puede verse influida por la estrategia de búsqueda de información que sigan los sujetos. La carga mental del operador dependerá en gran medida de las posibilidades que este tenga para resolver los conflictos entre demandas y recursos, mediante la selección de una estrategia apropiada a la tarea. Además, factores emocionales como la curiosidad, el gusto por el riesgo, etc., pueden influir en esta elección. Por tanto, la elección de una determinada estrategia y el dominio en el que se realiza la búsqueda depende de la situación específica, de los objetivos inmediatos del operador y de las características individuales (personalidad, capacidad, etc.) de los operadores. Además, el esquema general de la tarea puede afectar a la elección de determinados métodos de búsqueda de información.

La búsqueda por síntomas es más económica, pero, en la mayoría de los casos, y sobre todo cuando los sujetos no han sido bien entrenados, requiere la consideración de los múltiples fallos posibles del sistema. La generación de un conjunto de síntomas suele implicar una carga muy alta.

Un resultado particular de la carga mental es la tendencia (observada en contextos de la aviación, en plantas nucleares y en otros ambientes) de los operadores a focalizarse en una hipótesis o idea preconcebida de lo que está funcionando mal y de como solucionarlo, no tomando en cuenta más que la información que satisface su hipótesis, pudiendo pasar de largo información de alta relevancia para el diagnóstico.

4. Parece que el factor de complejidad del sistema que produce mayores niveles de carga es la razón de aparición de nuevas tareas. Cuando el operador se encuentra que tiene que realizar un gran número de tareas y no puede atenderlas a todas, la percepción subjetiva de la carga es extrema. El operador humano de un sistema complejo puede dividir su tiempo entre las tareas o asignar sus esfuerzos sensorial, motor y cognitivo entre varias tareas o demandas independientes, como una estrategia de reducción de la carga y, por tanto, de mejora de la ejecución.

La ejecución será peor a medida que aumente el tamaño del problema (Brooke y Duncan, 1981). Además, la incapacidad mostrada por los sujetos para tomar en cuenta información sobre lo que está funcionando adecuadamente, se acrecienta a medida que aumenta la complejidad del sistema.

Además, la complejidad del sistema influye en el tiempo necesario para realizar el diagnóstico (Wohl, 1981).

5. Los displays de sistemas complejos pueden mejorar la ejecución humana si ellos se ajustan a los modelos mentales que los humanos tienen del sistema, o si ayudan a construir modelos mentales correctos. Cuando la información se muestra en un formato que se corresponde con los modelos mentales que posee el operador, la carga mental se verá considerablemente reducida (Roscoe, 1968).

6. La codificación redundante en diferentes formatos es de gran importancia como factor que reduce la carga mental (Garner y Fefoldy, 1970; Booher, 1975). Sin embargo, no se han encontrado ventajas objetivas del color, como información redundante, en la investigación experimental en situaciones de tarea más complejas que la búsqueda visual (Hale y Billmeyer, 1988).

7. En la medida en que la frecuencia de muestreo requerida de una fuente o fuentes aumenta, así incrementa la probabilidad de que la atención se vea sobrecargada, y cuanto más sobrecargada esté la atención menos posible será que se observe una variable cuyo valor es inadecuado.

Por otro lado, las expectativas de los sujetos determinan la información a la que prestan atención, lo cual determina su respuesta. La evidencia que es esperada (preconcepciones) por el operador es rápidamente aceptada como evidencia conclusiva, y aquella que no es esperada tiende a no ser tomada en cuenta. La resistencia de los observadores humanos a aceptar evidencia contraria a sus hipótesis preconcebidas está relacionada con los sesgos de respuesta (Taylor, 1975; Gai y Curry, 1978; Moray, 1980). Parece que las personas quieren confirmar, más que explorar, la precisión de sus hipótesis.

8. Aunque la experiencia se ha mostrado como un importante factor que afecta a la detección y el diagnóstico de fallos del sistema, las diferencias en ejecución como un resultado de la experiencia son menos aparentes cuando se proporcionan guías para la solución del problema.

Por otro lado, el nivel de entrenamiento también afecta al modelo interno o representación que el sujeto forma del problema. La diferencia existente entre los modelos internos construidos por los sujetos expertos y los no expertos es la responsable de las diferencias en la ejecución de ambos grupos. Los modelos mentales de los sujetos inexpertos son incompletos e imprecisos.

Recientemente, no se han encontrado diferencias en el tiempo requerido para realizar la tarea entre sujetos expertos y superexpertos (Koubek y Salvendy, 1991). Sin embargo, otras investigaciones han encontrado diferencias en los tiempos de tarea entre no expertos y expertos (Vessey, 1985). Por tanto, parece que las diferencias en ejecución, y carga percibida en función del nivel de experiencia de los sujetos, sólo se ponen de manifiesto cuando se comparan grupos extremos. Sin embargo, se han encontrado diferencias en las estrategias de búsqueda de información utilizadas por los expertos y los superexpertos. Parece que los superexpertos buscan en primer lugar la información de tipo más general, mientras que los expertos buscan información más específica.

9. El diagnóstico correcto de los problemas que se dan en sistemas complejos requiere un considerable esfuerzo cognitivo y bastante cantidad de tiempo. Si el tiempo disponible es muy limitado, la probabilidad de que se produzca un error de diagnóstico aumentará. La gran mayoría de los autores incluyen el tiempo disponible para realizar la tarea como una fuente destacada de carga mental.

10. En la mayoría de las investigaciones sobre detección y diagnóstico de fallos se ha encontrado un factor de diferencias individuales de gran importancia.

Los estudios resaltan que de todas las habilidades, las que tienen una mayor relación con la ejecución en diagnóstico de problemas son aquellas relativas al conocimiento del trabajo específico y que las medidas de aptitud general no son muy predictivas de la ejecución en este tipo de tareas. Además, es difícil encontrar diferencias en ejecución entre el grupo de sujetos con altas puntuaciones en tests o pruebas de habilidad y aptitud, y el grupo con puntuaciones medias (Henneman y Rouse, 1984). Parece que, dado un determinado nivel de aptitud, las medidas de estilos cognitivos tienen mayor correlación con el nivel de ejecución.

En este sentido, la mayor parte de los estudios sobre detección y diagnóstico de fallos encuentran una alta relación entre los estilos cognitivos dependencia-independencia de campo y flexibilidad-impulsividad, con la ejecución. Estos estilos cognitivos suelen ser medidos mediante los tests Test de Figuras Ocultas y Emparejamiento de Figuras Familiares, respectivamente. Los resultados permitirían concluir que el efecto de los estilos cognitivos sólo aparece en las fases de aprendizaje de la tarea, pero no se muestra una vez que la tarea ha sido suficientemente practicada (Brooke y col., 1983).

Por otro lado, los resultados de Gallwey (1982) pusieron de manifiesto que el Test de figuras Ocultas era el mejor predictor de la ejecución en una tarea de inspección visual y detección de fallos de diversos tipos.

Sin embargo, otras investigaciones han encontrado efectos muy poco significativos o incluso nulos de estas variables sobre la ejecución en tareas de diagnóstico de fallos (Rose, Fingerman, Wheaton, Eisner y Kramer, 1974).

Diversas investigaciones sobre detección y diagnóstico de fallos incluyen como posibles variables predictoras de la ejecución, aspectos de la personalidad de los sujetos. La extroversión destaca como la variable más tenida en cuenta en estos estudios (Colquhoun, 1960; Wilkinson, 1961; Davies y Hockey, 1966). Sin embargo, y debido a las discrepancias encontradas entre los resultados obtenidos en estas investigaciones, no hay posibilidad de extraer conclusiones definitivas sobre el poder predictivo de estas variables.

CAPITULO 7. INTRODUCCION EXPERIMENTAL

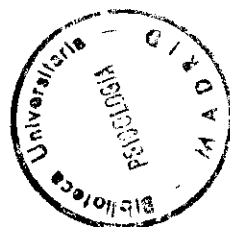
7.1. INTRODUCCION Y OBJETIVOS

El análisis de diferentes accidentes ocurridos en grandes y complejos sistemas (plantas nucleares, control aéreo, etc.) ha permitido concluir que en la mayor parte de las ocasiones la causa está en el error humano que se produce a la hora de detectar y diagnosticar fallos en el sistema. Debido a la importancia y a las graves consecuencias que implican estos errores multitud de autores (Rasmussen y Rouse, 1981; Wickens y Kessel, 1981) han dedicado sus esfuerzos a investigar la carga mental que imponen las tareas de detección y diagnóstico de fallos.

Al investigar la carga impuesta al operador por los procesos de detección de fallos, se deben considerar las características tanto de la tarea como del humano. Además, el análisis de la carga impuesta por este tipo de tareas debe tomar en cuenta la multidimensionalidad de los recursos de procesamiento humanos, aspecto subyacente al concepto de carga (Wickens, 1979). Esta multidimensionalidad ha sido el tema central de numerosas investigaciones (p.e., Navon y Gopher, 1979; Wickens, 1980) y de la conferencia de la OTAN sobre carga (Moray, 1979).

Los niveles altos de carga contribuyen a la generación de una situación de gran peligro. Cuando el operador está sometido a un alto nivel de estrés, su fiabilidad es cero (Swain, 1975, 1981), es decir la probabilidad de que se cometa un error es la total.

Por estas razones, esta investigación pretende como objetivos más generales determinar los factores de complejidad de la tarea y de presentación de la información que afectan a la ejecución de una tarea de diagnóstico de fallos específica de contexto. Por otro lado, también se pretende determinar hasta que punto estos factores afectan a la percepción subjetiva de la carga y si ambos, ejecución y carga subjetiva, producen evaluaciones similares de la carga de la tarea o si por el contrario disocian (Gopher y Braune, 1984; Vidulich y Wickens, 1986; Yeh y Wickens, 1988).



Las técnicas subjetivas de evaluación de la carga han mostrado ser sensibles a las variaciones de la dificultad de la tarea y a las modificaciones en el formato de presentación de la información. Además, son una de las técnicas más utilizadas y cuyo uso por los diseñadores de sistemas y evaluadores de la carga es cada día más creciente. No son intrusivas, son fácilmente utilizadas y aceptadas por los sujetos y no necesitan grandes requisitos de implementación ni entrenamiento de los operadores (O'Donnel y Eggemeier, 1986). Debido a que en esta investigación no se pretende determinar la capacidad o los recursos de los sujetos, sino que se trata de evaluar diferentes formatos de display (numérico, verbal y gráfico), y los efectos de la presión temporal y de otros factores de complejidad (perceptiva y de solución) de una tarea de diagnóstico de fallos específica de contexto, sobre la ejecución y los niveles de carga, se optó por la utilización de una técnica subjetiva de evaluación de la carga, que se describe más adelante.

Además, las técnicas subjetivas son las que necesitan una mayor investigación y validación, pues todavía no se ha encontrado una que permita medir la carga de manera válida y precisa. En este sentido, la utilización en esta investigación de la técnica propuesta permitirá evaluar la validez de la misma en función de si se encuentran o no discrepancias entre las puntuaciones de carga subjetiva obtenidas y las medidas de la ejecución (disociaciones) y en función de la capacidad de la técnica para detectar diferentes niveles de dificultad de la tarea determinados a priori (Vidulich y Tsang, 1986).

Por otro lado, no existen datos concluyentes sobre el poder predictivo de distintas variables de aptitud, estilos cognitivos y de personalidad de los operadores, sobre la ejecución. Mientras que algunos autores han encontrado diferencias en la ejecución debidas a algunas de estas variables, en otras investigaciones la predicción de la ejecución a partir de este tipo de variables fue muy pequeña o incluso nula. Los resultados más contradictorios se encuentran, sobre todo, en relación a las variables de personalidad. En este sentido, en los diferentes experimentos realizados en la

presente investigación se incluyeron diversas medidas de aptitud, personalidad y estilo cognitivo, con el fin de clarificar en la medida de lo posible, el peso que estas variables tienen sobre la ejecución.

En relación a los objetivos anteriormente expuestos, se plantean las siguientes hipótesis generales:

1. El aumento en la complejidad, tanto perceptiva como de solución del problema, de la tarea producirá un claro aumento en el tiempo necesitado por los sujetos para realizarla. Además, la probabilidad de emitir un diagnóstico incorrecto será mayor a medida que se incremente la dificultad de la tarea.
2. El aumento en la complejidad de la tarea producirá mayores niveles de carga subjetiva. Este efecto será principalmente debido al incremento en la valoración subjetiva de la demanda mental de la tarea.
3. La técnica subjetiva utilizada para la medida de la carga será sensible a las variaciones de la dificultad de la tarea. Además, no se espera encontrar grandes disociaciones entre la ejecución y la carga medida con esta técnica.
4. Se espera encontrar una relación aceptable entre el estilo cognitivo dependencia-independencia de campo (DIC) de los sujetos y las medidas de la ejecución. Sin embargo, las variaciones en ejecución no se deberán a las diferencias en la personalidad de los sujetos, ni a su capacidad de razonamiento.

7.2. ELABORACION DE LOS MATERIALES

7.2.1. Pretest

Para elaborar los materiales que compondrían los distintos ensayos de los experimentos se realizó un pretest, en el que participaron 29 sujetos estudiantes de Psicología, con edades comprendidas entre los 19 y los 29 años, de los cuales 7 fueron varones y 22 mujeres. Se elaboró una prueba escrita (ver apéndice 1) en la que aparecían 14 problemas relativos al mal funcionamiento de un sistema de conducción de agua formado por los siguientes componentes:

- .llave de paso general
- .depósito de agua
- .bomba eléctrica, y
- .tubería distribuidora del agua a los distintos pisos del edificio.

La tarea de los participantes consistía en enumerar todas las posibles causas, debidas al mal funcionamiento de alguno de los componentes del sistema, para cada uno de los problemas que formaban la prueba. Así mismo, los sujetos debían indicar al lado de cada una de las causas, la probabilidad de ocurrencia de las mismas, valorándola de 0 a 10.

Al final de la prueba aparecía un espacio en blanco, titulado "otros", en el que el sujeto podía sugerir otros posibles problemas, distintos a los que formaban la prueba. En el caso de que el sujeto respondiera a este espacio en blanco, debía además enumerar las causas posibles, y su probabilidad de ocurrencia, para cada uno de los problemas que había sugerido.

En la corrección de la prueba se eliminaron aquellas causas que hacían referencia a otros componentes distintos a los del sistema concreto que se había propuesto, o que no eran en realidad causas posibles del mismo. En general, los participantes siguieron con bastante fidelidad las instrucciones, por lo que estas eliminaciones fueron muy infrecuentes, y algunas de estas se utilizaron para elaborar los ensayos en los que en realidad la causa del problema no se debía al mal funcionamiento del sistema (falsas alarmas). Por otro lado, ningún sujeto sugirió problemas adicionales, distintos a los que formaban la prueba.

7.2.2. Resultados

Para cada problema se contabilizó el número de causas posibles dadas por los sujetos, así como el número de personas que habían respondido cada causa. Además, para cada una de las causas, se calculó el número de sujetos que habían valorado la probabilidad de ocurrencia con un 1, el número de sujetos que habían valorado la probabilidad de ocurrencia con un 2, con un 3, etc.. hasta 10. Con estos datos se calculó la probabilidad media estimada de ocurrencia de cada una de las causas para cada problema, según la fórmula siguiente:

$$FM = \frac{\sum n_j X_j}{n_c} \times \frac{n_c}{N} = \frac{\sum n_j X_j}{N}$$

donde:

. n_j es el número de sujetos que dieron esa probabilidad.

. X_j es el valor de la probabilidad (1-10).

. n_c es el número total de sujetos que dieron esa causa como posible.

. N es el número total de sujetos ($N=29$).

La posibles causas de cada problema, que fueron emitidas por los sujetos, así como la probabilidad media de cada una de ellas, se muestran en la tabla 7.1.

PROBLEMA	CAUSA	PM
El agua tarda bastante tiempo en salir	1. Bomba estropeada.....	6.36
	2. Llave semicerrada.....	1.85
	3. Tubería con aire.....	1.66
	4. Depósito semivacío.....	0.74
	5. Tubería con escapes.....	0.45
	6. Tubería mal distribuida.....	0.24
Solo hay agua en los pisos bajos pero no en los más altos	1. Poca potencia bomba.....	6.72
	2. Tubería rota.....	3.15
	3. Depósito semivacío.....	0.76
	4. Llave semicerrada.....	0.55
	5. Mal desnivel tubería.....	0.45
6. Tubería muy ancha.....	0.07	
No hay agua en ningún piso de edificio	1. Llave cerrada.....	5.17
	2. Tubería rota.....	3.62
	3. No funciona la bomba.....	3.45
	4. Depósito vacío.....	3.00
El flujo de agua es discontinuo	1. Poca potencia bomba.....	6.45
	2. Llave en mal estado.....	1.55
	3. Tubería con aire.....	1.45
	4. Depósito semivacío.....	0.96
Se ha encontrado humedad cerca de la entrada general al edificio	1. Depósito en mal estado.....	2.69
	2. Tubería rota.....	1.55
	3. Llave en mal estado.....	1.31
	4. Escape en la bomba.....	0.28
El agua tiene mal olor	1. Depósito sucio.....	4.79
	2. Tubería oxidada.....	4.34
	3. Bomba sucia.....	0.17
Sale más cantidad de agua de lo habitual	1. Mucha potencia en bomba.....	3.28
	2. Llave muy abierta.....	2.36
	3. Tubería roturada.....	0.65
El agua sale más caliente de lo normal	1. Mucha potencia en bomba.....	1.79
	2. Depósito mal aislado.....	0.61
	3. Tubería mal aislada.....	0.21
Si agua sale sucia	1. Tubería oxidada.....	5.14
	2. Depósito sucio.....	2.67
El vecino del mismo piso tiene manchas de humedad en el techo de su vivienda	1. Tubería rota.....	4.25
	2. Ventilador roto.....	1.54
El suelo del piso más bajo se ha levantado	1. Tubería rota.....	4.93
	2. Depósito roto.....	2.28
El agua tiene un color extraño	1. Depósito sucio.....	6.07
	2. Tubería oxidada.....	3.96
Si tercer piso tiene humedad en el techo	1. Tubería rota.....	7.45
	2. Ventilador roto.....	1.54
El segundo piso tiene humedad en la pared	1. Tubería rota.....	7.71
	2. Ventilador roto.....	1.54

Tabla 7.1. Causas posibles de cada problema y probabilidad media (PM) de cada una de ellas.

7.2.3. Selección de los ítems

A partir de estos resultados se eligieron los problemas que mejor permitían distinguir entre niveles de complejidad de solución del mismo. En primer lugar, se seleccionaron aquellos problemas con un número de causas mayor o igual a 3, ya que, por un lado, en los que sólo tienen una causa posible no se pueden distinguir niveles de probabilidad de ocurrencia de la solución, y por otro, aquellos con dos causas estas hacen, en todos los casos, referencia a los mismos componentes del sistema (tubería y depósito).

De esta forma, se eligieron los dos problemas con 6 posibles causas, y otros dos de los que tenían entre 3 y 4 causas, y que estimamos de similar frecuencia de ocurrencia en el mundo real. Estos cuatro problemas fueron los siguientes:

- . El agua tarda bastante tiempo en salir.
- . Solo hay agua en los pisos bajos pero no en los más altos.
- . Se ha encontrado humedad cerca de la entrada general al edificio.
- . Sale más cantidad de agua de lo habitual.

Como causas de estos problemas se escogieron aquellas entre las que había una diferencia de PM igual para cada nivel del factor "número de causas". Siguiendo este razonamiento se eligieron, para los problemas con 6 causas, dos de ellas entre las cuales esta diferencia fuera de aproximadamente 6.15 unidades. La diferencia entre la PM de las causas para los otros dos problemas fue de aproximadamente 2.5. El resultado de esta selección se muestra en la tabla 7.2.

PROBLEMA	CAUSAS
El agua tarda bastante tiempo en salir.	1.Bomba estropeada 2.Tubería mal distribuida
Solo hay agua en los pisos bajos pero no en los más altos.	1.Poca potencia bomba 2.Llave semiabierta
Se ha encontrado humedad cerca de la entrada general al edificio.	1.Depósito en mal estado 2.Escape en la bomba
Sale más cantidad de agua de lo habitual.	1.Mucha potencia en bomba 2.Tubería contraída

Tabla 7.2. Problemas y causas seleccionadas para construir los materiales. Las causas 1 tienen mayor PM que las 2.

7.2.4. Construcción de los materiales

Con estos problemas se elaboraron 32 ensayos (4 problemas x 2 causas x 4 niveles de complejidad perceptiva).

Por otro lado, se elaboraron 8 ensayos en los que el problema no se debía a un mal funcionamiento del sistema (falsas alarmas). En estos ensayos cada uno de los cuatro problemas aparecía dos veces: una en la que se presentaban 4 componentes y 4 variables de estado, y otra en la que se mostraban 3 componentes y tres variables. La solución correcta a estos ensayos siempre era "No hay fallo", y siempre aparecía un valor incorrecto o inadecuado en la primera casilla (primer componente y primera variable), lo cual indicaba que el fallo se producía en algún otro sistema o lugar anterior al sistema en cuestión.

Se establecieron, para cada ensayo, cuatro alternativas de respuesta. Una de ellas siempre era la de no fallo. Las restantes tres alternativas variaban entre problemas, pero para un mismo problema siempre eran las mismas. Dos de estas

alternativas hacían referencia a las causas que aparecen en la tabla 7.2. La restante alternativa nunca era la causa real del fallo, sin embargo también era una causa posible (ver tabla 7.1). Las alternativas de respuesta de cada problema aparecen en la tabla 7.3. El orden en el que aparecen estas alternativas en la tabla es el mismo que el que tenían en el experimento. La última alternativa siempre era la de no fallo.

PROBLEMA	ALTERNATIVAS
El agua tarda bastante tiempo en salir.	.Arreglar Bomba .Llenar Depósito .Nueva Tubería
Solo hay agua en los pisos bajos pero no en los más altos.	.Soldar Tubería .Abrir Llave .Regular Bomba
Se ha encontrado humedad cerca de la entrada general al edificio.	.Nueva Llave .Nueva Bomba .Nuevo Depósito
Sale más cantidad de agua de lo habitual.	.Aislar Tubería .Regular Bomba .Cerrar Llave

Tabla 7.3. Alternativas de respuesta utilizadas para cada problema.

Como variables más adecuadas para establecer el estado del sistema se eligieron las siguientes:

- . Presión de entrada al componente del agua.
- . Presión de salida del componente del agua.
- . Nivel o flujo del agua en ese componente (en %)
- . Número de reparaciones que ha tenido el componente.

El número de reparaciones se eligió como variable irrelevante para el diagnóstico, pero cuya presentación podía tener un efecto de aumento de la complejidad perceptiva. En aquellos niveles de complejidad perceptiva en los que se presentaban solamente tres componentes, estos eran los mismos que los que aparecían en las alternativas de respuesta para cada problema (al no aparecer entre las alternativas, el componente restante era irrelevante para el diagnóstico). En los niveles en los que se presentan solamente tres variables, la variable irrelevante era la que se eliminaba de la presentación.

Los valores adecuados de la presión (tanto de entrada como de salida) podían oscilar entre 0 y 9, significando 0 que no había presión (tampoco habría agua), y 9 que la presión era muy elevada. Para representar una presión excesiva se utilizaba el valor 10. El nivel o cantidad de agua venía expresado en porcentajes y sus valores adecuados eran del 90% al 95%. En las instrucciones aparecían estos valores relativos al buen funcionamiento del sistema y se advertía que la bomba solo funcionaría adecuadamente cuando aumentara la presión del agua, pero en dos unidades como máximo.

En principio, se espera que a medida que el rango de causas posibles es mayor, los sujetos necesitarán más tiempo para encontrar el fallo. De igual forma, también se espera que el aumento en el número de elementos del display (número de componentes/variables del sistema) produzca un incremento en el tiempo necesitado para emitir el diagnóstico. Por último, a medida que la probabilidad de ocurrencia de la solución sea más elevada, el tiempo de realización de la tarea se verá reducido.

7.3. PROCEDIMIENTO DE MEDIDA DE LA CARGA SUBJETIVA UTILIZADO

El instrumento de medida utilizado en la presente investigación parte de los resultados encontrados en diversas investigaciones sobre las técnicas de medida de carga subjetiva y sobre las dimensiones de la misma. La mayoría de los autores utilizan versiones modificadas de diferentes técnicas con el fin encontrar aquella que parezca más adecuada a las finalidades de cada investigación, en las que se recogen los resultados más actuales sobre las diferentes dimensiones de carga y sobre los procedimientos para medirla.

En cuanto a las dimensiones de la carga subjetiva, la mayor parte de los autores han llegado a la conclusión, algunos a partir del escalamiento multidimensional de los datos, de que estas se pueden reducir a tres (Jahns, 1973; Johannsen y col., 1979; Kahneman, 1973; Moray, 1982; Reid, Eggemeier y Shingledecker, 1982; Reid, Shingledecker y Eggemeier, 1981; Sheridan, 1980; Sheridan y Simpson, 1979). La primera haría referencia a la actividad mental y perceptiva que requiere la tarea. La segunda dimensión se relaciona con los aspectos de presión temporal de la tarea y la tercera fuente de carga sería la emocional (también denominada estrés).

Las técnicas de evaluación de la carga subjetiva más utilizadas y estudiadas son SWAT, TLX y la escala bipolar de la NASA (la escala de Cooper-Harper (1969) y sus versiones modificadas han sido también objeto de mucha investigación, sin embargo su aplicación válida y fiable está restringida al campo de la aviación). Una de las diferencias entre las tres técnicas mencionadas radica en el número de dimensiones que analizan. SWAT (Reid y col., 1982) asume que la carga subjetiva puede estar adecuadamente representada por las tres dimensiones de carga mencionadas anteriormente (mental, temporal y estrés). La escala bipolar de la NASA descompone la carga en nueve dimensiones (dificultad de la tarea, presión temporal, ejecución, esfuerzo mental/sensorial, esfuerzo físico, frustración, estrés, fatiga y tipo

de actividad). Por último, TLX (Hart y Stavelan, 1988) distingue seis dimensiones (demanda mental, demanda física, demanda temporal, ejecución, esfuerzo, y nivel de frustración).

En cuanto al procedimiento, en las tres técnicas este consta de dos fases: una en la que recogen los parámetros o pesos iniciales de cada dimensión, y otra en la que se estima el nivel de cada dimensión que ha requerido cada ensayo. TLX y la escala bipolar de la NASA utilizan el mismo procedimiento en ambas fases, esto es, los pesos iniciales son calculados a partir del número de veces que cada dimensión ha sido elegida, para cada sujeto, como mayor fuente de carga en el total de comparaciones de pares que se puede establecer con el conjunto de dimensiones (15 y 36 respectivamente). La recogida de esta información se realiza anterior al experimento. Las estimaciones de cada dimensión se realizan inmediatamente después de haber completado cada ensayo, situando una línea vertical en un punto de una escala cuyo rango va de 0 (muy bajo) a 100 (muy alto).

El procedimiento de la técnica SWAT es algo más complejo en el cálculo de los pesos iniciales. En esta técnica se distinguen tres niveles para cada dimensión (bajo, medio, alto), que generan 27 ($3 \times 3 \times 3$) comparaciones binarias posibles de dimensiones y niveles. El desarrollo de la escala requiere que los sujetos ordenen en función del nivel de carga las 27 combinaciones. Mediante la metodología de medida conjunta (Krantz y Tversky, 1971), de esta ordenación se obtiene una escala a nivel de intervalos que asigna una puntuación de 0 a 100 a cada una de las 27 combinaciones. Los sujetos son divididos en subgrupos en función de la dimensión a la que dieron más importancia en la ordenación y se obtiene una escala separada de carga para cada uno de estos subgrupos. En la segunda fase, los sujetos valoran la carga de cada ensayo experimental asignando un 1 (bajo), 2 (medio), o 3 (alto) a cada una de las tres dimensiones. Estas puntuaciones son convertidas en una puntuación de carga que se deriva de la anterior fase de desarrollo de la escala.

Como señalan Vidulich y Tsang (1986), estas técnicas son igualmente sensibles a las variaciones de la dificultad de la tarea, particularmente a las demandas de procesamiento perceptivo/central (Vidulich, 1986; Wickens y Yeh, 1982). Sin embargo, la variabilidad entre-sujetos fue significativamente más alta con SWAT, debido principalmente al procedimiento utilizado para desarrollar los pesos iniciales. El procedimiento de ponderación totalmente individualizado utilizado por la escala bipolar de la NASA y por TLX, fue más efectivo en la reducción de la variabilidad entre los sujetos, produciendo una estimación más estable de la dificultad de la tarea, frente al procedimiento grupal utilizado por SWAT.

Otro de los costos relacionados con SWAT, que señalan Vidulich y Tsang (1986) es su dificultad de uso. La ordenación de las 27 combinaciones es una tarea tediosa que normalmente requiere más de 20 minutos para ser realizada.

A partir de los anteriores resultados, se pensó que sería interesante utilizar una nueva técnica en la que se combinaran las ventajas encontradas en las anteriores. De esta forma la técnica que se utilizó en esta investigación combina las dimensiones de SWAT con el procedimiento seguido por TLX y la escala bipolar de la NASA.

Las tres dimensiones utilizadas se denominaron "demanda mental", "demanda temporal" y "demanda física", y se definieron como sigue:

Demanda Mental: Nivel de actividad mental (recordar, decidir, memorizar, calcular, etc...) y perceptiva (mirar, buscar, distinguir, etc...) que requiere la tarea. ¿Se trata de una tarea fácil o difícil, simple o compleja, exigente o relajada, respecto a esta dimensión?.

Demanda Temporal: Presión temporal sentida debida al tiempo disponible para realizar cada ensayo de la tarea. ¿Es mucho tiempo, poco, suficiente?. ¿En que medida el tiempo disponible para realizar la tarea te hace sentirte inseguro/a o molesto/a?.

Demanda Física: Nivel de actividad física que requiere la tarea (tirar, empujar, pulsar, mover, girar, deslizar, etc...) para alcanzar los objetivos o metas establecidas por el investigador (o por ti mismo). ¿Se trata de una tarea fácil o difícil, aburrida o amena, demasiado lenta o muy rápida, descansada o laboriosa?. ¿Estas satisfecho/a con tu nivel de rendimiento?. ¿Hasta que punto te sientes molesto/a o agradado/a con la tarea?.

La tercera dimensión (demanda física) incluye tanto aspectos de la demanda de ejecución de la respuesta como de sentimientos del sujeto hacia la tarea (molestia, insatisfacción, etc.).

Siguiendo el procedimiento utilizado por TLX y la escala bipolar de la NASA, los sujetos realizaron con anterioridad al experimento las tres comparaciones de pares posibles, indicando cual de las dos dimensiones del par era, a su juicio, mayor fuente de carga. Inmediatamente después de haber realizado cada ensayo el sujeto estimaba en una escala de 0 a 100, separada para cada dimensión, el nivel de carga del ensayo que acababa de realizar debido a cada dimensión (ver apéndice 1). En todo momento, los sujetos disponían de la definición de las dimensiones.

Posteriormente, se obtenían los pesos iniciales para cada dimensión individualizadamente en función del número de veces que había sido elegida como mayor fuente de carga en las comparaciones binarias. Estos pesos podían variar entre 0 (la dimensión no había sido elegida ninguna vez) y 2 (la dimensión fue elegida en todas las comparaciones en las que aparecía), y su suma para todas las dimensiones debía ser igual a 3. Multiplicando la estimación de la dimensión por su peso inicial se obtenía la estimación o puntuación ponderada de cada dimensión.

Para calcular el índice global de carga para cada ensayo, la suma de las tres puntuaciones ponderadas se dividía por 3 (número de comparaciones de pares posibles), según la fórmula siguiente:

$$IC_j = \frac{\sum p_i X_i}{3}$$

donde:

- IC_j : es el índice global de carga del ensayo j.
- p_i : es el número de veces que la dimensión i fue elegida en las combinaciones de pares. ($p=0,1,2$).
- X_i : es la valoración subjetiva estimada para la dimensión i.

**CAPITULO 8. EFECTOS DEL FORMATO DEL SISTEMA,
DEL TIPO DE FEEDBACK, Y DE LA COMPLEJIDAD DE LA
TAREA SOBRE LA CARGA. INFORMACION NUMERICA.**

8.1. INTRODUCCION

En el capítulo seis se han mencionado aquellos aspectos que en mayor o menor medida inciden sobre la carga mental impuesta por las tareas de detección y diagnóstico de fallos. Siguiendo la definición propuesta de carga mental, estos aspectos podrían dividirse en aquellos que hacen referencia a la tarea y aquellos relativos al individuo.

En este sentido, uno de los factores con mayor incidencia sobre la ejecución en tareas de este tipo era la complejidad del sistema a diagnosticar. Diversos autores (Rouse y Rouse, 1979; Brooke y Duncan, 1981) han distinguido dos principales fuentes de complejidad: a) Complejidad perceptiva y b) Complejidad de solución del problema. La primera de ellas se relaciona con aspectos relativos a la presentación, fundamentalmente visual, del sistema, y viene dada principalmente a partir de las modificaciones en el número de elementos del display, el número de conexiones entre elementos, la forma de dichas conexiones (líneas rectas, curvas, etc.), etc. Por otro lado, la complejidad de solución del problema se relaciona con la habilidad para comprender la estructura del problema o las interrelaciones existentes entre sus elementos, más que con la forma en que este es mostrado al sujeto. Por tanto, el formato del problema influye en la complejidad perceptiva, la cual afecta, a su vez, a la complejidad global de la tarea, pero no es una característica particular del problema en sí mismo.

En cuanto a las diferencias individuales, la evidencia de este factor es contradictoria. Con el fin de controlar su posible efecto, en este experimento se tomaron diversas medidas. Debido a que la tarea requiere que el sujeto elabore una serie de hipótesis sobre el componente del sistema que falla y a que estos están causalmente muy interrelacionados, se midió la capacidad de razonamiento de los sujetos, ya que, en principio, los sujetos con baja capacidad de razonamiento no serían capaces de realizar la tarea adecuadamente. Por otro lado, uno de los estilos cognitivos más estudiado y que parece haber demostrado tener un mayor efecto sobre la ejecución en tareas de diagnóstico de fallos, ha sido el de dependencia-independencia

de campo (DIC), por lo cual también se incluyó una medida del tipo de estilo de los sujetos. Por último, también se tomaron medidas de la personalidad de los sujetos, ya que, aunque su efecto sobre la ejecución es muy poco claro, es un factor que suelen incluir la mayor parte de las investigaciones en este campo.

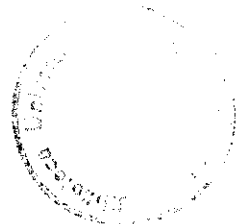
Por otro lado, la mayor parte de las investigaciones utilizan un único tipo de feedback, pero es posible que a diferentes tipos de feedback se obtengan diferentes niveles de ejecución y distintas estimaciones de la carga, a medida que transcurre el experimento (Berkowitz, Lewis y Drury, 1983). Por este motivo, en este experimento se incluyen dos tipos de feedback. El primero de ellos informa únicamente al sujeto de que su respuesta ha sido correcta o no correcta, mientras que el segundo informa de cual era la respuesta correcta. El segundo tipo de feedback es más informativo, por lo que cabe esperar que su efecto de ayuda al aprendizaje y a la ejecución sea superior.

Los objetivos principales de este experimento son los siguientes:

1. Analizar los efectos diferenciales de la complejidad tanto perceptiva como de solución del problema, sobre la ejecución y sobre la percepción subjetiva de la carga.
2. Analizar el posible efecto sobre la ejecución y la carga subjetiva de dos diferentes formatos de representación del sistema.
3. Analizar los posibles efectos que tiene el tipo de feedback proporcionado a los sujetos sobre las estimaciones subjetivas de carga, así como de factor de aprendizaje.
4. Establecer diferentes aspectos, cognitivos, perceptivos y de personalidad, de los individuos, que pueden afectar a los niveles de ejecución.

En relación a estos objetivos, y en función de los trabajos revisados se pueden establecer las siguientes hipótesis:

1. Se espera encontrar un importante efecto de la complejidad tanto perceptiva como de solución del problema sobre la ejecución. Es decir, se espera que el aumento en la complejidad perceptiva y/o de solución del problema, incremente el tiempo necesario para emitir el diagnóstico, el número de datos relativos al estado del sistema que observen los sujetos, y que disminuya el número de aciertos. Además, también se espera encontrar efectos significativos de las interacciones entre los factores, pues como se ha mencionado en la introducción, están teóricamente muy relacionados.
2. No se espera encontrar efectos del formato de presentación del problema (matriz normal o traspuesta).
3. Se espera que el efecto del tipo de feedback se ponga de manifiesto solamente como factor que ayuda al aprendizaje de la tarea. El tipo de feedback correcto-no correcto es *menos informativo*, por lo cual, los sujetos que reciben como feedback la respuesta correcta irán mejorando su ejecución más rápidamente a medida que transcurre el experimento. Pero estas diferencias no se pondrán de manifiesto en los primeros ensayos.
4. En cuanto a las medidas individuales, se espera que sólo la puntuación en DIC sea predictora del nivel de ejecución.



8.2. METODO

8.2.1. Sujetos

Participaron en este experimento 20 sujetos estudiantes de Psicología, con edades comprendidas entre los 20 y los 30 años, de los cuales 3 eran varones y 17 mujeres. Todos ellos recibieron un aumento de 0,25 puntos en la nota final de una de sus asignaturas del curso.

8.2.2. Diseño

El diseño fue un factorial mixto de cinco factores, dos de ellos entresujetos y tres intrasujetos. Los factores entresujetos fueron el tipo de feedback proporcionado (respuesta correcta vs. correcto-no correcto) y el formato de presentación de la información (matriz normal vs. matriz traspuesta). El primer tipo de feedback consistía en informar al sujeto sobre cual era la respuesta correcta a cada ensayo, mientras que el segundo solo hacía referencia a la calidad de la ejecución (es decir, si la respuesta del sujeto había sido la correcta o no). La información sobre el estado del sistema aparecía sobre la pantalla del ordenador en forma de una matriz en la cual cada celdilla contenía el valor de una determinada variable para un componente específico del sistema. Se consideró que la matriz era normal si las variables aparecían formando las columnas y los componentes las filas de la matriz, y que era traspuesta en el caso contrario. Los factores intrasujetos sirvieron para elaborar los distintos problemas o ensayos. Estos factores fueron el nivel de complejidad perceptiva (4x4, 4x3, 3x4, 3x3), el número de causas posibles del problema (3-4 vs. 6), y la probabilidad de ocurrencia de la causa (alta vs. baja). Los niveles de complejidad perceptiva se refieren al número de componentes x número de variables presentado y que componían la matriz de información sobre el estado del sistema. Los niveles de los dos últimos factores están relacionados con los resultados del pretest realizado con anterioridad a la elaboración de los materiales del experimento, y que se describe en el capítulo anterior. Ambos

factores, número de causas del problema y probabilidad de ocurrencia de la solución, hacen referencia a lo que Rouse y Rouse (1979) denominan complejidad de solución del problema.

Además se consideraron otras variables que podían tener relevancia para la investigación. Estas fueron medidas del estilo cognitivo dependencia-independencia de campo (DIC) de los participantes, así como de su personalidad y de su capacidad de razonamiento.

Como variables dependientes se recogieron las valoraciones subjetivas o estimaciones de la carga de la tarea, el tiempo total necesitado para realizar la tarea, el número de información considerada por los sujetos para emitir su diagnóstico y el número de aciertos o diagnósticos correctos emitidos.

8.2.3. Tarea

La tarea del sujeto consistía en, a partir de la información sobre el estado del sistema, encontrar la causa de un problema que tenía lugar en un sistema de conducción de agua. Es decir, el sujeto debía actuar como un técnico de control y mantenimiento de este tipo de sistemas para diversos edificios. Para emitir su diagnóstico el sujeto debía elegir una de las cuatro posibles soluciones a dicho problema, que aparecían en la parte inferior de la pantalla.

En cada uno de los ensayos aparecía, en primer lugar, una breve descripción del problema (p.e., sólo hay agua en los pisos más bajos pero no en los más altos). En la siguiente pantalla aparecía, en forma de matriz, el valor que tenía cada componente del sistema en una serie de indicadores.

Durante la fase experimental, para cada ensayo aparecía en primer lugar uno de los cuatro posibles problemas. Una vez que el sujeto había comprendido el problema debía bajar el cursor hasta el final de la pantalla y pulsar un botón del ratón

para pasar a la pantalla siguiente, en la que aparecía la matriz de información correspondiente. Las celdillas de la matriz aparecían cerradas, con su valor oculto, y para descubrirlo el sujeto sólo tenía que posicionar el cursor dentro de ellas. De esta forma el sujeto buscaba la información que considerase adecuada para el problema determinado que tenía que diagnosticar. Para elegir una de las cuatro alternativas de respuesta el sujeto debía posicionar el cursor sobre ella y pulsar un botón del ratón. Una vez que el experimentador había proporcionado el feedback del tipo correspondiente, el sujeto debía bajar el cursor al final de la pantalla y pulsar un botón del ratón para pasar a la siguiente. Para la mitad de los ensayos, a continuación aparecían tres escalas consecutivas en las que el sujeto debía estimar cada una de las tres dimensiones de carga subjetiva. Para la otra mitad de los ensayos se pasaba directamente a resolver otro problema, a otro ensayo.

El experimento se compuso de 40 ensayos, resultantes del doble del producto del número de niveles de los factores intrasujetos ($2 \times 4 \times 2 \times 2$), más 8 ensayos en los que el problema que aparecía no se debía a un fallo en el sistema, y que se introdujeron con el fin de aumentar la validez ecológica del experimento.

No había tiempo límite para realizar la tarea.

8.2.4. Procedimiento

En primer lugar el sujeto era informado de la duración prevista para la sesión experimental (entre 45 y 60 minutos en total, incluidos los test). Si el sujeto estaba conforme y disponía del tiempo preciso pasaba a completar el test de Figuras Ocultas y, una vez que había transcurrido el tiempo permitido para la realización del mismo (12 minutos), respondía al test de razonamiento. Una vez habían transcurrido los 6 minutos permitidos para la realización de este test, el sujeto era instruido en el manejo del ratón, necesario para la realización del experimento. Una vez el sujeto había comprendido y se había familiarizado con el manejo y funcionamiento del ratón, aparecían en la pantalla del ordenador las instrucciones del experimento (ver apéndice 2). Una vez leídas las instrucciones se resolvían las cuestiones que el sujeto pudiera

tener y se hacía hincapié en la importancia que tenía la memorización y recuerdo de los valores que deberían tener las distintas variables para cada componente para que el sistema estuviera funcionando adecuadamente. Además, previo a la fase experimental se realizaba un ensayo de prueba. A continuación, y previamente a la fase experimental, el sujeto completaba las tres comparaciones de pares de la escala de carga subjetiva, con el fin de determinar los pesos iniciales de cada una de las tres dimensiones que componían dicha escala (demanda mental/ perceptiva, demanda temporal y demanda física/ emocional). A continuación comenzaba la fase experimental. Inmediatamente después de que el sujeto eligiera una respuesta, el experimentador le proporcionaba el feedback del tipo correspondiente. Si se trataba de un ensayo en el que se medía la carga subjetiva, se pasaba a la estimación de cada una de las dimensiones de la misma. En caso contrario, se pasaba directamente a realizar el siguiente ensayo.

Una vez el sujeto había completado todos los ensayos, pasaba a responder el cuestionario de personalidad utilizado.

8.2.5. Instrumentos

Para medir el estilo cognitivo dependencia-independencia de campo se utilizó el Test de figuras Ocultas de Witkin y col. (Witkin y col., 1982). El nivel de razonamiento fue medido utilizando la escala de razonamiento o Factor R del Test de Aptitudes Primarias (PMA) de Thurstone (1981). Se utilizó el EPI (Eysenck y Eysenck, 1987) para medir la personalidad (neuroticismo, extraversión y sinceridad) de los sujetos participantes en este experimento.

Para la presentación de las instrucciones y demás aspectos del experimento se utilizó un ordenador PC/XT con disco duro y un ratón GM-6 PC Mouse. El experimento fue generado y controlado mediante el programa Mouselab System versión 4.2 (Johnson, Payne, Schkade y Bettman, 1989), el cual también se encargó de recoger los datos (tiempos de inspección, secuencia, respuesta, valoraciones, etc.).

8.3. RESULTADOS

8.3.1. Análisis de la ejecución

Con el fin de determinar los efectos de los factores entresujetos, se realizó un MANOVA para cada una de las variables dependientes: tiempo de tarea (TT), número de datos observados (ND) y número de aciertos (NA). En ninguno de estos análisis se encontraron efectos significativos simples de los factores entresujetos (formato y tipo de feedback), ni de su interacción (ver Tabla 1, apéndice 2).

Además, se realizó un MANOVA de los factores intrasujetos (complejidad perceptiva, CP, número de causas del problema, NC, y probabilidad media de la solución, PM) para cada una de las medidas de la ejecución. Los resultados de este análisis se muestran a continuación (ver además Tabla 2, apéndice 2).

8.3.1.1. Tiempo de tarea

Los tres factores de complejidad del problema (complejidad perceptiva (CP), número de causas posibles del problema (NC) y probabilidad media (PM) de ocurrencia de la solución), tuvieron efectos simples significativos [$F(3,48)=23.59, p<0.001$, para CP; $F(1,16)=18.65, p<0.005$, para NC; $F(1,16)=35.21, p<0.001$, para PM] sobre el tiempo de tarea. Así mismo, resultaron significativas las interacciones CPxNC [$F(3,48)=16.97, p<0.001$], CPxPM [$F(3,48)=7.94, p<0.001$] y CPxNCxPM [$F(3,48)=17.15, p<0.001$].

La figura 8.1. muestra el efecto de la interacción entre el número de causas posibles del problema y la complejidad perceptiva del display, sobre el tiempo necesitado por los sujetos para emitir su diagnóstico. La figura muestra como, en

general, a medida que aumentó el número de causas posibles del problema, los sujetos necesitaron menos tiempo para realizar la tarea. Este resultado puede deberse a que las soluciones seleccionadas para construir los materiales de los problemas con mayor número de causas tenían una mayor probabilidad de ocurrencia.

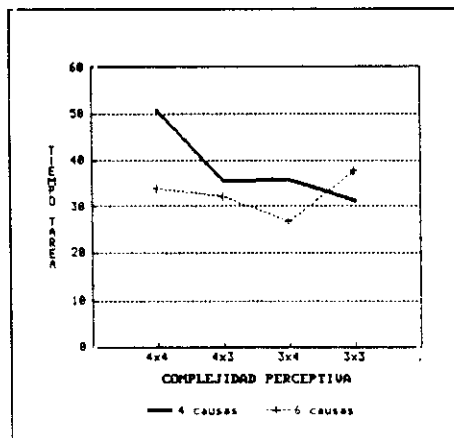


Figura 8.1. Interacción entre la complejidad perceptiva y el número de causas posibles del problema, para el tiempo de tarea (sg.).

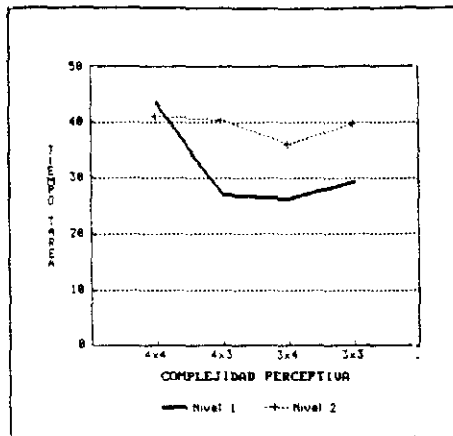


Figura 8.2. Interacción entre la complejidad perceptiva y la probabilidad media de la solución, para el tiempo de tarea (sg.). PM1=prob. alta; PM2= prob. baja.

La figura 8.2. representa el efecto de la interacción entre la complejidad perceptiva y la probabilidad de ocurrencia de la solución, sobre el tiempo de diagnóstico. Se observa claramente que a medida que aumentaba la probabilidad de ocurrencia de la solución, el tiempo necesitado para emitir el diagnóstico era menor.

Se realizaron comparaciones de medias entre los tiempos de tarea de cada par de las combinaciones formadas por los tres factores de complejidad, por el método de Scheffé. La tabla 8.1. muestra el valor del estadístico T para cada contraste significativo. El valor de la diferencia crítica o diferencia mínima significativa (S), para $p < 0.05$, fue 6.7794. La tabla permite observar que en los ensayos pertenecientes a la combinación 141 los sujetos tardaron significativamente más tiempo en realizar la tarea. Además, a medida que disminuye la complejidad perceptiva del display, aumenta el número de comparaciones no significativas

	141	142	161	162	241	242	261	262	341	342	361	362	441	442	461
141															
142	5.7														
161	10.1	4.4													
162	5.8	NS	4.2												
241	9.2	3.6	NS	3.4											
242	5.5	NS	4.6	NS	3.8										
261	11.1	5.4	NS	5.2	NS	5.5									
262	9.1	3.4	NS	3.2	NS	3.5	2.0								
341	9.2	3.6	NS	3.5	NS	3.8	NS	NS							
342	5.6	NS	4.5	NS	3.7	NS	5.5	3.5	3.7						
361	10.6	4.9	NS	4.7	2.9	5.1	NS	NS	NS	5.0					
362	6.4	NS	3.7	NS	NS	NS	4.7	2.7	2.5	NS	4.2				
441	11.0	5.3	NS	5.1	2.9	5.3	NS	NS	NS	5.4	NS	4.6			
442	6.4	NS	3.7	NS	NS	NS	4.7	2.7	2.5	NS	4.2	NS	4.6		
461	7.6	NS	2.5	NS	NS	2.0	3.5	NS	NS	NS	3.0	NS	3.4	NS	
462	6.0	NS	4.0	NS	3.2	NS	5.1	3.1	3.3	NS	4.6	NS	5.0	NS	NS

Tabla 8.1. Valores del estadístico T de Scheffé para cada contraste. El primer valor de cada combinación representa el nivel de la complejidad perceptiva (1=4x4; 2=3x3; 3=4x3; 4=3x3), el segundo representa el número de causas del problema, y el tercero probabilidad de la solución (1=PM alta; 2=PM baja). Los valores fueron significativos al menos para $p < 0.05$. NS= No Significativo.

Igualmente, se realizaron comparaciones de medias de los tiempos de diagnóstico para cada par de combinaciones de la complejidad perceptiva con la probabilidad media de la solución, siguiendo el método de Sheffé ($S=7.084$, $p < 0.05$). La tabla 8.2. muestra los valores del estadístico T que resultaron significativos.

	11	12	21	22	31	32	41	42
11								
12	NS							
21	4.9	4.2						
22	2.1	NS	2.8					
31	4.7	4.0	NS	2.5				
32	NS	NS	4.0	NS	3.8			
41	4.1	3.3	NS	NS	NS	3.2		
42	NS	NS	3.8	NS	3.6	NS	2.9	

Tabla 8.2. Valor del estadístico T de Scheffé para cada contraste significativo entre los tiempos de tarea de cada combinación formada por la complejidad perceptiva y la probabilidad de la solución. NS=No Significativo.

La tabla 8.3. muestra los valores del estadístico T de Scheffé que resultaron significativos en las comparaciones de los tiempos de tarea medios, para cada par de las combinaciones de la complejidad perceptiva (CP) con el número de causas posibles del problema (NC) ($S=6.867$, $p<0.05$). En la tabla se observa que los tiempos de la combinación 14 fueron significativamente más altos que los de las demás combinaciones.

	14	16	24	26	34	36	44	46
14								
16	5.1							
24	4.5	NS						
26	7.1	2.1	2.6					
34	4.5	NS	NS	2.6				
36	5.5	NS	NS	NS	NS			
44	5.7	NS	NS	NS	NS	NS		
46	3.8	NS	NS	3.2	NS	NS	NS	

Tabla 8.3. Valor del estadístico T de Scheffé para cada contraste significativo entre los tiempos de tarea de cada combinación formada por la complejidad perceptiva y el número de causas del problema. NS=No Significativo.

Por otro lado, la comparación del tiempo medio de tarea para cada nivel de complejidad perceptiva sólo resultó significativa para los contrastes en los que aparecía el nivel más alto (ver tabla 8.4). La diferencia crítica (S) del contraste fue igual a 5.659 ($p < 0.05$).

	4x4	3x4	4x3	3x3
4x4				
3x4	4.0			
4x3	3.1	NS		
3x3	2.8	NS	NS	

Tabla 8.4. Valor del estadístico T de los contrastes significativos entre los tiempos medios de tarea de cada nivel de complejidad perceptiva. NS=No Significativo.

8.3.1.2. Aciertos

Sólo los factores de complejidad de solución del problema (NC y PM) tuvieron efectos simples significativos sobre el número de aciertos [$F(1,16)= 8.17, p<0.05$, para NC; $F(1,16)= 18.15, p<0.005$, para PM], resultando la complejidad perceptiva significativa en interacción con NC [$F(3,48)= 6.77, p<0.005$].

La figura 8.3 muestra como a medida que el número de causas posibles del problema era mayor, la probabilidad de cometer un error disminuía. Las figuras 8.1 y 8.3 muestran que, debido a que la probabilidad de ocurrencia de las soluciones de los problemas con seis posibles causas era mayor, estos problemas fueron solucionados más correctamente y en menos tiempo.

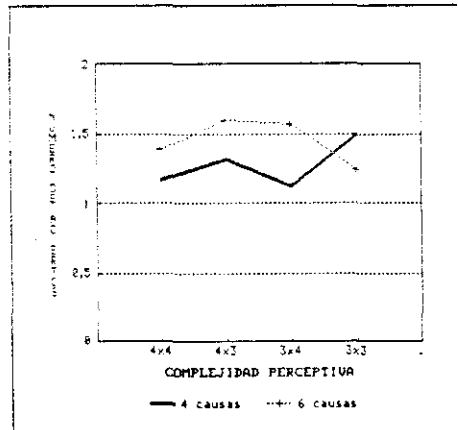


Figura 8.3. Interacción entre la complejidad perceptiva y el número de causas posibles del problema, para el número de aciertos.

Se realizaron comparaciones entre las medias del número de aciertos para las combinaciones de los niveles de complejidad perceptiva por los del número de causas, siguiendo el método de Scheffé ($S=0.3238$, $p<0.05$). La tabla 8.5. muestra los valores del estadístico T que resultaron significativos.

	14	16	24	26	34	36	44	46
14								
16	NS							
24	NS	NS						
26	2.5	NS	2.8					
34	NS	NS	NS	NS				
36	2.6	NS	2.9	NS	NS			
44	2.0	NS	2.3	NS	NS	NS		
46	NS	NS	NS	NS	NS	2.1	NS	

Tabla 8.5. Valores del estadístico T de Scheffé para los contrastes significativos, $p<0.05$, entre el número de aciertos medio de cada combinación CPxNC. NS=No significativo.

8.3.1.3. Número de datos observados

Para la variable dependiente número de datos observados se obtuvieron efectos simples significativos de los tres factores intrasujetos [$F(3,48)= 9.79$, $p<0.001$, para PC; $F(1,16)= 11.37$, $p<0.005$, para NC; $F(1,16)= 111.78$, $p<0.001$, para PM], sin embargo no fueron significativas ninguna de sus interacciones.

La tabla 8.6. muestra las medias y desviaciones típicas de las tres medidas de ejecución.

	TT		NA		ND	
	Media	ST.	Media	ST.	Media	ST.
TOTAL	35.45	15.53	1.36	.65	9.01	3.85
CP 4x4	42.36	18.79	1.28	.65	9.75	4.57
NC 4	51.02	19.71	1.17	.67	10.37	4.56
PM 1	60.59	21.21	1.40	.59	8.25	2.99
PM 2	41.44	12.42	.95	.68	12.50	4.93
NC 6	33.71	13.15	1.40	.63	9.13	4.55
PM 1	26.60	9.08	1.45	.68	7.50	4.17
PM 2	40.83	12.90	1.35	.58	10.77	4.42
CP 3x4	31.13	14.22	1.35	.65	9.10	3.77
NC 4	35.62	15.50	1.12	.68	9.47	3.48
PM 1	29.25	9.25	1.50	.60	8.85	2.62
PM 2	41.99	17.95	.75	.55	10.10	4.14
NC 6	26.64	11.31	1.57	.54	8.72	4.05
PM 1	23.23	12.05	1.70	.47	7.50	4.52
PM 2	30.05	9.63	1.45	.60	9.95	3.17
CP 4x3	33.76	12.16	1.46	.63	9.20	3.80
NC 4	35.49	11.62	1.32	.61	10.10	3.69
PM 1	29.15	9.84	1.50	.51	8.82	3.07
PM 2	41.84	9.77	1.15	.67	11.37	3.89
NC 6	32.02	12.58	1.60	.63	8.31	3.73
PM 1	24.94	10.65	1.65	.58	7.20	4.27
PM 2	39.11	10.30	1.55	.68	9.42	2.78
CP 3x3	34.53	14.16	1.37	.68	8.00	2.96
NC 4	31.32	13.55	1.50	.64	8.82	3.47
PM 1	23.50	7.26	1.70	.57	7.52	2.71
PM 2	39.14	13.98	1.30	.65	10.12	3.72
NC 6	37.74	14.19	1.25	.70	7.17	2.07
PM 1	35.10	16.77	1.35	.67	6.07	1.42
PM 2	40.39	10.85	1.15	.74	8.27	2.06

Tabla 8.6. Medias y desviaciones típicas de las tres medidas de ejecución en cada una de las combinaciones intrasujeto. TT=tiempo, NA=Número de aciertos, ND=Número de datos, CP=Complejidad perceptiva, NC=Número de causas, PM=Probabilidad media de ocurrencia de la solución (1=Alta; 2=Baja).

8.3.2. Medidas de carga subjetiva

En primer lugar se realizó un MANOVA de los dos factores entresujetos para cada medida de la carga subjetiva (para el índice global de carga y para las estimaciones particulares de cada dimensión de carga). No resultó significativo ninguno de los dos factores, ni tampoco su interacción (ver tabla 3, apéndice 1).

8.3.2.1. Índice global de carga subjetiva

Se calculó un índice global de carga para cada combinación a partir de la suma de las estimaciones de cada una de las dimensiones (mental, temporal y física) ponderadas por el número de veces que cada dimensión había sido elegida como mayor fuente de carga en las tres comparaciones de pares, y dividido por el número de comparaciones, según el procedimiento descrito en el capítulo anterior.

Para establecer los efectos de las variables experimentales sobre este índice de carga se realizó un análisis de varianza, en el cual solo resultó significativo el efecto simple de la probabilidad de ocurrencia de la solución [$F(1,16)=20.34, p<0.001$] (ver Tabla 4, apéndice 2).

En la tabla 8.9. se puede observar como a medida que aumenta la probabilidad de ocurrencia de la solución, la carga global percibida por los sujetos disminuye.

8.3.2.2. Dimensiones de carga subjetiva

Posteriormente, con el fin de determinar los efectos de las distintas variables experimentales sobre las estimaciones de las distintas dimensiones de carga subjetiva se realizó otro análisis de varianza en el cual resultaron significativos los efectos simples de los factores intrasujetos, y alguna de las interacciones entre ellos (ver Tabla 4, apéndice 2).

La valoración subjetiva de la demanda mental se vio afectada por la complejidad perceptiva del problema [$F(3,48) \approx 2.90, p < 0.05$] y por la probabilidad media estimada de ocurrencia de la solución [$F(1,16) = 13.90, p < 0.005$].

Por otro lado, resultaron significativos, sobre la estimación de la demanda temporal, los efectos simples de los factores "complejidad perceptiva" [$F(3,48) \approx 3.64, p < 0.05$] y "probabilidad media de la solución" [$F(1,16) = 13.32, p < 0.005$] así como la interacción entre la complejidad perceptiva y el número de causas posibles del problema [$F(3,48) \approx 3.91, p < 0.05$].

Se realizaron comparaciones múltiples de las medias de la valoración de la demanda temporal de las combinaciones CPxNC por el método de Scheffé ($S=7.9087, p < 0.05$). La tabla 8.7. muestra los valores del estadístico T para los contrastes significativos.

	14	16	24	26	34	36	44	46
14								
16	NS							
24	NS	NS						
26	2.3	NS	NS					
34	NS	NS	NS	2.6				
36	NS	NS	NS	NS	NS			
44	2.1	NS	NS	NS	2.4	NS		
46	NS	NS	NS	NS	NS	NS	NS	

Tabla 8.7. Valores del estadístico T de Scheffé para los contrastes significativos, $p < 0.05$, entre las medias de la demanda temporal de cada combinación CPxNC. NS=No significativo.

En cuanto a la demanda física, resultaron significativos los efectos simples de la complejidad perceptiva [$F(3,48) = 3.86, p < 0.05$] de la probabilidad media de la solución [$F(1,16) = 11.64, p < 0.005$] y de la interacción entre la complejidad perceptiva y el número de causas [$F(3,48) = 4.19, p < 0.05$].

Para las combinaciones CPxNC se realizaron comparaciones de medias de la valoración de la demanda física por el método Scheffé ($S = 6.4275, p < 0.05$). La tabla 8.8 muestra los valores de T que resultaron significativos.



	14	16	24	26	34	36	44	46
14								
16	NS							
24	NS	NS						
26	NS	NS	NS					
34	NS	NS	NS	NS				
36	2.2	2.4	NS	NS	2.5			
44	NS	NS	NS	NS	NS	NS		
46	NS	NS	NS	NS	NS	NS	NS	

Tabla 8.8. Valores del estadístico T de Scheffé para los contrastes significativos, $p < 0.05$, entre las medias de la demanda física de cada combinación CPxNC. NS=No significativo.

En resumen, parece que la variable que más incide o afecta a la estimación subjetiva de la carga es la probabilidad de ocurrencia de la solución del problema, la cual como ya se ha mencionado, interviene en la complejidad de solución del problema.

Las figuras 8.4 y 8.5 muestran los efectos de interacción encontrados. A partir de la tabla 8.9, se pueden observar las diferencias encontradas en las estimaciones de la carga debidas a las modificaciones de la complejidad de la tarea. En general, a medida que aumenta el número de elementos del display (complejidad perceptiva), las estimaciones de las tres dimensiones de carga subjetiva son mayores. De igual forma, a medida que la probabilidad de ocurrencia de la solución es más alta, se producen menores estimaciones de las dimensiones de carga subjetiva.

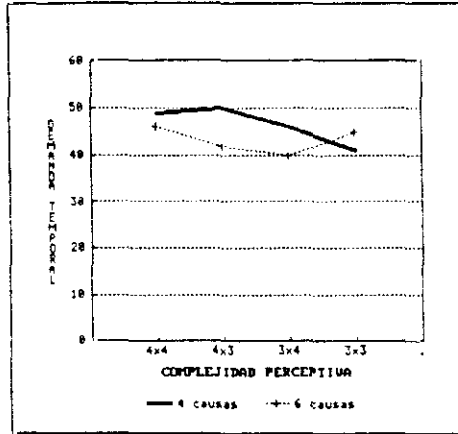


Figura 8.4. Interacción entre CP y NC, para la estimación de la demanda temporal.

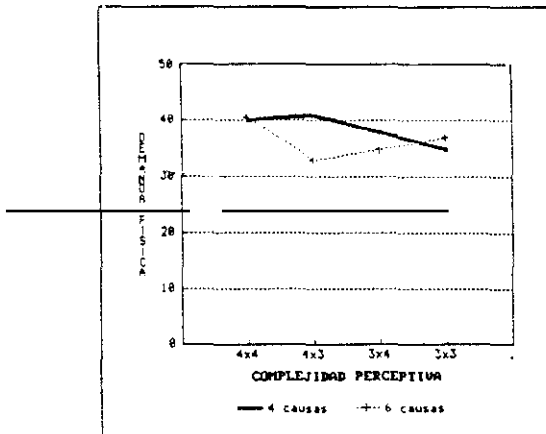


Figura 8.5. Interacción entre la CP y NC, para la estimación de la demanda física.

	DM		DT		DF		INDICE	
	Media	ST.	Media	ST.	Media	ST.	Media	ST.
TOTAL	51.12	16.92	44.91	20.49	37.28	24.69	49.51	17.22
CP 4x4	53.47	18.50	47.59	22.55	40.02	26.55	51.72	19.06
NC 4	54.77	18.92	48.95	23.50	39.62	26.61	52.63	19.99
PM 1	53.86	22.57	49.21	28.16	40.38	28.41	51.90	23.37
PM 2	55.68	14.95	48.70	18.45	38.96	25.40	53.36	15.88
NC 6	52.17	18.22	46.23	21.78	40.42	26.82	50.81	18.05
PM 1	48.08	17.22	42.68	21.40	38.28	24.52	47.18	16.08
PM 2	56.25	18.70	49.77	22.11	42.56	29.42	54.45	19.16
CP 3x4	50.04	16.80	42.92	19.83	36.35	24.34	48.46	16.90
NC 4	52.63	16.85	45.94	20.59	38.04	24.46	51.19	17.08
PM 1	52.92	19.58	44.39	23.31	36.48	26.47	50.38	19.88
PM 2	52.33	14.12	47.49	17.94	39.60	22.86	52.00	13.68
NC 6	47.46	16.55	39.91	18.80	34.66	24.42	45.74	16.28
PM 1	44.73	18.40	37.23	17.94	35.38	26.99	43.19	17.48
PM 2	50.19	14.43	42.59	19.71	33.93	22.23	48.30	14.56
CP 4x3	52.14	15.93	46.11	20.40	36.75	23.96	49.98	16.35
NC 4	56.65	16.12	50.05	20.59	40.80	26.79	54.13	16.70
PM 1	53.13	17.07	45.90	19.45	35.99	23.94	50.52	16.23
PM 2	69.17	14.69	54.19	21.35	45.61	29.17	57.74	16.39
NC 6	47.62	14.57	42.17	19.67	32.70	20.28	45.84	14.88
PM 1	45.46	16.76	39.93	20.52	32.32	20.50	43.87	16.48
PM 2	49.79	12.05	44.40	19.04	33.38	20.57	47.82	12.79
CP 3x3	48.81	16.26	43.01	19.01	35.99	24.06	47.88	16.19
NC 4	49.39	15.72	40.71	17.05	34.90	23.37	47.64	14.47
PM 1	45.85	16.17	39.09	19.78	31.44	21.14	44.35	15.52
PM 2	52.93	14.80	42.33	14.13	38.37	25.49	50.94	12.52
NC 6	48.23	16.96	45.32	20.75	37.07	24.98	48.13	17.75
PM 1	43.49	17.50	40.63	20.85	31.96	22.11	43.52	18.59
PM 2	52.98	15.40	50.00	20.07	42.19	27.14	52.74	15.57

Tabla 8.9. Medias y desviaciones típicas de las tres dimensiones de carga en cada una de las combinaciones intrasujeto. DM=Demanda mental; DT=Demanda temporal; DF=Demanda física; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad de ocurrencia de la solución (1=Alta; 2=Baja).

8.3.3. Diferencias individuales

Con el fin de determinar la influencia de las diferencias individuales sobre la ejecución, se realizó un análisis de regresión por el método Stepwise, para cada una de las variables dependientes, utilizadas en anteriores análisis, tomando como variables predictoras la puntuación obtenida por los sujetos en el Test de Figuras Ocultas (FO), en la prueba de razonamiento del PMA (R), y en cada una de las variables de personalidad (Neuroticismo, NEU, Extroversión, EXT, y Sinceridad, SIN) del cuestionario EPI (ver Tablas 5 a 7, apéndice 2).

Para la variable criterio "tiempo de tarea", entraron a formar parte de la ecuación de regresión la puntuación en el Test de Figuras Ocultas [T=2.606, $p<0.01$], la puntuación en razonamiento [T=-4.200, $p<0.0001$], y el neuroticismo [T=3.323, $p<0.005$; $R^2=0.07903$; R^2 ajustada= 0.07029].

En la ecuación de regresión en la que el "número de datos observados" era la variable criterio, entraron la puntuación en el Test de Figuras Ocultas [T=2.043, $p<0.05$], y el razonamiento [T=-3.126, $p<0.005$; $R^2=0.03362$; R^2 ajustada= 0.02752].

Cuando la ecuación de regresión se realizaba para predecir el número de aciertos, sólo entró a formar parte de la misma, la puntuación en razonamiento [T=5.506, $p<0.0001$; $R^2=0.08705$; R^2 ajustada= 0.08418].

	FO	R	NEU	EXT	SIN
TT	.0715	-.1717*	.1495*	-.0473	-.0536
ND	.0619	-.1445*	-.0672	.0303	.0131
NA	.1076	.2950**	-.0543	.0017	.0424

Tabla 8.10. Matriz de correlaciones entre las variables predictoras y cada una de las medidas de la ejecución (TT=Tiempo de tarea; ND=Número de datos observados; NA=Número de aciertos). Un * significación para $p<0.01$, y dos ** para $p<0.001$.

La tabla 8.10 muestra las correlaciones obtenidas, y su significación, entre las medidas de la ejecución y las de características individuales. El nivel de razonamiento es la que obtiene correlaciones más altas (aunque no muy elevadas) con las medidas de la ejecución, principalmente con el número de diagnósticos correctos emitidos. El signo de las correlaciones indica que los sujetos con mayor capacidad de razonamiento necesitaron observar un menor número de fuentes de información y tardaron menos tiempo en realizar la tarea. Además, cometieron menos errores de diagnóstico que los sujetos con puntuaciones más bajas en razonamiento.

8.3.4. Feedback como factor de aprendizaje

El análisis de la figura 8.6 (a y b) muestra que el feedback no actuó como factor de aprendizaje, ya que no se observa una clara tendencia al aumento en el número de aciertos a medida que transcurría el experimento. De igual forma, permite observar que el tipo de feedback no tuvo efectos sobre el número de aciertos, lo que ya se había puesto de manifiesto en el MANOVA realizado para los factores entresujeto.

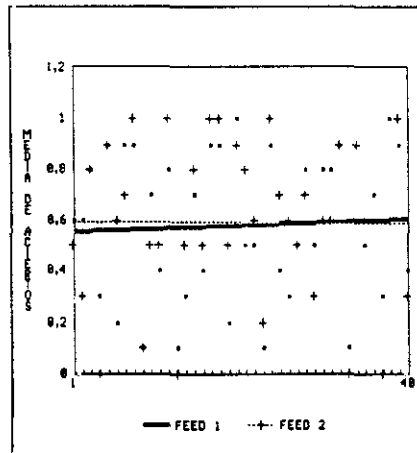


Figura 8.6a. Representación de la tendencia de los datos de la figura 6b.

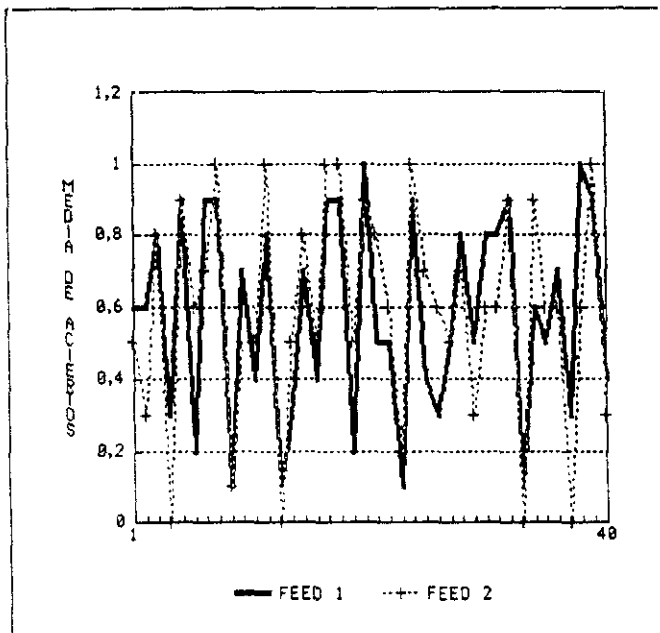


Figura 8.6b. Número de aciertos medio para cada ensayo, según el tipo de feedback recibido por los sujetos. Feed 1=respuesta correcta; Feed 2 =correcto/no correcto.

8.4. DISCUSION

Se utilizaron tres medidas de la ejecución: tiempo necesitado para realizar cada ensayo, número de datos que observaron los sujetos antes de emitir su diagnóstico, y el número de diagnósticos correctos para cada tipo de problema. El efecto de la complejidad perceptiva del display resultó significativo sobre las dos primeras medidas de la ejecución. También se encontraron efectos del número de causas posibles del problema y de la probabilidad de ocurrencia de la solución sobre el tiempo de tarea, así como efectos de las interacciones entre las tres variables de complejidad (perceptiva y de solución del problema). Sobre el número de aciertos sólo resultaron significativos los efectos de las variables de complejidad de solución del problema, y de la interacción entre la complejidad perceptiva y el número de causas posibles del problema.

Por otro lado, en cuanto al índice global de carga sólo resultó significativo el efecto de la probabilidad de ocurrencia de la solución. Las estimaciones de cada una de las tres dimensiones de carga (mental, temporal y física) resultaron afectadas por las variaciones en la complejidad perceptiva y en la probabilidad de ocurrencia de la solución. Además la interacción entre complejidad perceptiva y número de causas posibles del problema resultó significativa para las estimaciones de la demanda temporal y de la física.

Aunque algunas de las medidas individuales (fundamentalmente el estilo cognitivo DIC, la puntuación en razonamiento y el neuroticismo) entraron a formar parte de la ecuación de regresión para predecir la ejecución, estas variables explicaron una parte muy pequeña de la varianza de la ejecución. Esto indica que, en contra de los resultados de otras investigaciones (Henneman y Rouse, 1984; Hunt, Henneman y Rouse, 1981; Rouse y Hunt, 1984; Rouse y Rouse, 1979, 1982), las variaciones en la ejecución no se debieron a diferencias individuales en razonamiento, personalidad y estilo cognitivo.

En contra de lo previsto, el estilo cognitivo DIC no tuvo correlaciones significativas con ninguna de las medidas de la ejecución. Por otro lado, la capacidad de razonamiento fue la variable que correlacionó más alto con las medidas de la ejecución.

Como se había previsto, la capacidad de razonamiento correlacionó negativamente con el tiempo y el número de fuentes de información observadas necesitadas para emitir el diagnóstico, y positivamente con el número de soluciones correctas dadas a los problemas presentados.

En este sentido, en la investigación realizada por Seppala y Salvendy (1985), anteriormente mencionada, se utilizaron como variables individuales la puntuación en el test de figuras ocultas, la extroversión y el neuroticismo, entre otras variables. Ninguna de estas variables fue predictora del porcentaje de errores en la tarea. Estas variables fueron predictoras del tiempo de ejecución, sin embargo su contribución a la varianza de la medida fue en algunos casos inferior a la encontrada en este estudio. Por ejemplo, la extroversión fue capaz de explicar solamente el 2.67% de la ejecución, y el neuroticismo el 1.37%. Sin embargo, la puntuación en el test de figuras ocultas fue capaz de explicar el 56.59% de la varianza en el tiempo de ejecución. Estos resultados, junto con los aquí encontrados, permitirían concluir que la personalidad de los sujetos no parece ser un factor que determine las diferencias en ejecución de los individuos. Sin embargo, las discrepancias encontradas sobre el efecto del estilo cognitivo DIC, *no permiten extraer conclusiones definitivas sobre el poder predictivo de esta variable.*

Por otro lado, los resultados indican que, como se había previsto, el número de datos observado por los sujetos para emitir su diagnóstico aumentó a medida que la complejidad perceptiva del sistema, establecida por el número de elementos del display, era mayor. Además, se encontraron fuertes diferencias entre el tiempo medio de tarea de los displays con 16 elementos y todos los demás (ver tabla 8.4). De igual

forma, observando la tabla 8.1 se puede ver que a medida que la complejidad perceptiva del display disminuía, se encontraba un menor número de diferencias significativas entre los tiempos medios de tarea de los distintos tipos de problemas.

Resultados similares han sido encontrados en otras investigaciones en las que se utilizaban tareas libres de contexto (Rouse, 1981; Brooke y Duncan, 1981). Por ejemplo, Rouse (1981) encontró dificultades a la hora de confirmar estos resultados cuando los sujetos realizaban una tarea específica de contexto, similar a la utilizada aquí. Siguiendo a Rasmussen (1981), Rouse interpretó estas diferencias por el cambio de una estrategia topográfica a una estrategia de búsqueda sintomática en la que, muy frecuentemente, los sujetos utilizan reglas altamente dominadas por el contexto, basadas en sus experiencias pasadas. Los sujetos que participaron en los experimentos descritos por Rouse eran estudiantes de Ingeniería, familiarizados con los problemas que se les presentaron, y que podríamos denominar "expertos". Sin embargo, en la investigación que aquí nos ocupa, los sujetos fueron estudiantes de Psicología, no familiarizados con los problemas de los sistemas de conducción de agua. Esta diferencia en el nivel de experiencia, podría ser la responsable de que los sujetos que participaron en este experimento utilizaran una estrategia mayoritariamente topográfica, a pesar de que la tarea fuera específica de contexto. El uso de esta estrategia, basada en el conocimiento estructural del sistema, hizo que los sujetos no tuvieran en cuenta las relaciones de dependencia jerárquica entre los componentes y las variables del sistema, por lo cual observaron información totalmente irrelevante para el diagnóstico y para su confirmación. Además, el hecho de que el tiempo de tarea también aumentara con la complejidad perceptiva del sistema (resultado que también se puso de manifiesto en las investigaciones antes mencionadas), confirma la idea de que los sujetos utilizaron una estrategia de este tipo.

Sin embargo, tanto el tiempo necesitado para realizar la tarea, como el número de datos observados por los sujetos para emitir su diagnóstico, resultaron afectados no sólo por la complejidad perceptiva, sino también por las dos variables de complejidad de solución del problema. Este resultado parece indicar que los sujetos utilizaron una mezcla de ambas estrategias (topográfica y sintomática). Como Rasmussen (1981) y

Rouse (1981) ponen de manifiesto, los sujetos pueden variar el tipo de estrategia utilizada en cualquier momento del experimento y/o de cada ensayo, para reducir los niveles de carga que experimentan (Rouse y Rouse, 1984).

En principio, cabría pensar que los sujetos comenzaron la búsqueda de información en función de una hipótesis previa sobre el origen del problema. Si los valores de la variable y del componente incluidos en su hipótesis la confirmaban, el sujeto emitía la respuesta. Pero si los datos observados no confirmaban esta hipótesis, el sujeto continuaba la búsqueda de más información siguiendo una estrategia topográfica. El uso de esta estrategia mixta explicaría que las interacciones entre la complejidad perceptiva y los factores de complejidad de solución del problema fueran tan altamente significativas para la medida del tiempo de tarea.

La utilización de esta estrategia parece ser también la responsable de que el tipo de matriz (normal/traspuesta) utilizada para representar el sistema, y el tipo de feedback proporcionado, no tuvieran efectos significativos sobre la ejecución ni sobre la carga subjetiva. La utilización de una estrategia topográfica supone una secuencia de juicios de tipo bueno/malo de la magnitud de las variables de estado en función de los datos de referencia (los proporcionados en las instrucciones), lo cual hace que la información contenida en cada observación sea utilizada de forma antieconómica. De esta forma, cuando no se confirmaba su hipótesis previa, los sujetos buscaban valores de las variables de estado del sistema distintos a los de referencia, y una vez encontrados observaban a que componente hacía referencia dicho valor. para así emitir su respuesta. La elección de esta estrategia está determinada por el principio de menor resistencia (Rasmussen, 1981), cuyo criterio subyacente es la reducción de la carga impuesta sobre la memoria de trabajo. Este principio unido a la falta de experiencia de los participantes, determinaron el cambio de una estrategia sintomática a otra topográfica.

El uso de esta estrategia mixta y la tendencia de los sujetos a no tener en cuenta la información proporcionada por el feedback, sobreestimando sus hipótesis

previas (Einhorn y Hogarth, 1978; Fischhoff, 1977; Fischhoff y MacGregor, 1981), hizo que el tipo de feedback no tuviera ningún efecto de aprendizaje. Por ejemplo, Fischhoff (1977) describe el fenómeno que denominó "vanidad cognitiva" (*cognitive conceit*), según el cual las personas tendemos a infraestimar la información proporcionada por los resultados de nuestra conducta (*feedback*), sobreestimando el valor de nuestras previas hipótesis, lo que da origen a la no revisión de los procesos y estrategias de decisión utilizados, hecho que se reafirma cuando, por azar, se produce un acierto (Einhorn y Hogarth, 1978).

Como se había previsto, el tiempo de diagnóstico y el número de errores, aumentaron a medida que disminuía la probabilidad de ocurrencia de la solución. Sin embargo, y en contra de las hipótesis iniciales, los sujetos necesitaron menos tiempo para realizar la tarea y emitieron un mayor número de diagnósticos correctos, cuando el número de causas posibles del problema era mayor. En un principio se pensó que cuando el rango de soluciones posibles de un problema era menor, la solución correcta sería más rápidamente encontrada. Sin embargo, un análisis más detallado de los ensayos demuestra que la probabilidad de ocurrencia de las soluciones de los problemas con seis posibles causas era muy superior a la de aquellos que tenían cuatro posibles causas. Posiblemente esta diferencia en la probabilidad de ocurrencia de las soluciones sea el origen de este resultado.

El hecho de que la complejidad perceptiva aumentara el número de datos observados por los sujetos pero no el número de aciertos demuestra que a medida que el número de fuentes de información crece, los sujetos tienden a observar más información de la necesaria (Wright, 1974), pero esta no es utilizada para tomar decisiones más precisas (Dawes, 1979; Dawes y Corrigan, 1974; Hayes, 1964; Schroeder y Benbassat, 1975; Oskamp, 1965).

Por otro lado, los niveles globales de carga percibida por los sujetos dependieron únicamente de las variaciones en la complejidad de solución del problema, establecida por la probabilidad de ocurrencia de la solución, de forma que

a medida que la probabilidad de la solución era menor, los sujetos experimentaban mayores niveles de carga. El efecto de la complejidad perceptiva se puso de manifiesto sobre las estimaciones individuales de cada dimensión de carga, pero menos significativamente que el de la probabilidad de la solución, y además no tuvo ningún efecto cuando las diversas estimaciones se combinaron en una puntuación de carga global. Esto hace pensar que, aunque el aumento de la complejidad perceptiva del sistema produjo mayores niveles de carga subjetiva, esta está principalmente determinada por factores que afectan más a la complejidad de la solución.

CAPITULO 9. EFECTOS DE LA PRESION TEMPORAL.

9.1. INTRODUCCION

El estudio de diversos accidentes que han tenido lugar en sistemas complejos (principalmente en plantas nucleares) ha puesto de manifiesto que la toma de decisiones es una de las funciones más importantes del operador humano en los sistemas hombre-máquina actuales, especialmente en situaciones de emergencia, en las que el operador debe tomar una decisión correcta sobre el estado del sistema y sobre el componente que falla bajo condiciones de alta presión temporal.

La mayor parte de los investigadores reconocen la presión temporal como un factor mas, entre otros muchos, causante de la carga mental. Sin embargo, para algunos de ellos (Parks, 1979; Tulga, 1978) la carga está determinada única y exclusivamente por aspectos temporales. Siguiendo esta visión, los autores definen la carga en función del tiempo utilizado o necesario para realizar la tarea y el tiempo disponible, y siguiendo la metodología de análisis de línea temporal, descrita en el capítulo 3, la miden, la analizan, e intentan predecirla.

Por ejemplo, Senders (1979) opina que en las tareas que no requieren presión temporal, no se puede hablar de carga subjetiva. Otros autores como Phillip, Reiche y Kirchner (1971), Schmidt (1978) o Hacker, Plath, Richter y Zimmer (1978) han encontrado que la percepción subjetiva de la carga depende esencialmente de la presión temporal establecida para realizar la tarea. Sin embargo, en el experimento anterior, la técnica de medida de la carga subjetiva utilizada, fue sensible a las modificaciones de la dificultad de la tarea, aunque la tarea no requería presión temporal.

De todas formas, y concediéndole mayor o menor importancia, la presión temporal ha resultado ser un factor con gran incidencia sobre la ejecución y la carga asociada a tareas complejas de detección y diagnóstico de fallos (Moray, 1982; Ward y Poturalski, 1982; Zakay y Wooler, 1984).

Por ejemplo, Zakay y Wooler (1984) concluyen que la presión temporal claramente disminuye la efectividad de la decisión. Sus resultados también indicaron que la transferencia del entrenamiento en toma de decisiones bajo condiciones de no presión temporal a otras en las que el tiempo está limitado, sería muy pobre. De igual forma, Ward y Poturalski (1982), utilizando una tarea de solución de laberintos, concluyeron que la presión temporal produjo disminuciones en el nivel de ejecución alcanzado por los sujetos, así como cambios en las estrategias de solución del problema utilizadas.

Por otro lado, Moray (1982) revisa algunos de los factores que inciden en la carga percibida por los sujetos que realizan tareas de monitorización y control de sistemas complejos, resaltando la importancia de la presión temporal.

El objetivo fundamental de este experimento consiste en establecer cómo se ven afectadas la ejecución y la carga subjetiva cuando la tarea se realiza bajo presión temporal. Por este motivo, este experimento es muy similar al anterior, la única variación radica en la inclusión de un factor de presión temporal. Con este experimento se pretende dar respuesta a las siguientes cuestiones:

- . ¿Hasta qué punto la presión temporal modificará las valoraciones de carga subjetiva?. Si estas estimaciones se ven alteradas, ¿cuál o cuáles de las tres dimensiones de carga (mental, temporal y física) se verán afectadas?.
- . En cuanto a la ejecución, ¿aumentará el número de errores?, ¿variará el número de datos recogidos o tenidos en cuenta por los sujetos para emitir su diagnóstico?.
- . ¿Se modificará la estrategia de búsqueda de información?.

En relación a estos objetivos se pueden establecer las siguientes hipótesis:

1. Se espera que el efecto simple de la presión temporal sea significativo. Se espera que el tiempo de tarea disminuya, pero que aumenten el número de errores y de datos observados.
2. Puesto que la presión temporal ha resultado ser un factor de complejidad en otras investigaciones, se espera que sus interacciones con los factores intrasujeto sean significativas.
3. La presión temporal producirá un aumento en la estimación global de la carga subjetiva de la tarea, debido principalmente al incremento en la estimación de la demanda mental de la tarea.
4. Debido a la presión temporal, se reafirmará la preferencia por la estrategia de búsqueda topográfica.
5. En cuanto al efecto de los demás factores experimentales, se espera encontrar resultados muy similares a los del experimento anterior.

9.2. METODO

9.2.1. Sujetos

En este experimento participaron 20 sujetos estudiantes de Psicología, distintos a los del primer experimento, con edades comprendidas entre los 20 y los 27 años, de los cuales 3 fueron varones y 17 mujeres. Todos ellos recibieron un aumento de 0,25 puntos en la nota final de una de sus asignaturas del curso.

9.2.2. Diseño

El diseño y las variables utilizadas fueron exactamente los mismos que en el primer experimento. Con la única diferencia de que en este caso el tiempo disponible para realizar cada ensayo fue limitado a 30 segundos. Se eligió este tiempo porque los resultados del experimento anterior indicaban que podría ser consistentemente difícil resolver los problemas en medio minuto.

9.2.3. Tarea y Procedimiento

El procedimiento y la tarea fueron los mismos que los descritos para el experimento anterior, con la única diferencia de que en este caso si el sujeto sobrepasaba los 30 segundos permitidos como tiempo límite para realizar cada ensayo, aparecía en la parte superior de la pantalla un mensaje de aviso y se producía un sonido si el sujeto intentaba seguir buscando información, en cuyo caso las celdillas permanecían cerradas. Por lo tanto, una vez sobrepasado el tiempo límite el sujeto estaba forzado a seleccionar una respuesta. Con el fin de que el sujeto conociera el tiempo de que disponía en todo momento, en la parte superior izquierda de la pantalla aparecía un círculo que, a modo de reloj, iba poco a poco desapareciendo a medida que transcurría el tiempo.

9.2.4. Instrumentos

Para este experimento se utilizaron los mismos instrumentos que en el experimento anterior.

9.3. RESULTADOS

9.3.1. Efectos de la presión temporal

Para determinar los efectos de la presión temporal sobre las medidas de ejecución y las de carga subjetiva, se realizó un MANOVA para cada una de las variables dependientes, utilizando los datos de este y el anterior experimento (N=40) (ver tablas 1 y 2, apéndice 3).

En este análisis el efecto de la presión temporal resultó significativo para el tiempo de tarea [$F(1,32)= 27.23, p<0.001$] y para el número de aciertos [$F(1,32)= 10.21, p<0.005$].

Además, para el tiempo de tarea resultaron significativas las interacciones siguientes: presión temporal x complejidad perceptiva [$F(3,96)= 6.86, p<0.001$], presión x número de causas [$F(1,32)= 11.78, p<0.005$], presión x probabilidad media de la solución [$F(1,32)= 23.13, p<0.001$], presión x complejidad perceptiva x número de causas [$F(3,96)= 4.68, p<0.005$], presión x complejidad perceptiva x probabilidad media de la solución [$F(3,96)= 2.79, p<0.05$], y presión x complejidad perceptiva x número de causas x probabilidad de la solución [$F(3,96)= 8.02, p<0.001$].

Se realizaron comparaciones múltiples de medias entre los tiempos medios de tarea de cada combinación PCxNCxPM entre los grupos con y sin presión, por el método de Scheffé. Para $p<0.05$, el valor de la diferencia crítica (S) fue igual a 3.771. La tabla 9.1. muestra el valor de la diferencia y del estadístico T para los contrastes que resultaron significativos.

	Diferencia	T
141	27.27	14.35
142	16.05	8.44
161	3.98	NS
162	14.92	7.85
241	6.52	3.43
242	17.64	9.28
261	4.52	2.37
262	6.70	3.52
341	4.76	2.51
342	15.10	7.94
361	2.89	NS
362	11.87	6.24
441	0.74	NS
442	16.50	8.68
461	6.87	3.61
462	12.97	6.82

Tabla 9.1. Valores de la diferencia en los tiempos de tarea con y sin presión temporal, y de T para cada combinación. NS=No Significativo.

Los resultados que aparecen en la tabla 9.1. muestran que en prácticamente todos los casos el tiempo de tarea se vio muy significativamente reducido por la presión temporal.

En relación al número de aciertos, resultó significativa la interacción entre la presión temporal, el número de causas y la probabilidad media de la solución [$F(1,32)= 5.49, p<0.05$].

En cuanto al número de datos observados, sólo se encontraron efectos de la presión temporal en interacción con la probabilidad media de la solución [$F(1,32)= 15.39, p<0.001$], y con la interacción entre el número de causas y la probabilidad de la solución [$F(1,32)= 5.52, p<0.05$].

Las figuras 9.1, 9.2 y 9.3 muestran que como se había previsto, el tiempo requerido para emitir el diagnóstico se vio reducido por el efecto de la presión temporal (tabla 9.1), mientras que el número de diagnósticos incorrectos emitidos y el número de fuentes de información observadas aumentaron con la presión temporal.

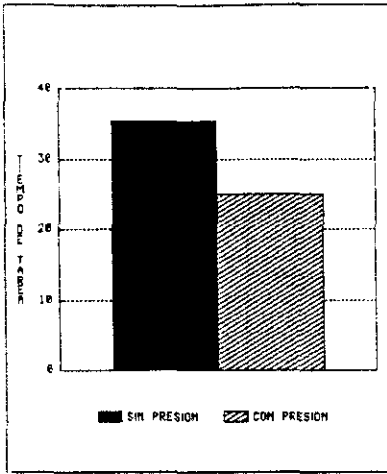


Figura 9.1. Tiempo medio de diagnóstico (sg.) con y sin presión temporal.

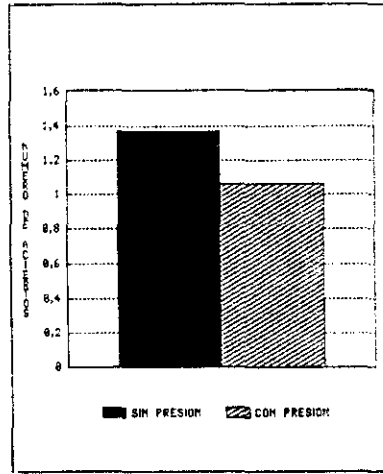


Figura 9.2. Número de aciertos medio con y sin presión temporal.

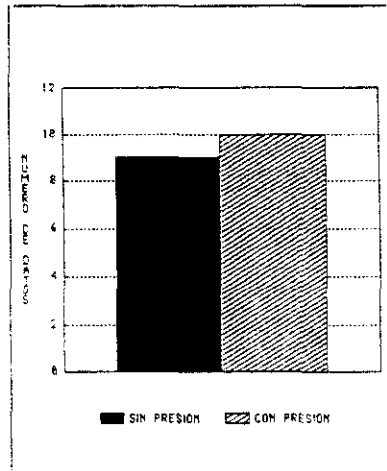


Figura 9.3. Número medio de datos observados con y sin presión temporal.

A partir de la comparación de la tabla 8.6. con la 9.6., se puede observar que la disminución en el tiempo de tarea y en el número de aciertos, debida a la presión temporal, tuvo lugar en todos los ensayos.

Sobre el índice global de carga, así como sobre las estimaciones de las tres dimensiones de carga, no se encontraron efectos simples, ni de interacción, significativos de la presión temporal. Sin embargo, la comparación de la tabla 8.9 y la 9.8. permite comprobar que en general los niveles de carga percibida por los sujetos fueron algo superiores bajo condiciones de presión temporal.

Debido a que la presión temporal tuvo efectos significativos sobre algunas de las variables, se realizaron MANOVA's para cada variable con los datos de los 20 sujetos que participaron en este segundo experimento. Los resultados de estos análisis aparecen a continuación.

9.3.2. Medidas de la ejecución

Al igual que en el experimento anterior y con el fin de determinar los efectos de los factores entresujetos, se realizó un MANOVA para cada una de las variables dependientes: tiempo de tarea (TT), número de datos observados (ND) y número de aciertos (NA). En ninguno de estos análisis se encontraron efectos significativos simples de los factores entresujetos (formato y tipo de feedback), ni de su interacción (ver Tabla 3. apéndice 3).

Además se realizó un MANOVA para cada una de las variables dependientes con el objetivo de determinar los efectos de los factores intrasujeto (ver Tabla 4, apéndice 3). Los resultados de este análisis se muestran a continuación.

9.3.2.1. Tiempo de tarea

Para el tiempo de tarea resultaron significativos los efectos de la complejidad perceptiva [$F(3,48) = 7.08$, $p < 0.001$] y de todas las interacciones posibles entre los tres factores de complejidad [$F(3,48) = 10.92$, $p < 0.001$, para CPxNC; $F(3,48) = 4.65$, $p < 0.01$, para CPxPM; $F(1,16) = 8.45$, $p < 0.05$, para NCxPM; $F(3,48) = 3.63$, $p < 0.05$, para CPxNCxPM].

La figura 9.4. muestra el efecto de interacción entre el número de causas posibles del problema y la complejidad perceptiva del display. Al igual que en el anterior experimento, en general los problemas con mayor número de causas posibles fueron resueltos más rápidamente.

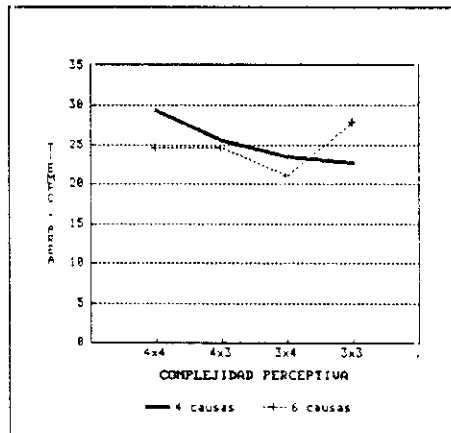


Figura 9.4. Interacción entre la complejidad perceptiva y el número de causas posibles del problema para el tiempo de tarea (sg.).

La figura 9.5. muestra la interacción entre la complejidad perceptiva y la probabilidad de ocurrencia de la solución. Al igual que en el experimento anterior y como se había previsto, a medida que aumenta la probabilidad de ocurrencia de la solución los sujetos necesitaron menos tiempo para realizar la tarea.

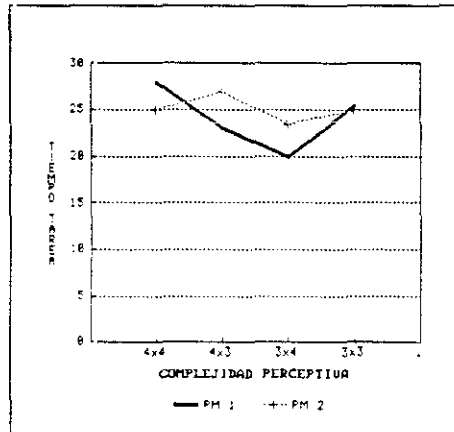


Figura 9.5. Interacción entre la complejidad perceptiva y la probabilidad de ocurrencia de la solución, para el tiempo de tarea (sg.). PM 1= Prob. alta; PM 2= Prob. baja.

La figura 9.6. muestra la interacción entre los dos factores de complejidad de solución del problema (número de causas posibles del problema y probabilidad de ocurrencia de la solución). Debido a que la disminución en el tiempo de diagnóstico encontrada para los problemas con seis causas se debe principalmente a la mayor probabilidad de ocurrencia de las soluciones más frecuentes de estos frente a las de los problemas con cuatro posibles causas, la interacción entre ambos factores indicaría que los dos hacen referencia al mismo tipo de dificultad. Es decir, en realidad ambos serían el mismo factor.

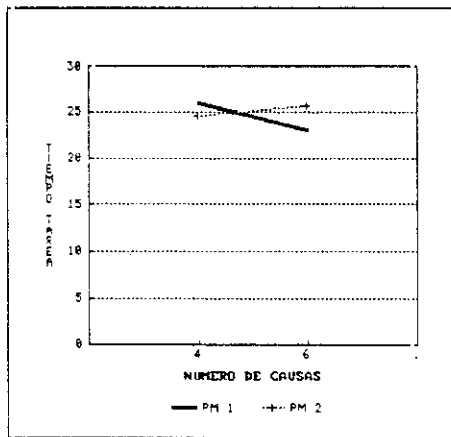


Figura 9.6. Interacción entre el número de causas posibles del problema y la probabilidad de ocurrencia de la solución, para el tiempo de tarea (sg.). PM1 = Prob. alta; PM2= Prob. baja.

Se realizaron comparaciones múltiples de medias en el tiempo de tarea de las combinaciones CPxNCxPM, por el método de Scheffé ($S=3.516$, $p<0.05$). La tabla 9.2. muestra el valor del estadístico T para los contrastes que resultaron significativos.

	141	142	161	162	241	242	261	262	341	342	361	362	441	442	461
141															
142	4.5														
161	5.7	NS													
162	4.2	NS	NS												
241	6.1	NS	NS	NS											
242	5.1	NS	NS	NS	NS										
261	2.3	3.8	2.6	4.1	2.3	3.2									
262	5.7	NS	NS	NS	NS	NS	2.6								
341	5.1	NS	NS	NS	NS	NS	3.2	NS							
342	3.7	NS	NS	NS	2.3	NS	4.6	NS	NS						
361	6.4	NS	NS	NS	NS	NS	NS	NS	NS	2.7					
362	3.5	NS	2.2	NS	2.6	NS	4.9	2.2	NS	NS	2.9				
441	6.1	NS	NS	NS	NS	NS	2.3	NS	NS	2.3	NS	2.6			
442	6.1	NS	NS	NS	NS	NS	2.2	NS	NS	2.2	NS	2.6	NS		
461	2.9	NS	2.8	NS	3.1	2.2	5.4	3.8	2.2	NS	3.5	NS	3.1	3.2	
462	3.4	NS	2.3	NS	4.7	NS	4.3	2.3	NS	NS	3.1	NS	2.3	2.7	NS

Tabla 9.2. Valores del estadístico T de Scheffé para cada contraste. El primer valor de cada combinación representa el nivel de la complejidad perceptiva (1=4x4; 2=3x4; 3=4x3; 4=3x3), el segundo representa el número de causas del problema, y el tercero la probabilidad de la solución (1=PM alta; 2=PM baja). Los valores fueron significativos al menos para $p<0.05$. NS= No Significativo.

La tabla 9.2 permite observar que en condiciones de presión temporal el número de comparaciones de medias significativas es menor (ver tabla 8.1). Sin embargo, el patrón de resultados es similar en ambas condiciones (con y sin presión temporal). Al igual que en el anterior experimento, la combinación 141 fue la que, en términos de tiempo de ejecución, resultó significativamente mucho más difícil a los sujetos. Otra combinación que parece destacar por su nivel de dificultad frente a los demás tipos de problemas es la 261.

Igualmente, se realizaron comparaciones de medias entre los tiempos de la combinación CPxNC, por el método de Scheffé ($S=3.6351$, $p<0.05$). La tabla 9.3 muestra los valores del estadístico T que resultaron significativos. En este caso, los problemas con cuatro posibles causas representados por el display con 16 elementos resultaron significativamente más difíciles que los demás.

	14	16	24	26	34	36	44	46
14								
16	2.6							
24	3.2	NS						
26	4.6	NS	NS					
34	2.1	NS	NS	2.5				
36	2.6	NS	NS	NS	NS			
44	3.7	NS	NS	NS	NS	NS		
46	NS	NS	2.4	3.7	NS	NS	2.8	

Tabla 9.3. Valor del estadístico T de Scheffé para cada contraste significativo entre los tiempos de tarea de cada combinación formada por la complejidad perceptiva y el número de causas del problema. NS=No Significativo.

La tabla 9.4. muestra los valores del estadístico T de Scheffé que fueron significativos en la comparación de los tiempos medios de tarea de las combinaciones CPxPM ($S=4.01$, $p<0.05$).

	11	12	21	22	31	32	41	42
11								
12	NS							
21	3.8	2.5						
22	2.2	NS	NS					
31	2.5	NS	NS	NS				
32	NS	NS	3.1	NS	NS			
41	NS	NS	2.4	NS	NS	NS		
42	NS	NS	2.1	NS	NS	NS	NS	

Tabla 9.4. Valor del estadístico T de Scheffé para cada contraste significativo entre los tiempos de tarea de cada combinación formada por la complejidad perceptiva y la probabilidad de la solución. NS=No Significativo.

Las comparaciones realizadas entre los tiempos de tarea de las combinaciones NCxPM no resultaron significativas. Para $p<0.05$ la diferencia crítica (S) de los contrastes fue igual a 4.04. Todas las diferencias fueron menores a este valor.

9.3.2.2. Aciertos

En cuanto al número de aciertos, sólo resultó significativo el efecto simple de la probabilidad de ocurrencia de la solución [$F(1,16)= 25.98, p<0.001$].

La tabla 9.6. muestra como a medida que aumenta la probabilidad de ocurrencia de la solución se produce un mayor número de diagnósticos correctos.

La tabla 9.5. muestra el valor del estadístico T de Scheffé para los contrastes significativos entre las medias del número de aciertos de las combinaciones CPxNC ($S=0.397, p<0.05$). Se observa que las únicas diferencias significativas se encuentran en dos de los contrastes en los que interviene la media de aciertos de los problemas con cuatro posibles causas y representados con 16 elementos en el display.

	14	16	24	26	34	36	44	46
14								
16	NS							
24	NS	NS						
26	2.3	NS	NS					
34	NS	NS	NS	NS				
36	NS	NS	NS	NS	NS			
44	2.1	NS	NS	NS	NS	NS		
46	NS	NS	NS	NS	NS	NS	NS	

Tabla 9.5. Valor del estadístico T de Scheffé para cada contraste significativo entre el número de aciertos medio de cada combinación formada por la complejidad perceptiva y el número de causas del problema. NS=No Significativo.



9.3.2.3. Número de datos observados

En cuanto al número de datos o fuentes de información observados por los sujetos, resultaron significativos los efectos de la complejidad perceptiva [$F(3,48)=6.61, p<0.005$] y de los dos factores de complejidad de solución del problema [$F(1,16)=8.28, p<0.05$, para el número de causas y $F(1,16)=36.46, p<0.001$, para la probabilidad media de la solución]. Además también resultó significativa la interacción entre los dos factores de dificultad de la solución del problema [$F(1,16)=8.95, p<0.01$]. La figura 9.7. muestra la interacción entre ambos factores.

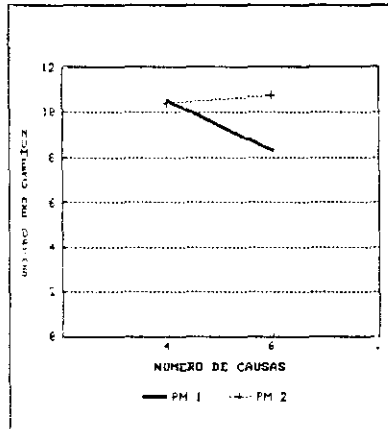


Figura 9.7. Interacción entre el número de causas posibles del problema y la probabilidad de ocurrencia de la solución, para el número de datos observados. PM1= Prob. alta; PM2= Prob. baja.

Se realizaron comparaciones de las medias del número de datos observados de las combinaciones entre el número de causas posibles del problema y la probabilidad media de la solución, por el método de Scheffé ($S=2.193$, $p<0.05$). Solamente resultó significativa la diferencia entre la media de los problemas con 6 causas y cuya solución era la más probable (combinación 61) y la media de los problemas con 6 causas y cuya solución era la menos probable (combinación 62) [$T=2.33$].

En la tabla 9.6. aparecen las medias y desviaciones típicas de las tres medidas de la ejecución utilizadas.

	TT		NA		ND	
	Media	ST.	Media	ST.	Media	ST.
TOTAL	24.91	8.27	1.06	.70	9.97	3.59
CP 4x4	26.99	9.22	.95	.67	10.53	4.31
NC 4	29.36	8.95	.80	.72	10.98	4.58
PM 1	33.32	7.82	.95	.82	9.97	5.17
PM 2	25.39	8.37	.65	.58	12.00	3.76
NC 6	24.62	8.97	1.10	.59	10.09	4.03
PM 1	23.32	9.55	1.20	.69	8.97	3.19
PM 2	25.91	8.40	1.00	.45	11.20	4.53
CP 3x4	22.28	7.75	1.08	.65	9.56	3.02
NC 4	23.54	7.51	.92	.69	9.75	2.96
PM 1	22.73	8.20	1.00	.72	10.00	3.19
PM 2	24.55	6.87	.85	.67	9.50	2.77
NC 6	21.03	7.87	1.25	.58	9.37	3.11
PM 1	18.71	7.46	1.45	.60	7.80	2.60
PM 2	23.35	7.77	1.05	.51	10.95	2.81
CP 4x3	25.10	7.84	1.06	.71	10.55	3.77
NC 4	25.56	7.31	1.07	.72	11.21	3.75
PM 1	24.39	7.39	1.25	.63	10.95	3.45
PM 2	26.74	7.23	.90	.78	11.47	4.11
NC 6	24.64	8.41	1.05	.71	9.90	3.73
PM 1	22.05	7.61	1.35	.74	8.77	3.40
PM 2	27.24	8.56	.75	.55	11.02	3.77
CP 3x3	25.26	7.64	1.13	.75	9.23	2.98
NC 4	22.69	6.13	1.22	.73	9.58	3.17
PM 1	22.76	6.27	1.35	.67	10.17	3.46
PM 2	22.64	6.15	1.10	.78	9.00	2.82
NC 6	27.82	8.19	1.05	.78	8.87	2.76
PM 1	28.23	8.99	1.40	.75	7.87	2.02
PM 2	27.42	7.53	0.70	.65	9.87	3.07

Tabla 9.6. Medias y desviaciones típicas de las tres medidas de ejecución en cada una de las combinaciones intrasujeto. TT=Tiempo; NA=Número de aciertos; ND=Número de datos; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad de ocurrencia de la solución (1=Alta; 2=Baja).

9.3.3. Análisis de la carga subjetiva

En primer lugar, se realizó un análisis de varianza con los dos factores entresujetos sobre el índice global de carga y sobre las estimaciones de cada dimensión de carga. En este análisis no resultaron significativos ninguno de los efectos simples, ni el de interacción (ver Tabla 5, apéndice 3).

A continuación, y siguiendo el mismo procedimiento que en el experimento anterior, se determinaron los efectos de los factores intrasujetos que inciden en la carga subjetiva (ver Tabla 6, apéndice 3). Los resultados de este análisis aparecen a continuación.

9.3.3.1. Índice global de carga subjetiva

En contra de lo observado en el experimento anterior, el índice global de carga solo resultó afectado por el factor de complejidad perceptiva [$F(3,48)= 3.23, p<0.05$].

La tabla 9.8. permite observar que a medida que aumenta el número de elementos del display se producen mayores niveles globales de carga subjetiva.

9.3.3.2. Dimensiones de carga

En este caso, únicamente resultó significativo el efecto de la probabilidad de ocurrencia de la solución, para la demanda mental [$F(1,16)= 7.56, p<0.05$].

La demanda temporal sólo resultó afectada por la complejidad perceptiva [$F(3,48)= 6.33, p<0.005$], y por la triple interacción entre los factores [$F(3,48)= 6.53, p<0.005$].

Se realizaron comparaciones de medias de las estimaciones de la demanda temporal entre las combinaciones formadas por CPxNCxPM, por el método de Scheffé ($S=8.6913$, $p<0.05$). La tabla 9.7. muestra el valor del estadístico T para las comparaciones que resultaron significativas.

	141	142	161	162	241	242	261	262	341	342	361	362	441	442	461
141															
142	3.2														
161	2.8	NS													
162	NS	3.1	2.7												
241	2.3	NS	NS	2.2											
242	3.8	NS	NS	3.6	NS										
261	2.7	NS	NS	2.6	NS	NS									
262	3.2	NS	NS	3.1	NS	NS	NS								
341	2.5	NS	NS	2.3	NS	NS	NS	NS							
342	NS	3.8	NS	NS	2.9	4.4	3.4	3.9	3.1						
361	3.9	NS	NS	3.8	NS	NS	NS	NS	NS	4.5					
362	3.8	NS	NS	3.6	NS	NS	NS	NS	NS	4.4	NS				
441	2.7	NS	NS	2.5	NS	NS	NS	NS	NS	3.3	NS	NS			
442	2.4	NS	NS	2.2	NS	NS	NS	NS	NS	2.9	NS	NS	NS		
461	4.2	NS	NS	4.1	NS	NS	NS	NS	NS	4.8	NS	NS	NS	NS	
462	2.2	NS	NS	NS	NS	NS	NS	NS	NS	2.3	NS	NS	NS	NS	2.1

Tabla 9.7. Valores del estadístico T de Scheffé para cada contraste. El primer valor de cada combinación representa el nivel de la complejidad perceptiva (1=4x4; 2=3x4; 3=4x3; 4=3x3), el segundo representa el número de causas del problema, y el tercero la probabilidad de la solución (1=PM alta; 2=PM baja). Los valores fueron significativos al menos para $p<0.05$. NS= No Significativo.

La demanda física sólo resultó afectada por el efecto simple de la probabilidad de ocurrencia de la solución [$F(1,16)= 6.63$, $p<0.05$].

La tabla 9.8. permite observar que a medida que la probabilidad de ocurrencia de la solución es más alta, las estimaciones de la demanda mental y la física disminuyen. Además, a medida que aumenta la complejidad perceptiva, la estimación de la demanda temporal toma valores más altos.

	DM		DT		DF		INDICE	
	Media	ST.	Media	ST.	Media	ST.	Media	ST.
TOTAL	58.79	20.08	53.37	21.47	40.04	23.80	57.08	19.40
CP 4x4	60.79	21.31	57.27	22.41	40.87	23.48	60.28	20.37
NC 4	58.81	22.42	57.07	23.31	41.37	23.39	59.58	21.04
PM 1	57.35	20.65	64.02	22.77	44.04	21.69	62.72	18.88
PM 2	60.27	24.51	50.12	22.24	38.70	25.25	56.45	22.57
NC 6	62.76	20.24	57.47	21.76	40.37	23.86	60.99	19.66
PM 1	57.81	22.59	51.62	23.10	37.75	23.23	55.72	21.57
PM 2	67.72	16.69	63.33	19.13	42.98	24.80	66.27	15.89
CP 3x4	59.40	19.90	50.91	21.44	39.13	24.71	56.08	19.49
NC 4	59.85	20.06	50.75	21.34	38.87	25.48	56.50	19.28
PM 1	59.68	20.09	53.77	20.59	37.46	22.77	58.44	18.94
PM 2	60.01	20.55	47.74	22.18	40.27	28.46	54.56	19.42
NC 6	58.96	19.99	51.06	21.80	39.40	24.24	55.66	19.69
PM 1	58.45	22.60	52.17	22.65	39.12	24.23	55.80	20.71
PM 2	59.48	18.31	49.96	21.45	39.69	24.87	55.53	18.62
CP 4x3	57.94	19.10	53.62	21.72	40.63	23.81	56.47	18.60
NC 4	60.63	18.32	59.98	20.09	42.59	21.55	60.74	17.15
PM 1	58.48	20.61	53.27	21.03	39.00	23.75	57.16	19.60
PM 2	62.78	15.94	66.68	17.07	46.19	19.02	64.33	13.37
NC 6	55.25	19.71	47.27	21.66	38.67	26.00	52.21	19.02
PM 1	50.91	20.65	46.81	21.11	35.92	24.74	49.49	18.63
PM 2	59.60	18.20	47.73	22.74	41.42	27.57	54.94	19.02
CP 3x3	57.04	20.11	51.66	20.10	39.54	23.58	55.52	18.73
NC 4	56.29	20.10	53.11	20.65	38.99	23.12	55.53	18.61
PM 1	55.78	15.08	52.38	18.88	37.23	23.37	54.64	14.89
PM 2	56.80	24.53	53.83	21.88	40.76	23.33	56.43	21.68
NC 6	57.78	20.34	50.21	19.70	40.08	24.31	55.51	18.86
PM 1	54.14	25.49	45.76	23.56	37.75	25.68	51.64	23.41
PM 2	61.43	13.10	54.66	14.12	42.41	23.28	59.39	11.56

Tabla 9.8. Medias y desviaciones típicas de las tres dimensiones de carga en cada una de las combinaciones intrasujeto. DM=Demanda mental; DT=Demanda temporal; DF=Demanda física; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad de ocurrencia de la solución (1=Alta; 2=Baja).

9.3.4. Diferencias individuales

Para examinar el efecto de las diferencias individuales sobre las medidas de ejecución, se realizó un análisis de regresión múltiple, para cada una de las variables dependientes, siguiendo el método Stepwise. En este análisis se incluyeron como posibles variables predictoras la puntuación en el Test de Figuras Ocultas (FO), la obtenida en la prueba de razonamiento del PMA (R), y las tres variables de personalidad del cuestionario EPI, neuroticismo (NEU), extroversión (EXT) y sinceridad (SIN) (ver Tablas 7 y 8, apéndice 3).

Para el tiempo de tarea, entraron a formar parte de la ecuación la puntuación en razonamiento [T= -2.762, $p<0.01$] y la extroversión [T= 3.164, $p<0.005$; $R^2= 0.04970$; R^2 ajustada= 0.04371].

En cuanto a la variable criterio "número de datos observados", únicamente resultó como variable predictora la puntuación en razonamiento [T= -2.306, $p<0.05$; $R^2= 0.01645$; R^2 ajustada= 0.01335].

Ninguna de las variables resultó predictora del número de aciertos.

	FO	R	NEU	EXT	SIN
TT	.0220	-.1493*	.0084	.1638*	.0341
ND	-.0535	-.1282	.0247	-.0112	.0370
NA	.0643	.0726	-.0267	-.0363	-.0360

Tabla 9.9. Matriz de correlaciones entre las variables predictoras y cada una de las medidas de la ejecución (TT=Tiempo de tarea; ND=Número de datos observados; NA=Número de aciertos). Un * indica significación para $p<0.01$, y dos ** para $p<0.001$.

La tabla 9.9. muestra que solamente se obtuvieron correlaciones significativas, aunque no elevadas, entre el neuroticismo, el razonamiento y el tiempo requerido para realizar la tarea. El signo de las correlaciones indica que los sujetos introvertidos con alta capacidad de razonamiento necesitarían menos tiempo para realizar la tarea.

9.3.5. Feedback como factor de aprendizaje

El análisis de la figura 9.8b muestra que, al igual que en el experimento anterior, no se puede hablar del tipo de feedback como factor de aprendizaje. Sin embargo, en este caso, parece observarse una tendencia no muy clara a mejorar la ejecución a medida que discurren los ensayos, cuando el feedback sólo informaba de que la respuesta era o no correcta (figura 9.8a), mientras que el informar a los sujetos de cual era la respuesta correcta produciría una disminución en el número de aciertos. De todas formas esta tendencia es muy poco clara y no significativa. De igual forma, permite observar que no hubo diferencias en el número de aciertos en función del tipo de feedback, como ya se puso de manifiesto en el MANOVA realizado.

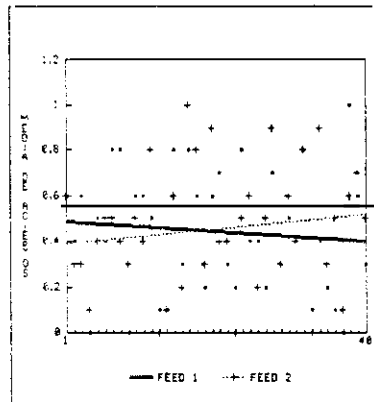


Figura 9.8a. Representación de la tendencia de los datos mostrados en la figura 9.8b.

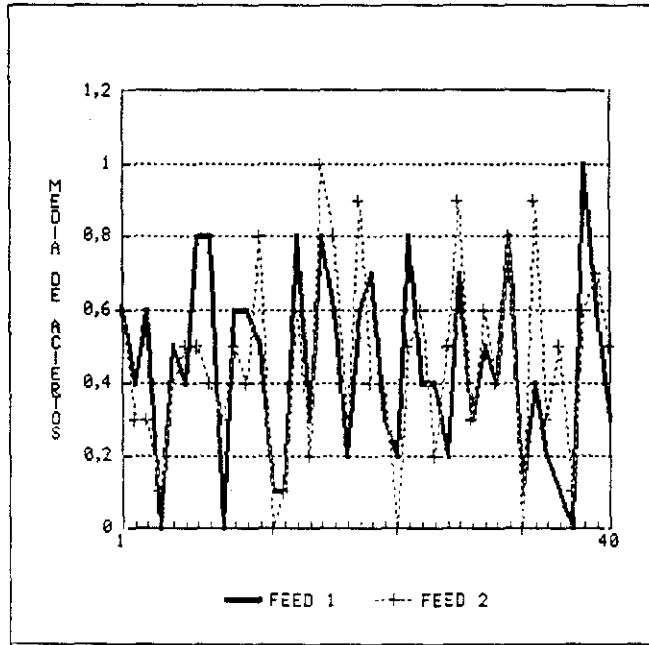


Figura 9.8b. Número medio de aciertos en cada uno de los ensayos, según el tipo de feedback proporcionado. Feed 1= Respuesta correcta; Feed 2= correcto/no correcto.

9.4. DISCUSION

En este experimento se pretendía, principalmente, analizar los efectos de la presión temporal sobre la ejecución y sobre la carga. En este sentido, los resultados indicaron que tanto el tiempo de tarea como el número de aciertos, principales medidas de la ejecución, se veían afectados por el factor de presión temporal. Sin embargo, el efecto de este factor no fue significativo para el número de datos observados. Por otro lado, las medidas de carga no se vieron significativamente afectadas por la inclusión de la presión temporal.

Se encontraron interacciones altamente significativas entre la presión temporal y los factores de complejidad (perceptiva y de solución del problema), sobre las medidas de la ejecución. Este resultado, unido a los encontrados para las variables de complejidad perceptiva y de solución del problema, confirma la hipótesis relativa a la visión de la presión temporal como un factor más de dificultad de la tarea, pero no como la única ni la mayor fuente de complejidad.

Además, aunque los niveles globales de carga subjetiva fueron generalmente superiores bajo condiciones de presión temporal, la diferencia no fue significativa. Por contra, los factores que tuvieron una mayor incidencia sobre la carga fueron aquellos relativos a la complejidad, tanto perceptiva como de solución del problema. Por tanto, en contra de la opinión de algunos autores como Senders (1979), Phillip, Reiche y Kirchner (1971), Schmidt (1978) o Hacker, Plath, Richter y Zimmer (1978), la carga percibida por los sujetos no está esencialmente determinada por la presión temporal. De forma que la evaluación de la carga subjetiva asociada a la tarea debería tener en cuenta además otros aspectos diferentes a los puramente temporales (Moray, 1982).

En cuanto a la ejecución, parecen confirmarse las previsiones. El tiempo necesitado por los sujetos para emitir su diagnóstico fue significativamente menor en prácticamente todos los ensayos (ver tabla 9.1) , sin embargo la presión temporal aumentó el número de diagnósticos incorrectos emitidos. Además, el número de datos observado por los sujetos aumentó con la presión temporal, aunque no

significativamente. Estos resultados confirman que la probabilidad de localizar el componente que falla decrece con el tiempo disponible, debido a la mayor dificultad para interpretar los valores de las variables que representan el estado del sistema (Ben Zur y Breznitz, 1981; Wohl, 1981; Zakay y Wooler, 1984).

Por otro lado, el tiempo necesario para realizar la tarea y el número de datos observados, aumentaron con el número de elementos del display, aunque esto no originó un mayor número de aciertos. Al igual que en el experimento anterior, estos resultados indican que los sujetos utilizaron una estrategia topográfica de búsqueda de información. Si la elección del tipo de estrategia está guiada por el criterio de mínima resistencia, el cual hace que sujetos inexpertos elijan la estrategia topográfica frente a la sintomática para reducir los niveles de carga (Rasmussen, 1981; Rouse, 1981), tiene sentido pensar que bajo presión temporal el deseo de reducir los niveles de carga será aun más fuerte, y por tanto la elección de la estrategia topográfica será más frecuente.

Sobre el tiempo necesitado para emitir el diagnóstico no se encontraron efectos simples significativos de los factores de complejidad de solución del problema, los cuales si fueron significativos en el experimento anterior, pero si de la complejidad perceptiva (Henneman y Rouse, 1984; Morris y Rouse, 1985). Este resultado refuerza la idea de que bajo condiciones de presión temporal la preferencia por el uso de la estrategia topográfica aumenta.

La utilización de la estrategia topográfica genera un alto número de errores, que aumenta bajo condiciones de presión temporal. Los sujetos buscan la mayor cantidad de información, aunque esta sea totalmente irrelevante para el diagnóstico, y no dedican el tiempo suficiente para procesarla. Este hecho ha sido puesto de manifiesto en los estudios realizados sobre diversos accidentes ocurridos en la realidad (por ejemplo en Three Mile Island) como el factor de error humano más frecuente en situaciones de emergencia y alto riesgo, que requieren que el operador emita un diagnóstico correcto del sistema en un período de tiempo muy breve (Maddox y Brickery, 1982).

En este experimento, las medidas de diferencias individuales fueron aun menos predictoras de la ejecución que en el anterior. Sólo se encontraron correlaciones pequeñas pero significativas entre el nivel de razonamiento y la extroversión con el tiempo requerido para realizar la tarea (ver tabla 9.9). En cuanto al estilo cognitivo DIC, parecen confirmarse los resultados encontrados en el experimento anterior. En resumen, estos resultado parecen indicar que bajo presión temporal estas características de los individuos son aún menos relevantes.

En relación a la carga subjetiva, en principio no se encontraron diferencias significativas debidas a la presión temporal. Sin embargo, se encontraron factores significativos diferentes a los del experimento anterior. En este sentido, mientras que en el experimento anterior el índice global de carga subjetiva resultó únicamente afectado por la complejidad de solución del problema (establecida en función de la probabilidad de ocurrencia de la solución), ahora, bajo condiciones de presión temporal, fue la complejidad perceptiva la que tuvo un mayor efecto sobre la carga global percibida por los sujetos (Sperandio, 1978). Este cambio no resulta sorprendente si tenemos en cuenta que cuando los sujetos utilizan mayoritariamente la estrategia topográfica, la carga está fundamentalmente determinada por el número de elementos del display.

En resumen, parece que bajo condiciones de presión temporal, se afianza el uso de la estrategia de búsqueda topográfica, y que a medida que el sujeto se encuentra con más fuentes de información, observa un mayor número de datos relativos al funcionamiento del sistema (aunque sean totalmente irrelevantes para el diagnóstico y para su confirmación), pero sin poder dedicar el tiempo suficiente al procesamiento de los mismos, lo que se traduce en un aumento de los errores y de los sentimientos de carga.

CAPITULO 10. EFECTOS DE LA COMPLEJIDAD DEL
PROBLEMA Y DE LA PRESENTACION VERBAL DE LA
INFORMACION SOBRE LA CARGA.

10.1. INTRODUCCION

En diversas investigaciones se han discutido los efectos sobre la ejecución y la carga del tipo de código de la información presentada. La mayor parte de ellas distinguen entre información numérica y gráfica pero muy pocas (Boles y Wickens, 1987), dentro de la Ergonomía, comparan los displays numéricos con los verbales.

Por otro lado, autores como Gallwey (1982) o Morrison y Duncan (1988), entre otros, han mostrado como frente a medidas de capacidad general, de todas las características o habilidades individuales, las que tienen una mayor relación con la ejecución son aquellas que se relacionan directamente con la tarea específica con la que se enfrentan los sujetos.

En este sentido, dado que no se encontraron importantes efectos de las variables individuales utilizadas en los experimentos anteriores, se pensó que, debido a que la tarea requería una respuesta manual y el uso de un ratón, quizás la habilidad manual necesaria para manejar correctamente este aparato podía tener efectos sobre el nivel de ejecución y, por tanto sobre los resultados encontrados.

Por estas razones, en este experimento las medidas de razonamiento, DIC y personalidad de los anteriores, fueron sustituidas por una medida de la habilidad manual inicial de los sujetos con respecto al manejo de este instrumento.

Los objetivos principales que guían este experimento son los siguientes:

1. Analizar las posibles diferencias en ejecución y carga percibida debidas al cambio en el tipo de información presentada. En anteriores experimentos la información era numérica, mientras que en este fue verbal.
2. Establecer en qué medida la ejecución está afectada por las diferencias individuales en la habilidad manual requerida para el manejo del ratón.



En relación a estos objetivos se pueden establecer las siguientes hipótesis:

1. No se espera encontrar diferencias en la ejecución debidas al cambio de información numérica a verbal.
2. Se espera que principalmente la medida global de carga subjetiva y la estimación de la demanda mental se vean afectadas por el cambio en el tipo de información. Debido a la escasa formación técnica de los sujetos que participaron en este experimento, se espera que la información verbal disminuya los niveles de carga percibida.
3. Se espera que la habilidad de manejo del ratón entre en la ecuación de regresión para predecir la ejecución, con mayor poder predictivo que las medidas de nivel de razonamiento, DIC y personalidad, utilizadas en los experimentos anteriores.

10.2. METODO

10.2.1. Sujetos

En este experimento participaron 15 sujetos estudiantes de Psicología, con edades comprendidas entre los 20 y los 25 años, de los cuales 4 fueron varones y 11 mujeres. Todos ellos recibieron un aumento de 0,25 puntos en la nota final de una de sus asignaturas del curso.

10.2.2. Diseño

El diseño fue un factorial 2x2x2 de medidas repetidas. Los tres factores utilizados fueron los mismos que en los experimentos anteriores, pero los cuatro niveles del factor de complejidad perceptiva fueron reducidos a 2 (4x3 vs. 3x3). La variable irrelevante para el diagnóstico (número de reparaciones) fue eliminada de todos los ensayos.

Las variables de DIC, personalidad y razonamiento fueron sustituidas por el tiempo tardado en realizar un test de movimientos y los errores cometidos durante el mismo.

Las variables dependientes utilizadas en los análisis, fueron las mismas que en los experimentos anteriores.

10.2.2. Tarea y Procedimiento

Básicamente la tarea y el procedimiento fueron los mismos que en los experimentos anteriores. Sin embargo, y debido a los cambios en el diseño, los ensayos que formaban la prueba fueron reducidos a 20. Además, en este experimento se utilizó información de tipo verbal para representar el estado del sistema. Los valores numéricos de los experimentos anteriores fueron sustituidos por las palabras: nulo/a, bajo/a, medio/a, alto/a y excesivo/a.

Como en el experimento anterior, los sujetos dispusieron de 30 segundos para realizar cada ensayo.

Los tests de DIC, personalidad y razonamiento fueron sustituidos por una prueba de manejo del ratón. Esta era una prueba en la que únicamente aparecía sobre la pantalla del ordenador un matriz de 16 celdillas (4x4) totalmente vacías. En primer lugar el sujeto debía ir pasando por dentro de las celdillas (abriéndolas), primero en vertical y después en horizontal. Seguidamente, el sujeto debía hacer un recorrido con el ratón similar al anterior, pero en este caso pasando el cursor entre casillas, sin que estas se abrieran (primero en vertical y después en horizontal). Como medidas de la ejecución en esta prueba se registraron el tiempo total necesitado por el sujeto para completar estos recorridos, así como el número de errores que había cometido.

10.2.4. Instrumentos

Para la presentación de las instrucciones y demás aspectos del experimento, así como para la realización de la prueba de manejo del ratón, se utilizó un ordenador PC/XT con disco duro y un ratón GM-6 PC Mouse. Al igual que en las demás ocasiones, el experimento y la prueba de manejo del ratón fueron generados y controlados mediante el programa Mouselab System versión 4.2 (Johnson, Payne, Schkade y Bettman, 1989), el cual también se encargó de la recogida de los datos.

10.3. RESULTADOS

10.3.1. Comparación entre tipos de información

En primer lugar, y con el fin de determinar las posibles diferencias sobre la ejecución y sobre la carga subjetiva de los dos tipos de información (numérica y verbal), se realizó un MANOVA con los datos de este experimento más los de 15 sujetos seleccionados aleatoriamente del experimento anterior ($N=30$), para cada una de las variables (ver Tablas 1 y 2, apéndice 4).

El tipo de información no tuvo ningún efecto significativo (ni simple, ni de interacción) sobre las medidas de ejecución (tiempo de tarea, número de aciertos y número de datos observados). Igualmente, no se encontró ningún efecto significativo de la interacción entre el tipo de información y los factores intrasujeto (complejidad perceptiva, número de soluciones posibles del problema y probabilidad media estimada de la ocurrencia de la solución).

Sin embargo, la figura 10.1 muestra como en todos los casos, excepto para la combinación 442, el tiempo requerido para realizar la tarea fue menor cuando la información era verbal. Aunque esta diferencia no fue estadísticamente significativa, parece haber una clara tendencia a la reducción del tiempo de tarea cuando el input de información es verbal.

La figura 10.2. muestra que el número de aciertos no varió sistemáticamente en función del tipo de información presentada, sino que las variaciones en el número de aciertos se debieron a los factores de complejidad de la tarea. Por tanto, ambos tipos de información proporcionaron, en general, similares niveles de fiabilidad humana. Este mismo resultado aparece en la figura 10.3 para el número de datos observados.

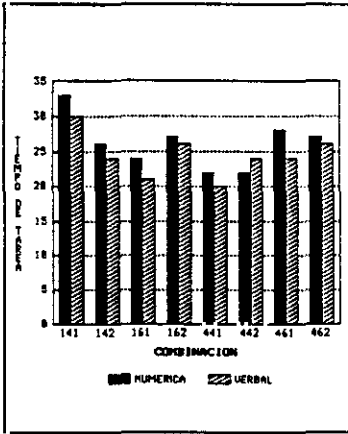


Figura 10.1. Tiempo medio de tarea (sg.) en cada combinación según el tipo de información presentada.

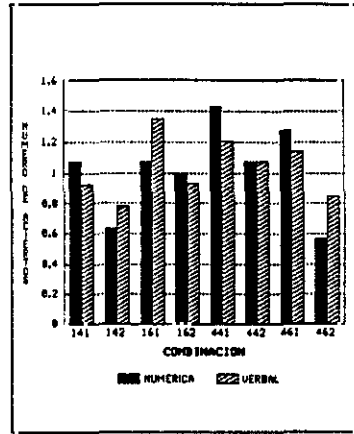


Figura 10.2. Número medio de aciertos en cada combinación según el tipo de información presentada.

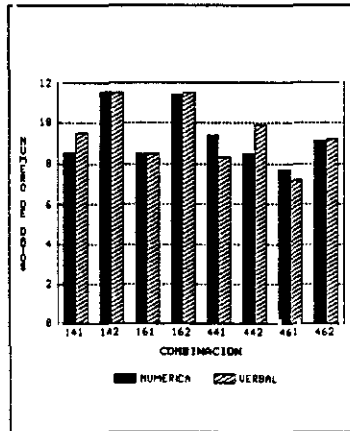


Figura 10.3. Número medio de datos observados en cada combinación según el tipo de información presentada.

En cuanto a las medidas de carga, el efecto simple del tipo de información resultó significativo sobre el índice global de carga [$F(1,28)= 4.57, p<0.05$], y sobre la estimación subjetiva de la demanda temporal de la tarea [$F(1,28)= 4.37, p<0.05$].

El efecto de interacción entre el tipo de información y el número de causas posibles de problema resultó significativo para la estimación de la demanda mental [$F(1,28)= 4.78, p<0.05$]. La estimación de la demanda física de la tarea no se vio afectada por el cambio en el tipo de información.

Las figuras 10.4, 10.5, 10.6 y 10.7, muestran como a pesar de que en algunos casos, como para la estimación de la demanda mental o de la física, no se encontraron efectos simples significativos del tipo de información, en todos los casos se encontraron mayores niveles medios de carga cuando la información sobre el estado del sistema se presentaba en forma numérica.

Por otro lado, la figura 10.8, muestra como el efecto del tipo de información varía en función de los niveles del factor de complejidad de solución del problema "número de causas posibles del problema". A partir de la figura, se puede observar, como ya se puso de manifiesto en el primer experimento, que a medida que aumenta el número de causas posibles del problema disminuye la demanda mental percibida por los sujetos. Este efecto se debe principalmente a que la probabilidad media de las soluciones de los problemas con seis posibles causas es bastante más alta que la de los problemas que tienen cuatro posibles causas.

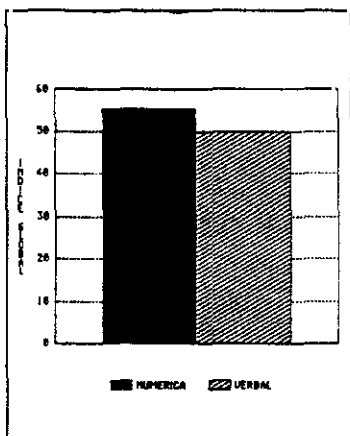


Figura 10.4. Media del índice global de carga según el tipo de información presentada.

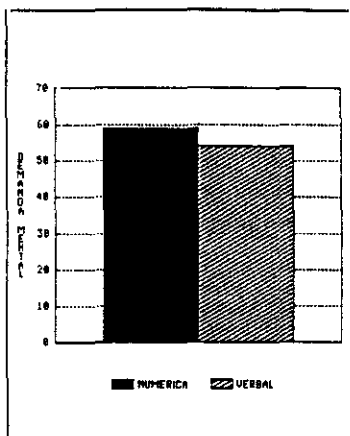


Figura 10.5. Media de las estimaciones de la demanda mental según el tipo de información presentada.

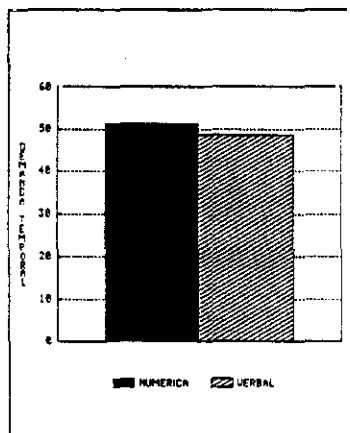


Figura 10.6. Media de las estimaciones de la demanda temporal, según el tipo de información presentada.

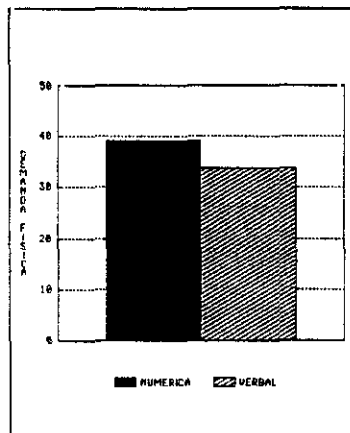


Figura 10.7. Media de las estimaciones de la demanda física, según el tipo de información presentada.

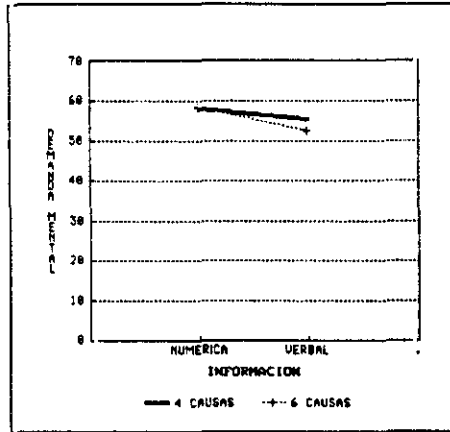


Figura 10.8. Interacción entre el número de causas posibles del problema y el tipo de información presentada.

10.3.2. Medidas de la ejecución

Debido a que no se establecieron diferencias significativas en ninguna de las medidas de la ejecución en función del tipo de información presentada, las medidas de la ejecución se vieron afectadas en general por los mismos factores que en el experimento anterior.

La tabla 10.1 muestra las medias y las desviaciones típicas de cada una de las medidas de la ejecución, en función de las diversas combinaciones de los factores de complejidad de la tarea, para este tercer experimento. Comparando la tabla 9.6 con la 10.1 se puede observar que no hubo grandes diferencias en las medidas de ejecución de ambos experimentos. Al igual que en el anterior, en este experimento el tiempo de tarea, el número de datos observados por los sujetos, y la probabilidad de cometer un

error, aumentaron a medida que crecía el número de elementos del display. Por otro lado, a medida que la probabilidad de ocurrencia de la solución aumenta, la proporción de aciertos es mayor mientras que disminuye el número de datos observados.

	TT		NA		ND	
	Media	ST.	Media	ST.	Media	ST.
TOTAL	24.72	7.94	1.04	.69	9.41	3.83
CP 3x4	25.82	7.55	1.00	.66	10.24	4.15
NC 4	27.83	7.41	.83	.64	10.44	3.59
PM 1	30.78	7.23	.93	.59	9.36	3.34
PM 2	24.89	6.54	.73	.70	11.53	3.61
NC 6	23.80	7.27	1.16	.64	10.03	4.70
PM 1	21.11	8.16	1.40	.63	8.40	3.95
PM 2	26.48	5.25	.93	.59	11.66	4.94
CP 3x3	23.63	8.22	1.08	.72	8.58	3.32
NC 4	22.18	7.21	1.16	.69	9.06	2.89
PM 1	20.23	7.85	1.26	.70	8.26	2.69
PM 2	24.13	6.15	1.06	.70	9.86	2.94
NC 6	25.08	9.01	1.00	.74	8.10	3.68
PM 1	23.76	11.80	1.13	.91	7.03	3.71
PM 2	26.40	5.03	.86	.51	9.16	3.45

Tabla 10.1. Medias y desviaciones típicas de las medidas de ejecución en cada una de las combinaciones intrasujeto.

Con el fin de establecer la significación estadística de los efectos de los factores sobre las medidas de la ejecución, se realizó un MANOVA para cada una de las variables dependientes, tomando los datos de los sujetos que participaron en este tercer experimento (N=15). Los resultados de este análisis aparecen a continuación (ver Tabla 3, apéndice 4).

10.3.2.1. Tiempo de Tarea

De forma similar a los resultados obtenidos en el anterior experimento, sobre el tiempo de tarea, resultó significativo el efecto de la complejidad perceptiva [$F(1,14)= 4.99, p<0.05$] y de las cuatro posibles interacciones entre los factores [$F(1,14)= 15.15, p<0.005$, para $CP \times NC$; $F(1,14)= 7.00, p<0.05$, para $CP \times PM$; $F(1,14)= 14.37, p<0.005$, para $NC \times PM$; $F(1,14)= 25.10, p<0.001$, para $CP \times NC \times PM$].

Se realizaron comparaciones de medias entre los tiempos medios de tarea de las distintas combinaciones $CP \times NC \times PM$ por el método de Scheffé. La diferencia crítica del contraste (S) para $p<0.05$ fue igual a 2.68. La tabla 10.2. muestra el valor de T para los contrastes que resultaron significativos. Al igual que en los anteriores experimentos, los problemas de la combinación 141 fueron los más difíciles, en términos del tiempo necesitado para emitir el diagnóstico.

	141	142	161	162	241	242	261	262
141								
142	4.7							
161	7.7	3.0						
162	3.4	NS	4.3					
241	8.4	3.7	NS	5.0				
242	5.3	NS	2.4	NS	3.1			
261	5.6	NS	NS	2.2	2.8	NS		
262	3.5	NS	4.2	NS	4.9	NS	NS	

Tabla 10.2. Valor del estadístico T de Scheffé para cada contraste significativo entre los tiempos de tarea de cada combinación formada por la complejidad perceptiva (1° valor) ($1=3 \times 4$; $2=3 \times 3$), el número de causas del problema (2°) y la probabilidad de la solución (3°) ($1=PM$ alta; $2=PM$ baja). NS=No Significativo.

Igualmente, se realizaron comparaciones entre los tiempos medios de tarea de las combinaciones formadas por NC y PM ($S=2.83$, $p<0.05$). La tabla 10.3. recoge los valores del estadístico T de Scheffé para cada contraste que resultó significativo.

	41	42	61	62
41				
42	NS			
61	2.3	NS		
62	NS	NS	3.0	

Tabla 10.3. Valor del estadístico T para cada contraste significativo entre los tiempos de tarea de cada combinación de NCxPM. NS=No Significativo.

La tabla 10.4. muestra los valores del estadístico T de Scheffé que resultaron significativos en las comparaciones de los tiempos medios de tarea en función de las combinaciones de la complejidad perceptiva y la probabilidad media de la solución ($S=2.87$, $p<0.05$).

	11	12	21	22
11				
12	NS			
21	2.9	2.8		
22	NS	NS	2.4	

Tabla 10.4. Valor del estadístico T para cada contraste significativo entre los tiempos de tarea de cada combinación de CPxPM. NS=No Significativo.

La tabla 10.5. muestra los valores del estadístico T de Scheffé que resultaron significativos en las comparaciones de los tiempos medios de tarea en función de las combinaciones de la complejidad perceptiva y el número de causas posibles del problema ($S=3.82$, $p<0.05$).

	14	16	24	26
14				
16	2.3			
24	3.2	NS		
26	NS	NS	NS	

Tabla 10.5. Valor del estadístico T para cada contraste significativo entre los tiempos de tarea de cada combinación de CPxNC. NS=No Significativo.

10.3.2.2. Aciertos

No se encontraron efectos significativos de ninguno de los factores ni de sus interacciones, sobre la medida del número de aciertos. Todo parece indicar que la no significación del factor "probabilidad ocurrencia de la solución" [$F(1,14)=3.22$, $p=0.09$], se debe únicamente a la reducción en el tamaño de las muestras con las que se realizaron los análisis de este experimento y el anterior.



10.3.2.3. Número de datos observados

En cuanto al número de datos observados por los sujetos para emitir su diagnóstico, resultó significativo el efecto de la complejidad perceptiva [$F(1,14)=30.92, p<0.001$] y de la probabilidad media de la solución [$F(1,14)=13.99, p<0.005$].

En conclusión, estos análisis permiten concluir que el patrón de resultados encontrados, para las tres medidas de ejecución con información verbal, es muy similar al encontrado utilizando información numérica. Sin embargo, la reducción en el tamaño muestral provocó que algunos efectos no fueran significativos.

10.3.3. Medidas de la carga subjetiva

Debido a que se encontraron diferencias significativas en función del tipo de información presentada, para algunas de las medidas de carga subjetiva, se realizó un MANOVA para cada una de ellas con los factores de complejidad del problema (perceptiva y de solución del problema), a partir de los datos de este tercer experimento ($N=15$). Los resultados de estos análisis se muestran a continuación (ver Tabla 4, apéndice 4).

10.3.3.1. Índice global de carga subjetiva

Este análisis reveló que el índice global de carga estaba afectado por la probabilidad media de la solución [$F(1,14)=7.35, p<0.05$], y por la interacción triple entre los factores [$F(1,14)=14.27, p<0.005$].

10.3.3.2. Dimensiones de carga subjetiva

En relación a las estimaciones de la demanda mental, resultó significativo el efecto simple de la probabilidad de la solución [$F(1,14)= 13.01, p<0.005$].

En cuanto a la demanda temporal resultó significativo el efecto de la interacción entre la complejidad perceptiva y el número de causas posibles del problema [$F(1,14)= 5.70, p<0.05$], y el de la triple interacción [$F(1,14)= 5.38, p<0.05$].

Se realizaron comparaciones entre las estimaciones medias de la demanda temporal, por el método de Sheffé, de las combinaciones CPxNC ($S=6.5834, p<0.05$). Ninguna de las diferencias resultó significativa.

La tabla 10.6. muestra los valores del estadístico T de Scheffé que resultaron significativos en la comparación de las medias de las estimaciones de la demanda temporal de cada combinación CPxNCxPM ($S=6.8465, p<0.05$).

	141	142	161	162	241	242	261	262
141								
142	NS							
161	2.8	NS						
162	NS	NS	2.5					
241	3.5	2.5	NS	3.1				
242	NS	NS	NS	NS	NS			
261	NS	NS	NS	NS	NS	NS		
262	NS	NS	NS	NS	2.6	NS	NS	

Tabla 10.6. Valor del estadístico T para cada contraste significativo entre las estimaciones de la demanda temporal de cada combinación CPxNCxPM. NS=No Significativo.

La demanda física sólo resultó afectada por el efecto de la triple interacción entre los factores [$F(1,14) = 8.86, p < 0.05$]. La tabla 10.7 muestra los valores del estadístico T de Scheffé para los contrastes de medias de la combinación CPxNCxPM que resultaron significativos ($S = 4.0335, p < 0.05$).

	141	142	161	162	241	242	261	262
141								
142	3.2							
161	2.6	NS						
162	NS	2.7	NS					
241	3.3	NS	NS	2.7				
242	2.2	NS	NS	NS	NS			
261	NS	NS	NS	NS	NS	NS		
262	NS	NS	NS	NS	NS	NS	NS	

Tabla 10.7. Valor del estadístico T para cada contraste significativo entre las estimaciones de la demanda física de cada combinación CPxNCxPM. NS=No Signif.

La tabla 10.8 muestra las medias y desviaciones típicas de las estimaciones de cada una de las dimensiones de carga (DM, demanda mental; DT, demanda temporal y DF, demanda física) y del índice global de carga subjetiva, en función de los distintos niveles de los factores intrasujeto.

	DM		DT		DF		INDICE	
	Media	ST.	Media	ST.	Media	ST.	Media	ST.
TOTAL	53.91	15.04	48.61	14.79	33.91	16.41	50.10	14.51
CP 3x4	54.42	14.73	49.92	14.73	34.42	17.18	50.95	14.27
NC 4	56.47	14.58	51.67	15.33	34.41	15.51	52.98	14.18
PM 1	52.87	14.52	53.18	15.17	37.45	14.74	52.51	13.33
PM 2	60.07	14.20	50.16	15.86	31.36	16.16	53.46	15.44
NC 6	52.37	14.84	48.17	14.16	34.44	18.97	48.91	14.30
PM 1	47.06	14.55	44.19	14.23	32.53	18.25	44.55	14.25
PM 2	57.68	13.57	52.14	13.38	36.36	20.11	53.27	13.41
CP 3x3	53.41	15.45	47.31	14.86	33.39	15.73	43.25	14.81
NC 4	54.33	16.12	45.39	14.83	32.28	17.06	48.79	15.32
PM 1	49.80	17.77	42.11	16.63	31.32	18.91	45.13	16.06
PM 2	58.85	13.35	48.67	12.49	33.24	15.59	52.45	14.12
NC 6	52.49	14.96	49.21	14.88	34.50	14.49	49.71	14.54
PM 1	50.58	17.79	47.86	18.34	34.17	14.46	48.53	17.48
PM 2	54.39	11.81	50.57	10.90	34.82	15.02	50.89	11.37

Tabla 10.8. Medias y desviaciones típicas de las medidas de carga subjetiva en cada una de las combinaciones intrasujeto.

10.3.4. Diferencias Individuales

Con el fin de determinar el efecto de las diferencias individuales iniciales en cuanto al manejo del ratón, sobre el nivel de ejecución alcanzado, se realizó un análisis de regresión múltiple, por el método Stepwise, para cada una de las medidas de ejecución, a partir del número de errores cometidos en la realización de la prueba de manejo del ratón y del tiempo tardado por los sujetos en realizarla (ver Tablas 5 a 7, apéndice 4).

Las dos variables predictoras entraron a formar parte de la ecuación de regresión para el tiempo total de tarea [$T=2.981$, $p<0.005$, para el número de errores, y $T= 4.121$, $p<0.0005$, para el tiempo; $R^2= 0.22839$; R^2 ajustada= 0.21520].

En cuanto al número de datos observados para emitir el diagnóstico, solo entró a formar parte de la ecuación de regresión el tiempo tardado en realizar la prueba de manejo del ratón [$T = -3.415$, $p < 0.001$; $R^2 = 0.08993$; R^2 ajustada = 0.08222], quedando fuera de la ecuación el número de errores cometidos en dicha prueba.

De igual forma, sólo el tiempo tardado en realizar la prueba fue predictor de la variable criterio "número de aciertos" [$T = -2.584$, $p < 0.05$; $R^2 = 0.05357$; R^2 ajustada = 0.04555].

A partir de la tabla 10.9 se observa como el tiempo necesitado para realizar la prueba de manejo del ratón correlacionó muy significativamente con todas las medidas de la ejecución, mientras que el número de errores cometidos en la prueba solo correlacionó con el tiempo de tarea. El signo de las correlaciones indica que a medida que los sujetos requerían menos tiempo para realizar la prueba de manejo del ratón y cometían menos errores en ella, necesitaban menos tiempo para realizar la tarea experimental. Por otro lado, a medida que la habilidad de manejo del ratón era menor, los sujetos emitían menor número de diagnósticos correctos y observaban un menor número de fuentes de información.

	Nº errores	Tiempo
TT	.4121**	.3412**
ND	-.1402	-.2999**
NA	-.0864	-.2315*

Tabla 10.3. Matriz de correlaciones entre las variables predictoras y las medidas de la ejecución (TT=Tiempo de tarea; ND=Número de datos observados; NA=Número de aciertos). Un * indica signif. para $p < 0.01$, y dos ** para $p < 0.001$.

10.4. DISCUSION

Como se había previsto, no se encontraron diferencias significativas en la ejecución entre el grupo de sujetos que utilizaron un display digital (numérico) y los que utilizaron el display verbal. Sin embargo, en general y no significativamente, el tiempo requerido para realizar la tarea fue menor cuando la información era verbal.

Por otro lado, el efecto del tipo de información se puso de manifiesto en la medida global de carga y en la estimación de la demanda temporal. La estimación de la demanda mental resultó afectada por el efecto de la interacción entre el tipo de información y el número de causas posibles del problema. Sin embargo, la estimación de la demanda física no resultó significativamente afectada por el tipo de información, ni por su interacción con los demás factores.

De forma muy similar al resultado encontrado en el primer experimento, el índice global de carga subjetiva resultó afectado por la probabilidad de ocurrencia de la solución y por la triple interacción entre los factores de complejidad de la tarea. Sobre la estimación de la demanda mental sólo resultó significativo el efecto de la probabilidad de ocurrencia de la solución. La estimación de la demanda temporal se vio afectada por los efectos de la complejidad perceptiva y del número de causas posibles del problema, así como por la triple interacción entre los factores, mientras que sobre la estimación de la demanda física de la tarea sólo resultó significativo el efecto de la triple interacción.

En cuanto a las diferencias individuales se encontraron correlaciones significativas entre el tiempo requerido para realizar la prueba de manejo del ratón y las tres medidas de la ejecución. Tanto el tiempo requerido para realizar la prueba de manejo del ratón como el número de errores cometidos en ella fueron predictores, y

explicaron algo menos del 23% de la varianza, del tiempo requerido para realizar la tarea de diagnóstico. Sin embargo, sólo el tiempo necesitado para completar la prueba de manejo del ratón entró en las ecuaciones de regresión para predecir el número de aciertos y el número de datos observados, y en ambas su poder predictivo fue pequeño (no llegaba a explicar el 6% y el 9% de la varianza respectivamente).

Aunque de forma no significativa, en general el tiempo de diagnóstico mejoró con la presentación verbal de la información. Sin embargo, no se cometieron menos diagnósticos incorrectos, ni parece que se modificaran en gran medida las estrategias de búsqueda de información utilizadas. Por tanto, estos resultados no permiten hablar de una mejora de la ejecución general de los sujetos cuando los inputs visuales de información son verbales.

Sin embargo, aunque la ejecución no mejoró, si se produjo una clara reducción en los niveles de carga percibida por los sujetos. Los niveles de carga subjetiva fueron muy similares a los obtenidos para el primer experimento, el cual se diferencia del segundo solamente en la limitación del tiempo disponible para realizar la tarea. Este resultado podría indicar que cuando la información se presenta en un formato verbal y bajo condiciones de presión temporal, los niveles de carga percibida por los sujetos se asemejan a los percibidos en situaciones de no presión temporal y con información de tipo numérico. Por tanto, el efecto de la presión temporal sobre la carga subjetiva quedaría compensado por la presentación verbal de la información.

En consecuencia, a la hora de elegir entre dos diseños de display alternativos, uno numérico y otro verbal, sería preferible el verbal ya que en situaciones de emergencia que requieren un diagnóstico preciso y rápido del fallo del sistema, se alcanzarán los mismos niveles de acierto pero disminuirá el tiempo requerido para el diagnóstico y los sentimientos de carga. La reducción en los niveles de carga redundará en un aumento de la aceptación del display por parte de los operadores.

Por otro lado, y como en los experimentos anteriores, el aumento en el número de elementos del display y la disminución en la probabilidad de ocurrencia de la

solución produjeron incrementos en el tiempo requerido para el diagnóstico y en el número de datos observados por los sujetos. Al igual que en los experimentos anteriores, esto indicaría que los sujetos modificaron el tipo de estrategia de búsqueda de información de una sintomática a otra topográfica, a medida que transcurría cada uno de los ensayos, con el fin de reducir la carga sobre la memoria de trabajo (Brooke y Duncan, 1981, 1983; Rouse, 1981; Rouse y Rouse, 1984; Rasmussen, 1981). Además, en este experimento también se confirmó la tendencia de los sujetos a utilizar un mayor número de información a medida que aumenta el número de fuentes de información, aunque esto no se traduzca en un mayor número de diagnósticos correctos (Dawes, 1979; Dawes y Corrigan, 1974; Hayes, 1964; Oskamp, 1965; Schroeder y Benbassat, 1975; Wright, 1974).

En cuanto a las medidas de capacidad individual, los resultados de este y anteriores experimentos confirman la idea de que de todas las habilidades individuales, aquellas que tienen una mayor relación directa con la tarea concreta que se ha de realizar, son las más predictoras de la ejecución (Demaree, Crowder y Morrison, 1955; Gallwey, 1982; Highland y col., 1956; Morrison y Duncan, 1988). Mientras que, en los experimentos anteriores, las medidas de capacidad general, de dependencia-independencia de campo y de personalidad no llegaron a explicar el 9% (en algunos casos no superaba el 2%) de la varianza, en este experimento, el 22.8% de la varianza del tiempo de diagnóstico se debió a la variación en el tiempo necesario para realizar una prueba de manejo del ratón y al número de errores cometido en la misma.

En este sentido, los sujetos con mayor habilidad de manejo del ratón (tardaron menos tiempo y cometieron menos errores en la prueba) necesitaron menos tiempo para realizar la tarea de diagnóstico, observaron más fuentes de información y emitieron un mayor número de diagnósticos correctos.

**CAPITULO 11. EFECTOS DE LA PRESENTACION GRAFICA DE
LA INFORMACION Y DE LA INTRODUCCION DE UNA TAREA
SECUNDARIA**

11.1. INTRODUCCION

Una gran línea de investigación dentro de la Ergonomía Cognitiva, se centra en las diferencias que sobre la ejecución y la carga tienen lugar cuando se utilizan displays gráficos y digitales o alfanuméricos.

En este sentido, la mayoría de los investigadores han encontrado mejoras en el tiempo necesitado para realizar la tarea cuando la información se presenta gráficamente (Boles y Wickens, 1987; Coury y Pietras, 1989; Schwartz y Howell, 1985). Además, autores como Eberts (1987) concluyen que la presentación gráfica de la información reduce los niveles de carga.

Sin embargo, autores como Remington y Williams (1986), encontraron que debido a la alta similaridad intragrupo de los símbolos gráficos, los displays numéricos producían mejoras en la ejecución.

En todas estas investigaciones se pone de manifiesto la importancia que sobre los resultados obtenidos tiene el tipo de tarea que se esté realizando, lo cual impide que se puedan extraer conclusiones generales sobre cual de los dos tipos de display es superior. En general, parece que la mejor representación del sistema viene dada por un display de tipo múltiple, en el que la misma información es presentada tanto gráfica como digitalmente (Coury y Pietras, 1989), ya que esta presentación redundante de la información suministra evidencia confirmatoria a los sujetos, que se traduce en una mejora de la calidad de la ejecución.

Otra de las conclusiones obtenidas en este sentido, hace referencia a la superioridad del display gráfico en aquellas tareas que requieren una visión global o

de conjunto del sistema. Sin embargo, cuando la tarea requiere la observación de cada una de las fuentes de información del display por separado para poder identificar la causa del fallo, los displays digitales producen niveles de ejecución más elevados (Casey, 1987; Coury, Boulette y Smith, 1989).

Los displays gráficos más utilizados son los diagramas de barras (Boles y Wickens, 1987; Barnett y Wickens, 1988; Buttigieg y col., 1988; Coury, Boulette y Smith, 1989; Coury y Purcell, 1988). Utilizando tareas de diagnóstico de fallos del sistema, el grupo de Coury (Coury, Boulette y Smith, 1989; Coury y Purcell, 1988) y Buttigieg y colaboradores (1988), entre otros, han encontrado que el display de gráfico de barras puede ser un formato muy versátil, y es la mejor elección para mostrar datos correlacionados y multidimensionales cuando el rango de incertidumbre sobre el estado del sistema es alto.

Por estos motivos, a diferencia de los anteriores, en este experimento se utilizaron gráficos de barras horizontales para representar el estado de cada componente del sistema.

Por otro lado, otro de los objetivos de este experimento hace referencia a los efectos de la inclusión de una tarea secundaria. Siguiendo el modelo de múltiples recursos (Wickens, 1984) el hecho de que dos tareas puedan ser realizadas a la vez sin que se produzcan decrementos en la ejecución implica que ambas demandan recursos de naturaleza diferente. Por contra, cuando dos tareas demandan recursos comunes, la ejecución conjunta de ambas producirá menores niveles de ejecución en la tarea primaria (Navon y Gopher, 1979; Norman y Bobrow, 1975; Gopher y Sanders, 1984; Sanders, 1983; Wickens, 1980, 1983, entre otros).

Cuando dos tareas demandan recursos de naturaleza diferente, los niveles de carga, medida por el procedimiento de tarea secundaria, no se ven afectados. Sin embargo, se han encontrado que frecuentemente las medidas de la ejecución y las medidas subjetivas de carga disocian (Derrick, 1981, 1988; Eggemeier, Crabtree y Reid, 1982; Herron, 1980; Vidulich y Wickens, 1986; Wickens y Yeh, 1983; Yeh y Wickens, 1988).

De este modo, con el fin de comprobar si la estructura de recursos demandados por la tarea primaria era aquella que en principio se asumía, en este experimento se introdujo una tarea secundaria que demandaba recursos de naturaleza distinta a los de la primaria. Si la inclusión de la tarea secundaria reduce los niveles de ejecución, ambas demandarían los mismos recursos. En caso contrario ambas tareas podrían ser realizadas a la vez sin perjuicio de la ejecución de la tarea primaria porque sus demandas serían, como se espera, diferentes.

Por otro lado, debido a las disociaciones encontradas entre las medidas de ejecución y las de carga subjetiva, las estimaciones subjetivas de la carga mental de la tarea primaria podrían verse afectadas por la inclusión de la secundaria independientemente de que esta afecte o no al nivel de ejecución.

Los objetivos principales de este cuarto y último experimento fueron los siguientes:

1. Analizar las posibles diferencias en la ejecución y la carga subjetiva debidas al tipo de información presentada (alfanumérica y gráfica).
2. Establecer, mediante la inclusión de una tarea secundaria, el tipo de recursos demandados por la tarea primaria.
3. Analizar en qué medida la valoración subjetiva de la carga se ve afectada por la inclusión de una tarea secundaria, que en principio no comparte recursos con la primaria.



En relación a estos objetivos se establecen las siguientes hipótesis:

1. Se espera encontrar diferencias en ejecución y carga debidas al tipo de información presentada. Siguiendo a Coury y Pietras (1989) se espera que la información gráfica reduzca el tiempo necesario para realizar la tarea, y que por tanto se produzcan menores niveles de carga subjetiva (Eberts, 1987).
2. Debido a que ambas tareas demandan en principio recursos diferentes, se espera que la inclusión de la tarea secundaria no tenga efectos significativos sobre la ejecución, ni sobre la carga subjetiva, de la tarea primaria (Wickens, 1984).

11.2. METODO

11.2.1. Sujetos

En este experimento participaron 14 sujetos estudiantes de Psicología, con edades comprendidas entre los 21 y los 29 años, de los cuales 2 fueron varones y 12 mujeres. Todos ellos recibieron un aumento de 0,25 puntos en la nota final de una de sus asignaturas del curso.

11.2.2. Diseño

El diseño fue un factorial $2 \times 2 \times 2$ de medidas repetidas. Los tres factores utilizados fueron los mismos que en el experimento anterior. La introducción de la tarea secundaria fue contrabalanceada entre los sujetos, de forma que la mitad de ellos realizaron la primera mitad de la tarea primaria en condiciones de tarea dual, mientras que para los siete sujetos restantes la tarea secundaria fue introducida durante la segunda mitad del experimento.

Las variables dependientes utilizadas en los análisis, fueron las mismas que en los experimentos anteriores.

11.2.2. Tarea y Procedimiento

Básicamente la tarea y el procedimiento fueron los mismos que en el experimento anterior. Sin embargo, en este experimento la información fue representada mediante gráficos. Los valores de los experimentos anteriores fueron sustituidos por filas de X's, de manera que se asemejaban a gráficos de barras horizontales.

Como en otros experimentos, los sujetos dispusieron de 30 segundos para realizar cada ensayo.

La tarea secundaria consistió en atender a una grabación en la que el sujeto debía distinguir un sonido de silbato, del ruido generado por el funcionamiento de diversas máquinas y otros sonidos. Cuando el sujeto escuchaba el silbato debía responder con la palabra "ahora". Esta tarea demanda recursos de procesamiento verbal, el input es auditivo y la modalidad de respuesta es oral.

En las instrucciones se indicaba al sujeto que durante la fase de tarea dual, debía atender a ambas tareas a la vez, manteniendo los niveles de ejecución lo mejor posible en ambas tareas.

11.2.3. Instrumentos

Para la presentación de las instrucciones y demás aspectos del experimento se utilizó un ordenador PC/XT con disco duro y un ratón GM-6 PC Mouse. Al igual que en las demás ocasiones, el experimento fue generado y controlado mediante el programa Mouselab System versión 4.2 (Johnson, Payne, Schkade y Bettman, 1989), el cual también se encargó de la recogida de los datos.

Para la tarea secundaria se utilizó un Micro Cassette Aiwa TP-M10.

11.3. RESULTADOS

11.3.1. Efectos de la inclusión de la tarea secundaria

En primer lugar, se realizaron análisis de varianza con el fin de determinar el efecto de la inclusión de la tarea secundaria, sobre el tiempo total de tarea. Como la mitad de los sujetos realizaron la primera mitad de la tarea primaria en condiciones de tarea dual, mientras que para la otra mitad de sujetos la tarea secundaria fue introducida durante la segunda mitad del experimento, se realizaron dos análisis de varianza, uno para la primera mitad de ensayos y otro para la segunda.

En el primer análisis se comparaba el tiempo tardado en realizar cada uno de los diez primeros de ensayos, entre el grupo de sujetos que había realizado estos ensayos en condiciones de tarea dual y el grupo que los había realizado en condiciones de tarea simple.

En el segundo análisis la comparación era la misma, pero tomando como variable dependiente el tiempo tardado en realizar los diez últimos ensayos. En ninguno de los dos análisis el efecto de la inclusión de la tarea secundaria resultó significativo [$F(1,12) = 0.03$, $p > 0.5$, para el primer análisis; y $F(1,12) = 2.43$, $p > 0.1$, para el segundo].

La figura 11.1. muestra que prácticamente en ningún caso se produjo un aumento en el tiempo necesitado para emitir el diagnóstico, en condiciones de tarea dual.

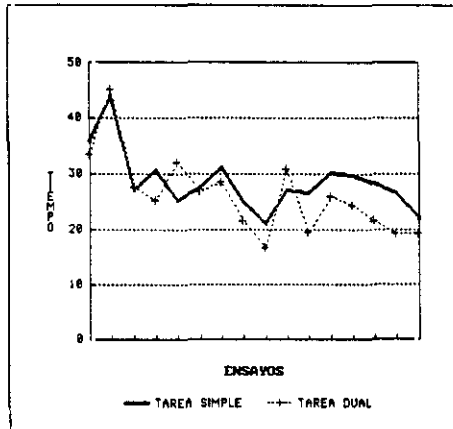


Figura 11.1. Media del tiempo de diagnóstico (sg.) para cada ensayo, en condiciones de tarea simple y de tarea dual.

En segundo lugar, este análisis fue repetido para el índice global de carga subjetiva y para las estimaciones de cada una de las dimensiones de carga. El efecto de la inclusión de la tarea secundaria no fue significativo en ninguno de los casos ($p > 0.1$).

Las figuras 11.2, 11.3 y 11.4 muestran como las estimaciones de la demanda mental y la temporal, así como el índice global de carga, son en casi todos los ensayos muy similares bajo ambas condiciones (simple y dual).

Sin embargo, a partir de la figura 11.5. se observa como las diferencias en la estimación de la demanda física son mucho más acusadas, tanto en cuanto a la condición como en función del ensayo. Mientras que en condiciones de tarea simple los sentimientos de carga física van disminuyendo a medida que transcurre el experimento, bajo condiciones de tarea dual la estimación de la demanda física va aumentando. En este punto hay que recordar que esta dimensión estaba definida por aspectos tanto manuales o esencialmente físicos (que no fueron modificados para cada ensayo), como por los sentimientos de frustración de los sujetos (ver capítulo 7). Por tanto, este resultado sugiere que bajo condiciones de tarea dual los sujetos se fueron sintiendo cada vez más molestos con la tarea, mientras que en condiciones de tarea simple la sensación de agrado hacia la tarea fue en aumento.

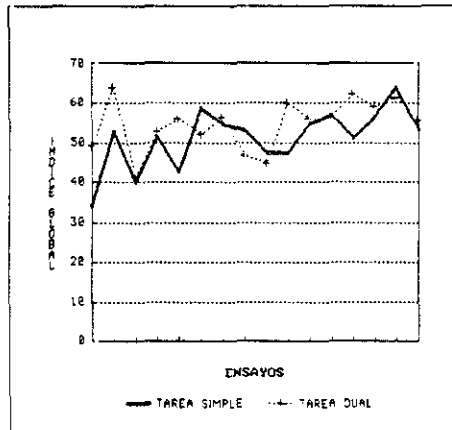


Figura 11.2. Media del índice global de carga para ensayo, en condiciones de tarea simple y dual.

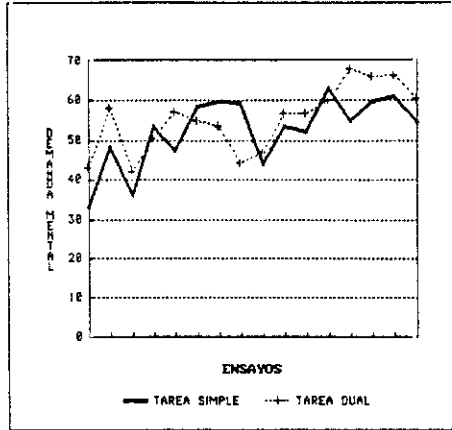


Figura 11.3. Media de la estimación de la demanda mental de cada ensayo, en condiciones de tarea simple y dual.

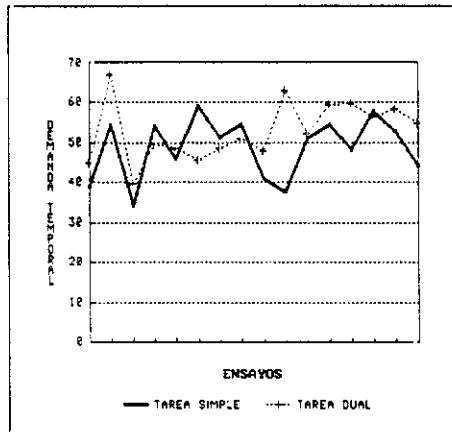


Figura 11.4. Media de la estimación de la demanda temporal de cada ensayo, en condiciones de tarea simple y dual.

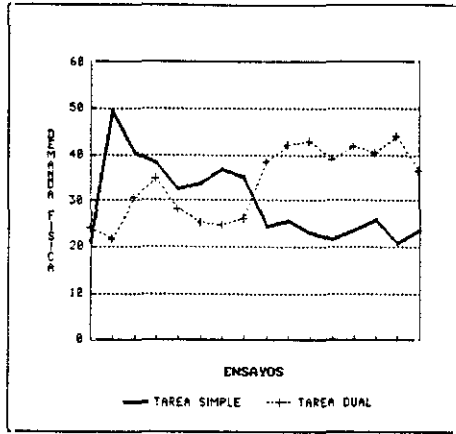


Figura 11.5. Media de la estimación de la demanda física de cada ensayo, en condiciones de tarea simple y dual.

11.3.2. Efectos del tipo de información

Para comprobar los efectos del tipo de información, se realizó un MANOVA para cada una de las medidas de ejecución (tiempo de tarea, número de aciertos y número de datos observados), utilizando los datos del segundo, del tercero y de este experimento (para que el tamaño de las muestras de cada celdilla fuera el mismo, se eliminó un sujeto aleatoriamente de aquellos que participaron en el experimento 2, y otro de los que participaron en el 3).

En este análisis el efecto simple del tipo de información no resultó significativo para ninguna de las variables dependientes (ver Tabla 1, apéndice 5). Solamente resultó significativo el efecto de la interacción entre el tipo de información, el número de causas posibles del problema y la probabilidad media de ocurrencia de la solución, sobre el número de aciertos [$F(2,39) = 4.25, p < 0.05$].

La figura 11.6. muestra el tiempo necesitado para emitir el diagnóstico en cada combinación, en función del tipo de representación de la información. Aunque no se encontraron diferencias significativas, la figura muestra que con la presentación gráfica de la información, el tiempo de diagnóstico fue similar al obtenido con información numérica. Los menores tiempo obtenidos con el display verbal parecen indicar que este sería el más aconsejable bajo condiciones de presión temporal.

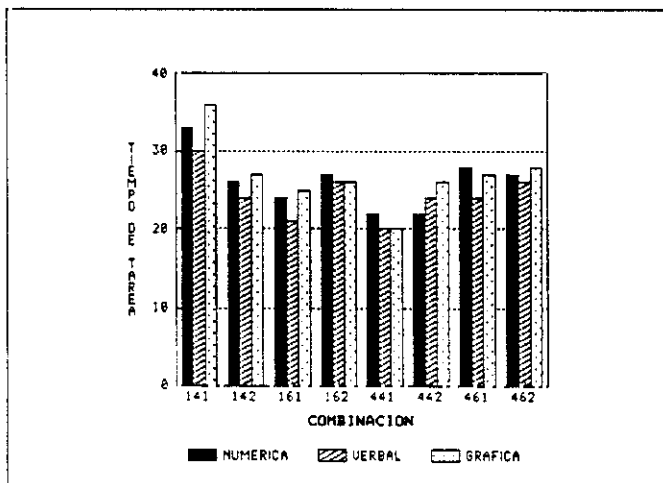


Figura 11.6. Medias del tiempo de diagnóstico (sg.) en cada combinación y para cada tipo de información.

En la figura 11.7. aparece el número medio de aciertos en cada combinación, en función del tipo de información presentada. Solamente en un caso (combinación 462) se emitieron más diagnósticos correctos con el display gráfico. Todo parece indicar que el error humano de diagnóstico no se debe al formato de presentación de la información, sino solamente a las variaciones en la complejidad de solución del problema.

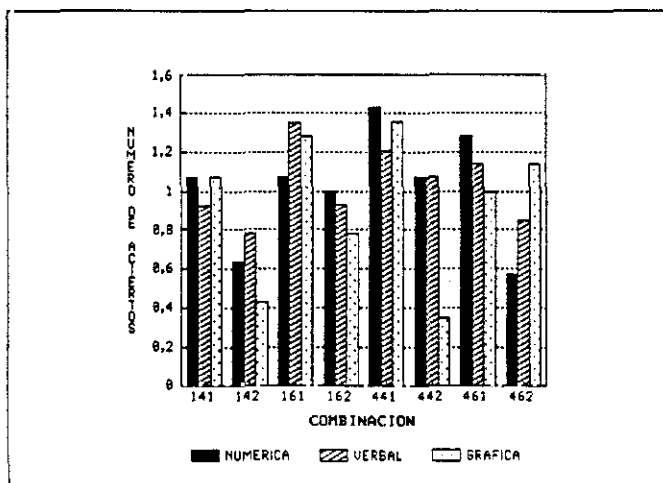


Figura 11.7. Media del número de aciertos para cada combinación y cada tipo de información.

La figura 11.8. muestra el número medio de fuentes de información observadas por los sujetos en cada combinación y en función del formato de presentación de la información. La figura parece indicar que con el display gráfico los sujetos necesitaron recoger un mayor número de datos que con los demás displays, aunque esta diferencia no fuera estadísticamente significativa.

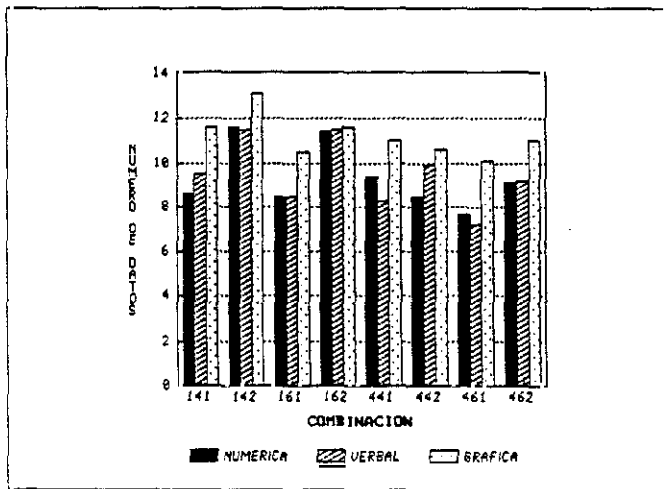


Figura 11.8. Media del número de datos observado para cada combinación y para cada tipo de información.

Por otro lado, se repitió el mismo análisis para las medidas de carga subjetiva (índice global de carga y estimaciones de la demanda mental de la tarea, de la temporal y de la física) (ver Tabla 2, apéndice 5). En este análisis, el efecto simple del tipo de información solo resultó significativo para el índice global de carga [$F(2,39)= 3.24, p=0.05$].

Sobre la estimación de la demanda mental de la tarea resultó significativo el efecto de la interacción entre el tipo de información y el número de causas posibles del problema [$F(2,39)= 6.00, p<0.01$].

No se encontraron efectos significativos sobre las estimaciones de la demanda temporal ni de la física.

La figura 11.9 muestra como la carga global percibida por los sujetos fue mayor con el display gráfico que con el verbal, pero menor que con el numérico. Algo muy similar sucedió con la estimación media de la demanda temporal (Figura 11.11).

La estimación media de la demanda mental fue muy similar con el display verbal que con el gráfico, pero menor que con el display numérico (Figura 11.10). Algo muy similar ocurrió con la estimación media de la demanda física (Figura 11.12).

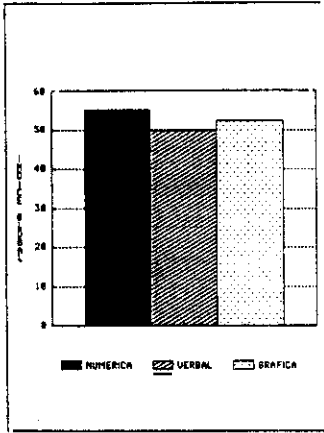


Figura 11.9. Media del índice global de carga en función del tipo de información.

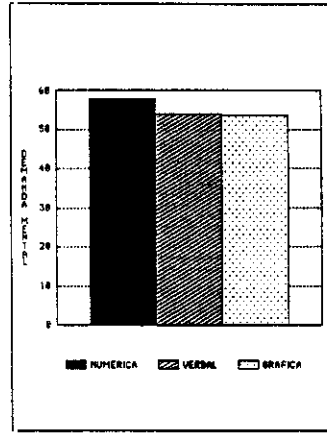


Figura 11.10. Media de las estimaciones de la demanda mental en función del tipo de información.

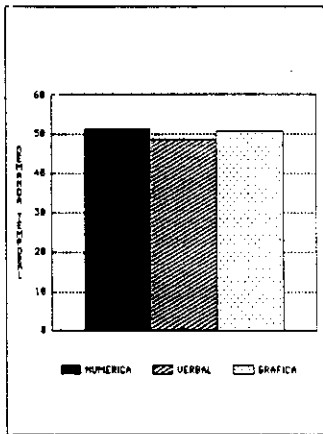


Figura 11.11. Media de la estimación de la demanda temporal en función del tipo de información.

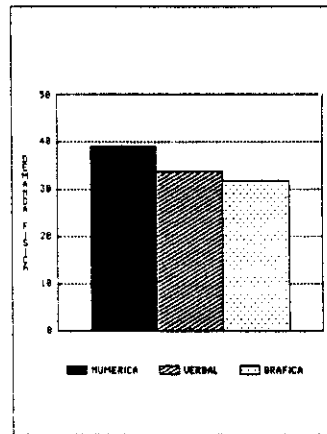


Figura 11.12. Media de la estimación de la demanda física en función del tipo de información.

11.3.1. Análisis de la ejecución

Con el fin de confirmar los resultados obtenidos en los experimentos anteriores, al igual que en ellos se realizó un MANOVA con los factores experimentales, para cada una de las medidas de la ejecución, tiempo de tarea (TT), número de aciertos (NA) y número de datos observados por los sujetos para emitir su diagnóstico (ND), utilizando los datos de los sujetos que participaron en este experimento (N=14). Los resultados de estos análisis se muestran a continuación (ver Tabla 3, apéndice 5).

11.3.1.1. Tiempo total de tarea

En cuanto al tiempo de tarea resultó significativo el efecto simple de la complejidad perceptiva [$F(1,13)= 5.000, p<0.05$], y los efectos de la interacción entre la complejidad perceptiva y el número de causas posibles del problema [$F(1,13)= 21.38, p<0.001$], de la interacción entre complejidad perceptiva y la probabilidad media de ocurrencia de la solución [$F(1,13)= 9.27, p<0.01$], y de la triple interacción entre los factores [$F(1,13)= 8.30, p<0.05$].

Se realizaron comparaciones de medias entre los tiempos de tarea de las combinaciones CPxNCxPM, por el método de Scheffé ($S=5.53, p<0.05$). La tabla 11.1. muestra los valores del estadístico T que resultaron significativos para cada comparación. Los resultados son muy similares a los de experimentos anteriores.

	141	142	161	162	241	242	261	262
141								
142	3.5							
161	4.4	NS						
162	3.9	NS	NS					
241	6.1	2.6	NS	2.2				
242	4.1	NS	NS	NS	NS			
261	3.5	NS	NS	NS	2.6	NS		
262	3.3	NS	NS	NS	2.8	NS	NS	

Tabla 11.1. Valor del estadístico T para cada contraste significativo entre los tiempos de tarea de cada combinación CPxNCxPM. NS=No Significativo.

La tabla 11.2. muestra el valor del estadístico T de Scheffé para cada contraste realizado entre los tiempos medios de tarea de las combinaciones CPxNC ($S=4.8929$, $p<0.05$).

	14	16	24	26
14				
16	2.3			
24	3.2	NS		
26	NS	NS	NS	

Tabla 11.2. Valor de T para cada contraste significativo entre los tiempos de tarea de cada combinación CPxNC. NS=No Significativo.

La tabla 11.3. muestra el valor del estadístico T de Scheffé para cada contraste realizado entre los tiempos medios de tarea de las combinaciones CPxPM ($S=4.8724$, $p<0.05$).

	11	12	21	22
11				
12	NS			
21	3.0	NS		
22	NS	NS	NS	

Tabla 11.3. Valor de T para cada contraste significativo entre los tiempos de tarea de cada combinación CPxPM. NS=No Significativo.

11.3.1.2. Aciertos

En cuanto al número de aciertos, resultó significativo el efecto simple de la probabilidad media de ocurrencia de la solución [$F(1,13)= 12.13$, $p<0.005$]. También resultaron significativas las interacciones entre número de causas y probabilidad media de la solución [$F(1,13)= 9.16$, $p<0.05$], así como la triple interacción [$F(1,13)= 4.79$, $p<0.05$].

Se realizaron comparaciones de medias por el método Scheffé del número de aciertos de cada combinación CPxNCxPM ($S=0.4966$, $p<0.05$). La tabla 11.4. muestra los valores de T para los contrastes que resultaron significativos.

	141	142	161	162	241	242	261	262
141								
142	2.8							
161	NS	3.7						
162	NS	NS	2.2					
241	NS	4.0	NS	2.4				
242	3.1	NS	4.0	NS	4.3			
261	NS	2.5	NS	NS	NS	2.8		
262	NS	3.1	NS	NS	NS	3.4	NS	

Tabla 11.4. Valor del estadístico T para cada contraste significativo entre el número de aciertos de cada combinación CPxNCxPM. NS= No Significativo.

La tabla 11.5 muestra el valor del estadístico T de Scheffé para cada contraste realizado entre el número medio de aciertos de las combinaciones NCxPM ($S=0.4618$, $p<0.05$).

	41	42	61	62
41				
42	3.8			
61	NS	3.5		
62	NS	2.7	NS	

Tabla 11.5. Valor de T para cada contraste significativo entre el número de aciertos de cada combinación NCxPM. NS=No Significativo.



11.3.1.3. Número de datos observados

El número de datos o información recogida por los sujetos antes de emitir su diagnóstico solo se vio afectado por el efecto de la complejidad perceptiva [$F(1,13)=4.97, p<0.05$].

En general, el patrón de resultados en cuanto a las medidas de la ejecución es el mismo que el encontrado en otros experimentos. Por lo que se confirma que las variaciones en ejecución se deben principalmente a las modificaciones en la complejidad, perceptiva y de solución del problema, de la tarea.

En la tabla 11.6 aparecen las medias y desviaciones típicas de las tres medidas de la ejecución en cada uno de los niveles de los factores de complejidad.

	TT		NA		ND	
	Media	ST.	Media	ST.	Media	ST.
TOTAL	27.44	9.69	.92	.73	11.23	4.73
CP 3x4	29.12	11.13	.89	.70	11.75	4.90
NC 4	32.16	12.27	.75	.70	12.39	5.33
PM 1	36.70	13.62	1.07	.73	11.64	5.10
PM 2	27.62	9.12	1.42	.51	13.14	5.63
NC 6	36.07	9.08	1.33	.69	11.10	4.43
PM 1	25.52	9.55	1.28	.72	10.57	4.34
PM 2	26.62	8.91	.78	.57	11.64	4.61
CP 3x3	25.76	7.73	.96	.76	10.72	4.53
NC 4	23.57	6.46	.85	.64	10.83	4.17
PM 1	20.98	6.11	1.35	.74	11.00	4.95
PM 2	26.15	5.90	.35	.63	10.67	3.39
NC 6	27.95	8.37	1.07	.66	10.60	4.95
PM 1	27.66	9.58	1.00	.67	10.17	5.59
PM 2	28.25	7.32	1.14	.66	10.03	4.38

Tabla 11.6. Medias y desviaciones típicas de las medidas de ejecución en cada una de las combinaciones intrasujeto.

11.3.2. Análisis de la carga subjetiva

Siguiendo el mismo procedimiento que en los experimentos anteriores, se realizó un MANOVA para la medida del índice global de carga subjetiva, y otro para cada una de las dimensiones de carga (mental, temporal y física), a partir de los datos de los 14 sujetos que participaron en este experimento. Los resultados de estos análisis se muestran a continuación (ver Tabla 4, apéndice 5).

11.3.2.1. Índice global de carga subjetiva

En el análisis realizado para establecer la influencia de los factores sobre el índice global de carga, resultaron significativos los efectos simples de la complejidad perceptiva [$F(1,13)= 6.53, p<0.05$] y del número de causas posibles del problema [$F(1,13)= 8.89, p<0.05$].

11.3.2.2. Dimensiones de la carga subjetiva

La valoración de la demanda mental de la tarea se vio afectada por la complejidad perceptiva del sistema [$F(1,13)= 4.95, p<0.05$] y por el número de causas posibles del problema [$F(1,13)= 11.20, p<0.01$].

De igual forma, la complejidad perceptiva y el número de causas posibles del problema tuvieron efectos significativos sobre la valoración de la demanda temporal [$F(1,13)= 5.04, p<0.05$, para la complejidad perceptiva; $F(1,13)= 6.62, p<0.05$, para el número de causas].

En cuanto a la valoración de la demanda física de la tarea, no resultó significativo ninguno de los efectos simples ni de la interacción entre los factores.

La tabla 11.7. muestra las medias y desviaciones típicas de la medida global de carga, así como de las estimaciones de cada una de las dimensiones de carga subjetiva, encontradas en este experimento.

	INDICE		DM		DT		DF	
	Media	ST.	Media	ST.	Media	ST.	Media	ST.
TOTAL	52.33	14.04	53.82	16.14	50.82	18.80	31.85	25.58
CP 3x4	54.04	11.94	55.37	15.94	53.06	20.00	32.53	26.01
NC 4	55.30	16.80	57.05	16.47	54.67	21.34	32.32	26.80
PM 1	53.39	16.50	52.73	15.04	53.96	20.82	32.99	25.17
PM 2	57.91	16.80	61.38	17.24	55.38	22.62	31.66	29.27
NC 6	52.79	14.47	53.70	15.51	51.46	18.81	32.74	25.69
PM 1	51.88	16.42	53.19	16.02	51.72	22.87	34.70	27.76
PM 2	53.70	12.16	54.20	15.57	51.19	14.55	30.78	24.33
CP 3x3	50.62	15.01	52.27	16.34	48.57	17.40	31.17	25.35
NC 4	54.50	14.72	57.13	17.37	52.36	17.38	31.73	27.09
PM 1	52.37	16.80	54.63	19.19	50.06	18.88	32.00	29.33
PM 2	56.64	11.93	59.63	15.66	54.67	16.10	31.47	25.77
NC 6	46.74	14.28	47.41	13.89	44.78	16.89	30.60	23.97
PM 1	46.27	15.38	47.77	15.70	45.84	18.56	26.41	21.05
PM 2	47.22	13.08	47.05	12.42	43.72	15.67	34.79	26.69

Tabla 11.7. Medias y desviaciones típicas de las medidas de carga subjetiva en cada una de las combinaciones intrasujeto.

11.4. DISCUSION

No se encontraron efectos significativos de la representación gráfica de la información frente a la verbal o a la numérica, sobre las medidas de la ejecución. Solamente resultó significativo el efecto de la interacción entre el tipo de información y los dos factores de complejidad de solución del problema, sobre el número de diagnósticos correctos emitidos. Sin embargo, se encontraron diferencias en el índice global de carga subjetiva debidas al formato de presentación de la información (numérica, verbal, gráfica).

Por otro lado, ni la ejecución ni la carga subjetiva resultaron significativamente afectadas por la inclusión de una tarea secundaria que, en principio, demandaba recursos de naturaleza diferente a los de la primaria.

En cuanto a la ejecución, utilizando el display gráfico se encontró un patrón de efectos de los factores de complejidad (perceptiva y de solución del problema) de la tarea muy similar al obtenido en los experimentos anteriores. Esto implica, que básicamente las estrategias de búsqueda de información utilizadas por los sujetos no se vieron modificadas por la inclusión de la tarea secundaria ni por el cambio en la presentación de la información. Sin embargo, la carga percibida por los sujetos fue significativamente mayor con el display gráfico que con el verbal, alcanzando niveles similares a los obtenidos con el display numérico.

En contra de los resultados obtenidos por Boles y Wickens (1987), Coury, Boulerte y Smith (1989), Coury y Purcell (1988) o Schwartz y Howell (1985), entre otros, la presentación gráfica de la información no redujo el tiempo requerido para realizar la tarea de diagnóstico. Además, tampoco parece confirmarse el resultado encontrado por Remington y Williams (1986) relativo a la superioridad del display numérico, ya que no se encontraron grandes diferencias en el tiempo de diagnóstico entre ambos tipos de display.

Aunque no resultó estadísticamente significativa, se encontró una mejora sistemática en el tiempo requerido para realizar la tarea cuando se utilizaba el display verbal (ver figura 11.6). Esta reducción en el tiempo requerido para el diagnóstico no produjo un aumento en el número de errores, lo que implica que el tiempo requerido para procesar la información verbal fue menor que el necesitado para alcanzar un mismo nivel de comprensión de la información numérica o la gráfica.

Además, en contra de los resultados obtenidos por Eberts (1987), los niveles globales de carga subjetiva fueron algo menores con el display verbal.

Estos resultados, al igual que los encontrados en las investigaciones mencionadas, son específicos del tipo de tarea y de las características individuales de los sujetos que las han de realizar. Mientras que en la mayor parte de las investigaciones revisadas participan estudiantes de Ingeniería, cuya formación es muy técnica, y por tanto cargada de aspectos gráficos y numéricos, en este estudio los sujetos fueron estudiantes de Psicología, cuya formación está fuertemente cargada de contenidos y aspectos más de tipo verbal. Esta podría ser la causa de que los resultados obtenidos no sean los que, en función de la literatura revisada, se esperaban.

Por otro lado, los resultados confirmaron la hipótesis previa relativa a la estructura de recursos demandados por la tarea primaria. La tarea secundaria utilizada demandaba recursos de procesamiento verbal, el input era auditivo y la modalidad de respuesta oral. Como la ejecución de la tarea primaria no resultó afectada por la introducción de la secundaria, se confirma que la tarea primaria demanda recursos de procesamiento de tipo espacial (información gráfica), la modalidad del input es visual y el tipo de respuesta es manual (Wickens, 1984; Wickens y Liu, 1988).

Al igual que la ejecución, las estimaciones subjetivas de la carga no se vieron afectas por la introducción de la tarea secundaria. Únicamente se encontró una pequeña variación en las estimaciones de la dimensión de carga denominada "física". Debido a que la dificultad manual de la tarea se modificó, y a que en la definición de esta dimensión también intervienen aspectos de tipo emocional (sentimientos de frustración, molestia, inseguridad, estrés, etc.), esto parece indicar que bajo condiciones de tarea dual se produjo un incremento en los sentimientos de desagrado hacia la tarea a medida que transcurría el experimento, mientras que el efecto fue al contrario bajo condiciones de tarea simple (ver figura 11.5).

**CAPITULO 12. EVALUACION DE LA TECNICA SUBJETIVA
UTILIZADA**

12.1. INTRODUCCION

Como ya se ha mencionado en anteriores capítulos, actualmente existe un gran interés en la búsqueda de una técnica de medida de la carga subjetiva que no disocie con las medidas de la ejecución. La validez y utilidad de estas técnicas se evalúa principalmente en función de su capacidad para detectar diferentes niveles de dificultad de la tarea (Vidulich y Tsang, 1986).

En múltiples ocasiones se han encontrado disociaciones entre las estimaciones subjetivas de la carga y las medidas de la ejecución en la tarea. Este problema ha sido explicado bajo dos puntos de vista: uno teórico y otro metodológico. La perspectiva teórica implica una serie de supuestos sobre la naturaleza y contenido de la experiencia consciente, y sus relaciones con la atención, el procesamiento de la información y la ejecución. Es decir, ¿La experiencia consciente, y por tanto las medidas subjetivas, incorpora todos los fenómenos de interés incluidos en el concepto de carga de trabajo?. Dar respuesta a esta cuestión trasciende los objetivos de la investigación presente.

Desde el punto de vista metodológico, el interés se centra en el desarrollo de una escala que imponga mínimas restricciones a priori, capaz de cuantificar la experiencia de los sujetos, y que permita la comparación entre e intra tareas (Gopher y Braune, 1984).

Siguiendo este punto de vista, la finalidad de este capítulo es evaluar las características de la técnica de medida de la carga subjetiva utilizada en esta investigación, a partir de su sensibilidad para detectar diferentes niveles de dificultad de la tarea, y su concordancia con las medidas de la ejecución (la descripción de esta técnica aparece en el capítulo 7).

12.2. ANALISIS

En este apartado se distinguirán dos tipos de análisis. El primer tipo de análisis de las características de la técnica será función de los resultados encontrados en los experimentos realizados a partir del análisis de la varianza de las estimaciones de carga. Este será completado con un análisis de regresión logística realizado con los datos del tercer experimento, en el que se pondrá en relación la estimación de la carga subjetiva con la probabilidad de acertar cada problema o ensayo.

12.2.1. Análisis de varianza

En primer lugar, en el primer y segundo experimentos se manipularon dos factores entresujetos (tipo de matriz y tipo de feedback). Ambos factores no tuvieron efectos simples ni de interacción significativos sobre la ejecución y tampoco sobre las estimaciones de la carga.

En segundo lugar, en todos los experimentos se utilizaron una serie de factores intrasujeto, que hacían referencia a la dificultad o complejidad de la tarea. Estos factores eran la complejidad perceptiva y la complejidad de solución del problema. El primer factor incluye aspectos de tipo perceptivo/sensorial, mientras que el segundo estaría relacionado con las demandas de procesamiento central (recordar, decidir, razonar, etc.). Siguiendo la definición dada a las distintas dimensiones de carga que se utilizaron, ambos tipos de complejidad forman parte de la demanda mental de la tarea. La técnica utilizada fue diferencialmente sensible a estas variaciones de la dificultad de la tarea. En las situaciones de menor complejidad (display verbal o situaciones de no presión temporal), el índice global de carga se vio afectado por la complejidad de solución del problema, mientras que en aquellas en las que la ejecución era peor (display numérico o gráfico, con presión temporal) la carga global percibida por los sujetos era función principalmente de la complejidad perceptiva de la tarea.

Por otro lado, en el primer experimento no había tiempo límite para realizar la tarea, mientras que en los demás se introdujo un factor de presión temporal. Esta modificación en el tiempo límite para realizar la tarea tuvo efectos sobre la ejecución, sin embargo las medidas de carga no fueron capaces de detectar significativamente estas diferencias.

En cuanto al tipo de representación de la información (numérica, verbal y gráfica), no se encontraron efectos simples de este factor sobre la ejecución, pero sí en interacción con la complejidad de solución del problema. Esta interacción también resultó significativa para la estimación de la demanda mental. Además el índice global de carga resultó afectado por el efecto simple del tipo de representación de la información.

Por último, la inclusión de una tarea secundaria que demandaba recursos de naturaleza diferente a los demandados por la primaria no tuvo efectos significativos ni sobre la ejecución ni sobre la carga experimentada por los sujetos.

12.2.2. Regresión logística

El procedimiento de regresión logística se fundamenta en los modelos logarítmico-lineales. Estos modelos se basan en una transformación logarítmica de la proporción (o probabilidad) de cada categoría de la variable criterio, que tiene como objetivo representar la dependencia entre esa variable criterio y las variables predictoras de forma que la probabilidad de cada categoría se mantenga entre 0 y 1. Esta transformación es la siguiente:

$$\lambda = \ln\left[\frac{P}{1-P}\right]$$

Lambda recibe el nombre de logística. Como la probabilidad varía entre 0 y 1, la correspondiente logística lo hace entre $-\infty$ y $+\infty$.

De esta forma, el modelo logístico-lineal puede ajustarse a un modelo de regresión lineal para variables criterio dicotómicas, cuyos valores posibles son 0 y 1. El modelo de regresión logística vendría dado por la siguiente fórmula:

$$\ln [P(y=1)/P(y=0)] = \beta_0 + \sum_{i=1}^p \beta_i x_i$$

; donde, $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ son los coeficientes de regresión a estimar. La estimación de estos parámetros se realiza mediante procedimientos de máxima verosimilitud.

La interpretación de los resultados de la regresión logística sigue la misma línea de razonamiento que en los análisis de regresión múltiple convencionales.

Con los datos de los 15 sujetos que participaron en el tercer experimento se realizó un análisis de regresión logística para cada ensayo. En estos análisis se pretendía establecer si las estimaciones dadas por los sujetos para las tres dimensiones de carga estaban relacionadas con la probabilidad de emitir un diagnóstico correcto del problema, y por tanto, con la localización correcta del fallo.

Para los ensayos 4, 8, y 13 no se pudo realizar este análisis ya que nunca fueron acertados por los sujetos. En siete de los restantes ensayos ninguna de las variables predictoras entraron en la ecuación de regresión. La estimación de la demanda física solo fue predictora del acierto en el ensayo 19. Para los restantes nueve ensayos fue la estimación de la demanda mental la única variable que entró en la ecuación de regresión. El porcentaje de casos correctos variaba en los ensayos desde el 66.67% al 93.33%.

Debido a la variabilidad entre ensayos encontrada y puesto que en todos ellos solo se modificó la complejidad mental (perceptiva/central), quedando constantes la presión temporal y la demanda física, se realizó el mismo análisis de regresión logística para todos los ensayos en conjunto. En este análisis sólo entró en la ecuación

de regresión la estimación de la dimensión "demanda mental" [$p < 0.0001$], siendo la estimación de las otras dos dimensiones no significativa (ver tabla 12.1.). El porcentaje de casos correctos clasificados a partir de la demanda mental fue del 70.67% (ver tabla 12.2).

	Chi-Square	df	Significance				
-2 Log Likelihood	340.455	298	.0455				
Model Chi-Square	71.103	1	.0000				
Improvement	71.103	1	.0000				
Goodness of Fit	299.146	298	.4704				
--- Variables en la ecuación ---							
Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
DM	-.0655	.0091	51.2725	1	.0000	-.3460	.9366
--- Variables no en la ecuación ---							
Chi2 residual	4.683 with 2 df		Sig = .0962				
Variable	Score	df	Sig	R			
DT	.8148	1	.3667	.0000			
DF	2.1254	1	.1449	.0175			

Tabla 12.1. Resultado del análisis de regresión logística realizado sin distinguir entre ensayos a partir de las tres dimensiones de carga. DM= Demanda mental; DT=Demanda temporal; DF= Demanda física.

		Pronosticados		Porcentaje de correctos
		0	1	
Observados	0	137	31	81.55%
	1	57	75	56.82%
		Total		70.67%

Tabla 12.2. Porcentaje de casos correctos pronosticados a partir de la estimación de la demanda mental. 1=Diagnóstico correcto; 0=Diagnóstico incorrecto.

La tabla 12.2. muestra que el porcentaje de casos correctamente clasificados por la ecuación de regresión es mucho mayor para los verdaderos negativos (aquellos sujetos a los que se les pronostica un diagnóstico incorrecto y que emitieron un diagnóstico incorrecto), que para los falsos positivos (aquellos a los que se les pronostica un diagnóstico correcto pero que emitieron un diagnóstico incorrecto). Sin embargo, la ecuación de regresión no serviría para discriminar de entre los sujetos que emitieron un diagnóstico correcto.

Además se repitió el mismo análisis tomando como variable predictora la puntuación obtenida en el índice global de carga. En este análisis resultó que la probabilidad de emitir un diagnóstico correcto podía ser pronosticada a partir del nivel global de carga percibido por los sujetos [$p < 0.0001$] (ver tabla 12.3). En esta ecuación de regresión se obtuvo un 61.87% de casos correctamente clasificados (ver tabla 12.4).

	Chi-Square	df	Significance
-2 Log Likelihood	375.692	297	.0013
Model Chi-Square	34.220	1	.0000
Improvement	34.220	1	.0000
Goodness of Fit	296.921	297	.4904

Variable	B	S.E.	Wald	df	Sig	R	Exp(B)
INDICE	-.0455	.0083	29.7898	1	.0000	-.2604	.955

Tabla 12.3. Resultado del análisis de regresión logística realizado sin distinguir entre ensayos a partir del índice global de carga subjetiva.

Observados	Pronosticados		Porcentaje de correctos
	0	1	
0	126	42	75.00%
1	72	59	45.04%
Total			61.87%

Tabla 12.4. Porcentaje de casos correctos pronosticados a partir del índice global de carga subjetiva. 1=Diagnóstico correcto; 0=Diagnóstico incorrecto.

La tabla 12.4 muestra que el porcentaje de casos correctamente clasificados con esta ecuación de regresión sigue la misma pauta que en el análisis anterior.

12.3. DISCUSION Y CONCLUSIONES

Aunque los factores que no tuvieron efectos sobre la ejecución no los tuvieron tampoco sobre las estimaciones de la carga subjetiva, se encontraron algunas disociaciones entre las medidas de la ejecución en la tarea primaria y las estimaciones subjetivas de la carga. En primer lugar, mientras que en todos los casos el tiempo necesitado para realizar la tarea resultó afectado por el número de elementos del display, sólo en aquellas situaciones más demandantes (display gráfico y numérico con presión temporal) la puntuación global de carga se vio afectada por este factor. En las situaciones menos demandantes (con el display verbal o en ausencia de presión temporal) la carga global percibida por los sujetos fue principalmente función de la complejidad de solución del problema.

De igual forma, el número de diagnósticos correctos emitidos por los sujetos estuvo mayoritariamente afectado por las variaciones en la complejidad de solución del problema. Sin embargo, el índice global de carga sólo resultó afectado por este factor en las situaciones de menor dificultad o menos demandantes.

Por tanto, el análisis de las disociaciones existentes entre las medidas de la ejecución y las de carga subjetiva debería tomar en cuenta el tipo de recursos demandados por la tarea, las capacidades de los sujetos y además la medida de la ejecución (rapidez o precisión) a partir de la cual se van a establecer estas disociaciones.

La presión temporal fue un factor que tuvo un fuerte efecto tanto sobre la rapidez como sobre la calidad de la ejecución, por lo cual puede ser considerada como un factor más de complejidad de la tarea. Sin embargo, este factor no tuvo un efecto explícito sobre la estimación de la carga subjetiva. Su efecto parece ser más indirecto. Es decir, aunque no se obtuvieron diferencias significativas en el nivel global de carga

debidas a la presión temporal, el factor determinante de los sentimientos de carga pasó de ser la complejidad de solución del problema a ser la complejidad perceptiva del display.

El tipo de representación de la información en el display tuvo efectos similares. En condiciones de presión temporal, y con un formato verbal de presentación, la carga global experimentada por los sujetos resultó principalmente determinada por la complejidad de solución del problema, mientras que con el display numérico y con el gráfico fue mayor el efecto de la complejidad perceptiva. Además, la carga global resultó afectada por la interacción entre ambos tipo de complejidad únicamente cuando se utilizó el display verbal.

Por tanto se puede concluir que las puntuaciones globales de carga subjetiva obtenidas a partir de la técnica utilizada son más capaces de detectar las variaciones en la complejidad perceptiva o en la de solución del problema en función de las características y condiciones de la tarea. En general, a medida que las características de la tarea la hacen más difícil para los sujetos que tiene que realizarla, la carga subjetiva será más sensible a las modificaciones en la complejidad perceptiva.

Por otro lado, de las tres dimensiones de carga (mental, temporal y física), sólo la primera, demanda mental, fue objeto de manipulación experimental. El hecho de que en el análisis de regresión logística realizado esta dimensión fuera la única de las tres que entró en la ecuación de regresión para predecir el acierto/fallo de los sujetos, demuestra que los sujetos fueron capaces de distinguir entre los tres tipos de demandas y de estimar adecuadamente la carga asociada a las variaciones en la demanda mental de la tarea. Además, el índice global de carga también resultó predictor de la calidad de la ejecución. Sin embargo, utilizando la estimación global se obtuvo un porcentaje de casos correctamente clasificados inferior al obtenido a partir de la estimación de la demanda mental.

Atendiendo al signo de la correlación que aparece en las tablas 12.1 (correlación entre acierto/fallo y la estimación de la demanda mental) y 12.3 (correlación entre el índice global de carga y el acierto/fallo), se puede concluir que los sujetos que fallaron valoraron con una puntuación más alta la carga mental y global de la tarea, mientras que cuando se producía un acierto las puntuaciones eran inferiores. En principio se podría pensar que este resultado es una consecuencia del feedback puesto que los sujetos conocían la calidad de su respuesta previamente a la valoración de la carga de la tarea. Si este fuera el caso, es decir, si la valoración de las distintas dimensiones de carga hubiera estado guiada única o principalmente por el resultado de la ejecución, las estimaciones de las tres dimensiones de carga habrían entrado en la ecuación de regresión. Además, no se habrían puesto de manifiesto los efectos de la complejidad del sistema sobre la carga que se encontraron mediante el análisis de varianza. Por tanto, a la luz de estos resultados se puede concluir que, aunque posiblemente el conocimiento de los sujetos de la calidad de su ejecución previo a la estimación de la carga de la tarea pudo influir en cierto grado sobre la misma, los niveles de carga percibida medidos con esta técnica fueron función principalmente de la dificultad de la tarea.

Por otro lado, las tablas 12.2. y 12.4. muestran que el porcentaje de casos correctamente clasificados es superior en el caso del fallo, mientras que no se podría clasificar adecuadamente los casos de aciertos. Una posible explicación a este resultado estaría relacionada con las diferencias en la capacidad o recursos de procesamiento de los sujetos. En principio es admisible pensar que los sujetos que diagnostican correctamente el fallo poseen una mayor cantidad de los recursos de procesamiento necesarios para la realización adecuada de la tarea que aquellos que emiten un diagnóstico incorrecto. Debido a que la carga no es función únicamente de la tarea sino también de la capacidad del individuo, parece adecuado pensar que los sujetos que fallan sientan una fuerte carga al realizar los ensayos más difíciles, mientras que sus recursos serán suficientes para realizar con éxito los ensayos más fáciles, experimentando niveles de carga menores. Por el contrario, aquellos sujetos con recursos suficientes para encontrar la solución correcta a los problemas difíciles

sentirán niveles de carga similares en todos los ensayos. Esta diferencia en la capacidad de los sujetos puede ser la responsable de que los sujetos "no eficaces" en la realización de la tarea puedan distinguir mejor entre distintos niveles de dificultad de los problemas.

En conclusión, y teniendo en cuenta las limitaciones encontradas mediante los análisis de varianza, la técnica de medida subjetiva de la carga utilizada permite distinguir entre niveles de complejidad mental de la tarea.

Además, la técnica fue bien aceptada y comprendida por los sujetos. No fue necesario entrenamiento de los participantes y no consumió mucho tiempo. Para completar las tres comparaciones binarias fue necesario menos de medio minuto, y la estimación de cada una de las dimensiones se realizaba en aproximadamente 10 segundos.

CAPITULO 13. CONCLUSIONES

En primer lugar, hay que hacer notar que, como en prácticamente la totalidad de las investigaciones realizadas dentro del campo de la Ergonomía, las conclusiones que aquí se obtienen son específicas tanto del tipo de tarea utilizada como de las características de los sujetos que participaron en la realización de los diferentes experimentos.

El aumento en la complejidad, tanto perceptiva como de solución del problema, de la tarea incrementó el tiempo requerido para emitir el diagnóstico y el número de fuentes de información analizadas por los sujetos. Además se encontraron fuertes efectos de interacción entre ambos tipos de complejidad. Sin embargo, en condiciones de presión temporal, la complejidad perceptiva del display tuvo una mayor incidencia sobre ambas medidas de la ejecución. Por otro lado, el número de soluciones correctas emitidas fue únicamente función de la complejidad de solución del problema. Estos resultados permiten concluir que:

1. La complejidad perceptiva del display es un factor que incide en la complejidad de solución del problema, aunque no es una característica particular del problema en sí mismo.
2. A pesar de que se trataba de una tarea específica de contexto, la no familiaridad de los sujetos con los problemas presentados, hizo que estos optaran por una estrategia de búsqueda mixta (comenzando a partir de una hipótesis sobre el componente que fallaba, y pasando después a una búsqueda topográfica). Sin embargo, en las condiciones más demandantes (bajo presión temporal o en las modalidades de display numérico o gráfico) el uso de una estrategia simplemente topográfica se agudizó.
3. Las personas tienden a observar una mayor cantidad de datos a medida que aumenta el número de fuentes de información presentadas en el display (aunque algunas de ellas sean total y claramente irrelevantes para el diagnóstico y su confirmación), sin embargo este aumento en la información observada no se traduce en una disminución del error de diagnóstico. Esta tendencia se agudiza en situaciones de emergencia.

Por otro lado, no se encontraron diferencias en los niveles de ejecución ni en los de carga subjetiva, debidas al tipo de matriz (normal, traspuesta) utilizada, ni al tipo de feedback proporcionado. Tampoco se encontró un efecto de mejora de la ejecución a medida que transcurría el experimento, en función del tipo de feedback. Estos resultados pusieron de manifiesto la tendencia de los sujetos a infraestimar la información proporcionada por el feedback, sobreestimando sus hipótesis previas, encontrada en la investigación sobre toma de decisiones. Esta tendencia origina la no revisión de los procesos y estrategias de decisión utilizados, hecho que se reafirma cuando, por azar, se produce un acierto.

En cuanto a la carga percibida por los sujetos, se puede concluir que a medida que los niveles de ejecución descendían (por el uso de un determinado display, presencia de la presión temporal, etc.) los índices globales de carga subjetiva aumentaban en función del número de elementos del display. Sin embargo, cuando se producía una disminución en la dificultad de la tarea el aumento de la carga era función principalmente del incremento en la complejidad de solución del problema. En ambos casos, los niveles globales de carga subjetiva obtenidos estuvieron principalmente determinados por la estimación subjetiva de la dimensión mental de la carga, en cuya definición intervienen aspectos tanto perceptivos como de solución del problema.

13.1. PRESION TEMPORAL

La presión temporal fue incluida en los tres últimos experimentos con el fin de determinar en que medida se veían modificadas tanto la ejecución como la carga subjetiva en situaciones de emergencia.

Como se había pronosticado, bajo condiciones de presión temporal se redujo el tiempo de diagnóstico pero el número de respuestas incorrectas emitidas aumentó, a pesar de que se recogió más información sobre los valores de las variables que representaban el estado del sistema. Se encontró que bajo presión temporal las

personas tienden a observar toda la información de la que disponen, aunque gran parte de esta sea totalmente irrelevante para el diagnóstico. Las limitaciones atencionales del sistema de procesamiento humano hacen que tanta información no pueda ser procesada en tan poco tiempo, lo cual produce altos índices de error en el diagnóstico.

Por otro lado, se encontraron fuertes interacciones entre la presión temporal y los factores de complejidad tanto perceptiva como de solución de la tarea, sobre la ejecución. Este resultado confirma la visión de la presión temporal como un factor más de dificultad de las tareas, pero no como la única ni la mayor fuente de complejidad.

Debido al aumento en la complejidad de la tarea, introducido por la presión temporal, los sujetos optaron por un uso más frecuente de la estrategia de búsqueda de información topográfica frente a la sintomática, con el fin de reducir la carga sobre la memoria a corto plazo. El uso de esta estrategia hizo que bajo presión temporal el tiempo necesitado para emitir el diagnóstico dependiera fundamentalmente de la complejidad perceptiva del display, mostrándose los efectos de la complejidad de solución del problema en interacción con ella.

En condiciones de presión temporal se produjeron índices de carga subjetiva más elevados que en la situación de no presión temporal, sin embargo esta diferencia no fue significativa.

13.2. TIPO DE DISPLAY

Se utilizaron tres tipos de display en función del formato de presentación de la información: numérico, verbal y gráfico. La información contenida en los tres displays era la misma, solo variaba el tipo de símbolos utilizados para representarla. Así, en el primero la información era mostrada al sujeto en forma de dígitos o símbolos numéricos. En el display verbal los valores numéricos fueron transformados

en palabras. Por último en el display gráfico se utilizaron diagramas de barras horizontales para representar la información. La ejecución y la carga con los tres displays fue medida bajo condiciones de presión temporal.

No se encontraron diferencias significativas en el nivel de ejecución debidas al tipo de display. Sin embargo, se observó una mejora sistemática en el tiempo requerido para realizar la tarea cuando se utilizaba el display verbal, que no se tradujo en un aumento de los errores, mientras que con el display gráfico se obtuvieron tiempos de diagnóstico muy similares a los del numérico.

Por otro lado, con el display verbal se obtuvieron los menores índices globales de carga subjetiva, mientras que el display numérico fue el que produjo mayores sentimientos de carga en los sujetos.

El que estos resultados fueran contrarios a los encontrados en otras investigaciones se debió principalmente a las diferencias cualitativas en la formación o estudios académicos de los sujetos. Mientras que en la mayor parte de las investigaciones participan estudiantes de Ingeniería, cuya formación es muy técnica, y por tanto cargada de aspectos gráficos y numéricos, en este estudio los sujetos fueron estudiantes de Psicología, cuya formación está fuertemente cargada de contenidos y aspectos de tipo verbal.

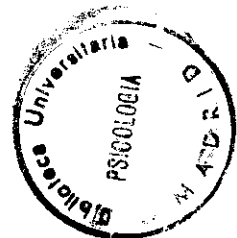
En resumen, el display verbal resultó ser el más eficaz para representar información sobre el estado del sistema a diagnosticar, en condiciones de presión temporal y cuando los sujetos que tienen que realizar la tarea no tienen una alta formación en aspectos técnicos.

13.3. CARACTERISTICAS INDIVIDUALES

En algunos de los experimentos se tomaron diversas medidas de la capacidad de los sujetos. En los dos primeros, estas fueron la capacidad de razonamiento, el estilo cognitivo dependencia-independencia de campo, y la personalidad (neuroticismo, extroversión y sinceridad). En el tercer experimento, se midió la habilidad inicial de los sujetos para manejar el ratón que se iba a utilizar durante la realización de la tarea de diagnóstico de fallos.

La personalidad y el tipo de estilo cognitivo DIC de los sujetos no fueron predictores de la ejecución. La capacidad de razonamiento tampoco resultó ser un buen predictor de la ejecución, ya que aunque entró en las ecuaciones de regresión realizadas, solo explicó una muy pequeña proporción de la varianza de la ejecución. La capacidad de razonamiento fue la variable que correlacionó más alto con las medidas de la ejecución. Como se había previsto, la capacidad de razonamiento correlacionó negativamente con el tiempo y el número de fuentes de información observadas necesitados para emitir el diagnóstico, y positivamente con el número de soluciones correctas dadas a los problemas presentados. Sin embargo, en condiciones de presión temporal estas correlaciones y el poder predictivo de estas variables fue aun menor. Por tanto, en situaciones de emergencia estas diferencias son aun menos relevantes.

Por otro lado, la habilidad inicial de manejo del ratón explicó el 22.8% de la varianza del tiempo de diagnóstico. Los sujetos con mayor habilidad de manejo del ratón (tardaron menos tiempo y cometieron menos errores en la prueba) necesitaron menos tiempo para realizar la tarea de diagnóstico, observaron más fuentes de información y emitieron un mayor número de soluciones correctas.



En conclusión, se confirma la idea de que de todas las habilidades o capacidades individuales, aquellas que tienen una mayor relación directa con la tarea concreta que se ha de realizar, son las más predictoras de la ejecución.

13.4. INCLUSION DE UNA TAREA SECUNDARIA

Con el fin de determinar la estructura de recursos atencionales de la tarea de diagnóstico y para establecer los posibles efectos de la inclusión de una tarea adicional sobre la percepción de la carga, en el cuarto experimento la tarea primaria fue realizada bajo condiciones de tarea simple y de tarea dual.

No se obtuvo ningún efecto significativo ni sobre la ejecución ni sobre la carga subjetiva de la inclusión de la tarea secundaria. La tarea secundaria utilizada demandaba recursos de procesamiento verbal, el input era auditivo y la modalidad de respuesta oral. Debido a que ambas tareas fueron realizadas a la vez sin perjuicio de la ejecución de la tarea primaria, se puede concluir que, como se había supuesto, la tarea primaria demanda recursos de procesamiento de tipo espacial (información gráfica), la modalidad del input es visual y el tipo de respuesta es manual.

La carga global percibida por los sujetos no se vio modificada por la inclusión de la tarea secundaria. Únicamente se encontró un aumento progresivo, pero no significativo, de los sentimientos de desagrado hacia la tarea en la condición dual, mientras que el efecto fue el contrario en la condición de tarea simple.

En resumen, la inclusión de una tarea secundaria que demanda recursos independientes a los demandados por la tarea primaria, no modifica significativamente la ejecución, ni los sentimientos de carga.

13.5. TÉCNICA SUBJETIVA DE MEDIDA DE LA CARGA

En general, la técnica utilizada permitió distinguir entre niveles de complejidad mental de la tarea. Además, fue bien aceptada y comprendida por los sujetos. No fue necesario entrenamiento de los participantes y no consumió mucho tiempo.

Sin embargo no fue directamente sensible a las variaciones en la complejidad producidas por presión temporal. Se encontró una sensibilidad diferencial hacia una forma u otra de complejidad de la tarea (perceptiva y de solución del problema) en función de las condiciones y características de la tarea.

En general, a medida que las características de la tarea la hacen más difícil para los sujetos que tiene que realizarla, la carga global percibida es más sensible a las modificaciones en la complejidad perceptiva, mientras que en aquellas situaciones menos demandantes (en las que mejora la ejecución), como por ejemplo utilizando el display verbal o en condiciones de no presión temporal, las puntuaciones globales de carga obtenidas con esta técnica resultaron mayoritariamente afectadas por las variaciones en la complejidad de solución del problema.

A partir de los resultados de los análisis de regresión logística se puede concluir que los sujetos valoraron la carga de la tarea en función de la complejidad de la misma, sin tener en cuenta el éxito de su ejecución. Por tanto, la técnica fue capaz de distinguir entre los niveles de dificultad, sin que los datos estuvieran contaminados por el efecto de otros factores.

BIBLIOGRAFIA

- ACKERMAN, P.L., SCHNEIDER, W. y WICKENS, C.D. (1984). Deciding the existence of time-sharing ability: a combined metodological and theoretical approach. *Human Factors*, 26, 71-82.
- ADELSON, B. (1981). Problem solving and the development of abstract categories in programming languages, *Memory & Cognition*, 9, 422-433.
- ADELSON, B. (1984). When novices surpass experts: the difficulty of a task may increase with expertise. *J. Exp. Psy.*, 10, 3, 483-495.
- ADELSON, B. y SOLOWAY, E. (1985). The role of domain experience in software design. *IEEE Trans. on Software Engineering*, 11, 1351-1360.
- ADELSON, B. LITTMAN, D. SHRLICH, K. BLACK, J. y SOLOWAY, E. (1984). Novice-expert differences in software design. *Interact'84: First IFIP Conference on H-C Interc.* Amsterdam: Elsevier.
- ALLPORT, D.A., ANTONIS, B. y REYNOLDS, P. (1972). On the division of attention: a disproof of the single channel hypothesis. *Quar. Jour. Exp. Psy.*, 24, 225-235.
- ALLWITT, L.F. (1981). Two neural mechanism related to modes of selective attention. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 324-332.
- ANDRE, A.D. y WICKENS, C. (1988). The interaction of spatial and color proximity in aircraft stability information displays. *Proc. Human Factors Society 32th Annual Meeting*, 1371-1375. Santa Monica, CA: Human Factors Society.
- ANDRE, A.D. y WICKENS, C.D. (1989). *Proximity compatibility and information display: The effects of space and color on the analysis of aircraft stall conditions.* (ARL-89-3/AHEL-89-2). Savoy: Illinois Aviation Research Lab.
- BADDELEY, A.D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology*, 18, 119-130.
- BADDELEY, A. D. y HITCH, G. (1974). Working memory. En G. Bower (ed.), *Recent advances in learning and motivation*, (Vol. 8). Nueva York: Academic.
- BADDELEY, A. D., y LIBERMAN, K. (1980). Spatial working memory. En R. S. Nickerson (Ed). *Attention and performance VII.*Hillsdale, N.J.: Erlbaum.

- BADDELEY, A.D., GRANT, S., WRIGHT, E. y THOMSON, N. (1975). Imagery and visual working memory. En P. Rabbitt y S. Dornic (eds.), *Attention and Performance V*. NY: Academic Press.
- BAHRICK, H.P. y SHELLY, C. (1958). Time-sharing as an index of automatization. *Journal of Experimental Psychology*, 56, 288-293.
- BAINBRIDGE, L. (1978). The process controller. En T. Singleton (Ed.), *The study of real skills*. New York: Academic.
- BAINBRIDGE, L. (1978). Forgotten alternatives in skills and workload. *Ergonomics*, 21, 169-185.
- BALDWIN, R.D. (1978). *Training the electronics maintenance technician* (HumPRO Prof. Paper 7-78). Alexandria, VA: Human Resources Research Organization.
- BARNETT, B. J. y WICKENS, C. D. (1988). Display proximity in multicue information integration: the benefits of boxes. *Hum. Factors*, 30, 15-24.
- BARNETT, B.J., GOETTL, B., KRAMER, A. y WICKENS, C. (1986). *The object display in decision making and judgement*. (CPL-86-7). Urbana, Illinois. Cognitive Psychophysiology Lab.
- BASS, B. M. (1983). *Organizational Decision Making*. Homewood, IL: Irwin.
- BASSILLI, J.N. y REGAN, D.T. (1977). Attributional focus as a determinant of information selection. *J. Soc. Psychology*, 101, 113-121.
- BAUER, D. W., EDDY, J. K. (1986). The representation of command language syntax. *Hum. Factors*, 28, 1-10.
- BEATTY, J. (1982). Task-evolved pupillary responses processing load and the structure of processing resources. *Psychological Bulletin*, 91(2), 276-292.
- BEATTY, J. y KAHNEMAN, D. (1966). Pupillary changes in two memory tasks. *Psychometric Science*, 5, 371-372.
- BEN ZUR, H. y BREZNITZ, S.I. (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica*, 47, 89-104.
- BERENBLUT, B.J. y WHITEHOUSE, H.B. (1977). A method for monitoring process plant based on a decision table analysis. *The Chemical Engineer*, 175-181.

- BERG, O. (1988). *Selected Topics on Development and Evaluation of Man-Machine Systems at the OECD Halden Reactor Project*. International Seminar on Human Interface. Febrero 1988, Japon.
- BERINGER, D. B. (1985). A peripheral integrated status display: Is it really giving the "big picture" or is it a miniseries?. *Proc. Hum. Factors Soc.*, 304-7.
- BERINGER, D. B., CHRISMAN S. E. (1987). A comparison of shape/object displays and conventional univariate indicators: integration benefits or the "nearer to thee" effect?. *Proc. 31st Ann. Meet. Hum. Factors Soc.*, 543-47.
- BERKOWITZ, D., LEWIS, R. y DRURY, C.G. (1983). The effect of computer-displayed knowledge of results on a two-variable optimization task. *Ergonomics*, 26, 10, 975-979.
- BISSERET. A. (1981). Application of signal detection theory to decision making in air traffic control. *Ergonomics*. 24, 81-94.
- BOLES, B. D., WICKENS, C. D. (1987). Display formatting in information integration and nonintegration tasks. *Hum. Factors*, 29, 395-406.
- BOND, N.A. (1981). Troubleshooting in the commercial computer industry: a success story. En J. Rasmussen y W.B. Rouse (eds.), *Human detection and diagnosis of system failures*. NY: Plenum Press.
- BOOHER, H. R. (1975). Relative comprehensibility of pictorial information and printed words in proceduralized instructions. *Human Factors*, 17, 266-277.
- BOULETTE, M. D., COURY, B. G., BEZAR, N. A. (1987). Classification of multidimensional data under time constraints: evaluating digital and configural display representations. *Proc. 31st Ann. Meet. Hum. Factors Soc.*, 116-20.
- BRADSHAW, J.L. (1968). Pupil size and problem-solving. *Quarterly Journal of Experimental Psychology*, 20, 116-122.
- BRATFISCH, O. (1972). *Experienced intellectual activity and perceived difficulty of intelligence test* (Report 30). Suecia: Universidad de Estocolmo, Instituto de Psicología Aplicada.
- BRATFISCH, O., BORG, G. y DORNICK, S. (1972). *Perceived item difficulty in three tests of intellectual performance capacity* (Report 29). Suecia: Universidad de Estocolmo, Instituto de Psicología Aplicada.

- BREHMER, B. (1987). Development of mental models for decision in technological systems. En J. Rasmussen, K. Duncan y J. Leplat (eds.), *New Technology and Human Error*. Chichester: Wiley.
- BROADBENT, D.E. (1958). *Perception and Communication*. Londres: Pergamon.
- BROADBENT, D.E. (1982). Task combinations and selective intake of information. *Acta Psychologica*, 50, 253-296.
- BROOKE, J.B. (1981). Tools for debugging computer programs. How much do they help?. En J. Rasmussen y W.B. Rouse (eds.), *Human detection and diagnosis of system failures*. NY: Plenum Press.
- BROOKE, J.B. Y DUNCAN, K.D. (1980). An experimental study of flowcharts as an aid to identification of procedural faults. *Ergonomics*,
- BROOKE, J.B. y DUNCAN, K.D. (1981). Effects of system display format on performance in a fault location task. *Ergonomics*, 24, 175-189.
- BROOKE, J. B., DUNCAN, K. D. (1983). Effects of prolonged practice on performance in a fault location task. *Ergonomics*. 26. 379-394.
- BROOKE, J.B., COOKE, J.F. y DUNCAN, K.D. (1983). Effect of computer aiding and pretraining on fault location. *Ergonomics*, 26, 669-686.
- BROOKS, L.R. (1968). Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, 22, 349-368.
- BROWN, I. D. (1968). Criticisms of time-sharing techniques for the measurement of perceptual-motor difficulty. *XVI International Congress of Applied Psychology*. Amsterdam: Swets and Zeitlinger, pp. 322-328.
- BRYAN, G.L., BOND, N.A., LAPORTE, H.R. Y HOFFMAN, L. (1956). *Electronics troubleshooting: a behavioral analysis*. Universidad del Sur de California, LA.
- BUTTIGIEG, M., SANDERSON, P.M. y FLACH, J. (1988). Object vs. separate displays for process failure detection: the emergent features approach. *Proc. of Human Factors Society 31st Annual Meeting*. Santa Monica, CA: Human Factors Society.
- CARD, S. K., MORAN, T. P., NEWELL, A. P. (1983). *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Erlbaum Assoc.

- CARROLL, J. M. y OLSON, J. (eds.). (1987). *Mental Models in Human-Computer Interaction: Research Issues About What the User of Software Knows*. Washington, DC: National Academic Press.
- CARSWELL, C. M. Y WICKENS, C. D. (1987). Information integration and the object display: an interaction of tasks demands and display superiority. *Ergonomics*, 30, 511-527.
- CARSWELL, C. M., WICKENS, C. D. (1988). Information integration and the object display. *Ergonomics*, 31, 511-528.
- CARTER, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 127-136.
- CASALI, J. G. y WIERWILLE, W. W. (1982). A sensitive/intrusion comparison of mental workload estimation techniques using a simulated flight task emphasizing perceptual pilot behaviors. *Proc. IEEE International Conference on Cybernetics and Society*, 598-602.
- CASALI, J. G. y WIERWILLE, W. W. (1983). Communications imposed pilot workload: a comparison of sixteen estimation techniques. *Proc. 2th Ohio State University Symposium on Aviation Psychology*, 223-235.
- CASALI, J. G. y WIERWILLE, W. W. (1984). On the measurement of pilot perceptual workload: A comparison of assessment techniques addressing sensitivity and intrusion issues. *Ergonomics*, 27, 1033- 1050.
- CASEY, E. J. (1987). Visual display representation of multidimensional systems: the effects of system structure and display integrality. *Proc. 31st Ann. Meet. Hum. Factors Soc.*, 112-15.
- CASEY, E.J. y WICKENS, C. (1986). *Visual display representation of multidimensional systems*. (CPL-86-2/MDA903-83-K-0255). Illinois. Cognitive Psychophysiology Lab.
- CHASE, W. G. (1986). Visual information processing. En K.R. Boff, L. Kaufman y J.P. Thomas (eds.), *Handbook of Human Perception and Performance*. NY: Wiley.
- CHI, M., FELTOVICH, P.J. y GLASER, R. (1981). Categorization and representation of physics problems by expert and novices. *Cognitive Science*, 5, 121-152.

- CHILDRENS, M. E., HART, S. C., y BORTOLUSSI, M. R. (1982). *Proceedings of the Human Factors Society*, 26, 319-323. The reliability and validity of flight task workload ratings.
- CHILES, W. D. (1982). Workload, task, and situational factors as modifiers of complex human performance. En E. A. Alluisi y E.A. Fleishman (Eds.), *Human performance and productivity*. Hillsdale, NJ: Erlbaum.
- CHRIST, R. E. (1975). Review and analysis of color coding research for visual displays. *Human Factors*. 17, 542-570.
- CHRIST, R. E. (1983). Research for evaluating visual display codes: An emphasis on colour coding. En R. S. Easterby and H. Zwaga (Eds.) *Information design*. Chichester, England: Wiley.
- CHRIST, R. E., CORSO, G. (1983). The effects of extended practice on the evaluation of visual display codes. *Hum. Factors*. 25, 71-84.
- CHRISTENSEN, J.M. Y HOWARD, J.M. (1981). Field experience in maintenance. En J. Rasmussen y W. Rouse (eds.), *Human detection and diagnosis of system failures*, NY: Plenum Press.
- COLQUHOUN, W.P. (1960). temperament, inspection efficiency and time of day. *Ergonomics*, 3, 365-366.
- COOMBS, C.H.; DAWES, R.M. y TVERSKY, A. (1970). *Mathematical Psychology: An Elementary Introduction*. Englewood Cliffs, N.J.: Prentice-Hall.
- COOPER, L.A. (1976). Individual differences in visual comparison processes. *Perception and Psychophysics*, 19, 433-444.
- COOPER, L.A. (1982). Strategies for visual comparison and representation: individual differences. En R.J. Sternberg (ed.), *Handbook of Human Intelligence*. Cambridge: Cambridge University Press.
- COOPER, L.A. y REGAN, D.T. (1982). Attention, perception and intelligence. En R.J. Sternberg (ed.), *Handbook of Human Intelligence*. Cambridge: Cambridge University Press.
- COOPER, G. E. y HARPER, R. P. (1969). *The use of pilot ratings in the evaluation of aircraft handling qualities* (NASA Ames Technical Report NASA TN-D-5153). Moffett Field, CA: NASA Ames Research Center.

- COURY, B.G. y PIETRAS, C.M. (1989). Alphanumeric and graphic displays for dynamic process monitoring and control. *Ergonomics*, 32, 11, 1319-1342.
- COURY, B.G. y PURCELL, J.A. (1988). The bar graph as a configural and separable display. Proc. Human Factors Society 32nd Annual Meeting. Santa Monica, CA: Human Factors Society.
- COURY, B. G., BOULETTE, M. D., ZUBRITZKY, M. C., FISHER, D. L. (1986a). Classification of multidimensional data: evaluating displays to aid identification of system states. *Proc. 16th IEEE Int. Conf.*, 1503-8.
- COURY, B. G., ZUBRITZKY, M. C., SMITH, R. A., CUQLOCK, V. G. (1986). Multidimensional scaling as a method for probing the conceptual structure of state categories. *Proc. 16th IEEE Int. Conf.*, 593-98.
- COURY, B.C., BOULETTE, M.G. y SMITH, R.A. (1989). Effect of uncertainty and diagnosticity on classification of multidimensional data with integral and separable displays of system status. *Human Factors*, 31, 551-569.
- CRABTREE, M.S. (1975). *Human Factors Evaluation of Several Control System Configurations, including workload sharing with Force Wheel Steering during Approach and Flare.* (AFFDL-TR-75-43), OH: USAF Flight Dynamics Lab.
- CRAIK, F. y LOCKHART, F. (1972). Levels of processing: a framework for memory research. *Jour. Verbal Learnng. & Verbal Behavior*, 11, 671-684.
- CROSBY, J. V. y PARKINSON, S. R. (1979). A dual task investigation of pilots' skill level. *Ergonomics*, 22, 1301-1313.
- CURRY, R. E. (1976). A model of the human observer in failure detection tasks. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-6, 85-91.
- CURRY, R. E. (1981). A model of human fault detection for complex dynamic processes. En Rasmussen y Rouse (Eds.), *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- DAMOS, D. (1978). Residual attention as a predictor of pilot performance. *Human Factors*, 20, 435-440.
- DAVIDOFF, J. (1987). The role of color in visual displays. En D. Osborde (Ed.), *International reviews of ergonomics*. 24-42. London: Taylor & Francis.
- DAVIDSON, D. (1977). The effect of individual differences of cognitive style on judgements of document relevance. *J. Am. Soc. Inform. Sci.*, 28, 5, 273-284.

- DAVIS, R. (1959). The role of "attention" in the psychological period. *Quarterly Journal of Experimental Psychology*, 11, 211-220.
- DAVIES, D.R. y HOCKEY, G.R.J. (1966). The effects of noise and doubling the signal frequency on individual differences in visual vigilance performance. *British Journal of Psychology*, 57, 381-389.
- DAWES, R.M. (1979). The robust beauty of improper linear models in decision making. *American Psychologist*, 34, 571-582.
- DAWES, R.M. y CORRIGAN, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81, 95-106.
- de KLEER, J. y BROWN, J. S. (1981). Mental models of physical mechanisms and their acquisition. En J. R. Anderson (ed.), *Cognitive Skills and their Acquisition*. Hillsdale, NJ: Erlbaum.
- DEMAREE, R.G., CROWDER, N.A. y MORRISON, E.J. (1955). *Proficiency of Q-24 radar mechanics: Summary of findings*. (Tech. Memorandum ASPRL TM-55-6) Lowry Air Force Base, CO: ASPRL.
- DERMER, J.D. (1973). Cognitive characteristics and the perceived importance of information. *Accounting Rev.*, 511-519.
- DERRICK, W. L. (1981). The relationship between processing resources and subjective dimensions of operator workload. *Proceedings of the Human Factors Society*, 25, 532-536.
- DERRICK, W. L. (1988). Dimensions of operator workload. *Hum. Factors*, 30, 95-110.
- DEUTSCH, J. A. y DEUTSCH, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.*, 70, 80-90.
- De VEGA, M. (1984). *Introducción a la Psicología Cognitiva*. Madrid: Alianza.
- DIMOND, S.J. y BEAMONT, J. G. (1972). Processing in perceptual integration between and within the cerebral hemispheres. *British Journal of Psychology*, 63, 509-514.
- DONCHIN, E. (1975). Brain electrical correlates of pattern recognition. En G. F. Inbar (Ed.) *Signal analysis and pattern recognition in biomedical engineering*. Nueva York: John Wiley.

- DONCHIN, E. (1984a). The dissociation of electrophysiology and behavior: Adisarter or a challenge? En E. Donchin (Ed), *Cognitive psychophysiology: Event-related potentials and the study of cognition*. The Carmel conferences (Vol. 1). Hillsdale, N. J.: Erlbaum.
- DONCHIN, E. (1984b). The use of ERPs to monitor nonconscious mentation. En S. G. Hart & E. J. Hartswell (Eds.), *Proceedings of the twentieth Annual Conference on Manual Control* (NASACP-2341).
- DONCHIN, E., KRAMER, A., y WICKENS, C. (1982). Probing the cognitive infrastructure with event-related brain potentials. En M. C. Frazier & R. B. Crombie (Eds), *Proceedings of the workshop on flight testing to identify pilot workload and pilot dynamics* (AFFTC-JR 82-5), pp. 371 -387.
- DONNELL, M.L. y col. (1981). *A systems operability measurement algorithm (SOMA): application, validation and extensions* (TR-81-11-156). McLean, Va.: Decision and Desings, Inc.
- DORNIC, S. (1980 a). Language dominance, spare capacity, and perceived effort in bilinguals. *Ergonomics*, 23, 366-377.
- DORNIC, S. (1980 b). *Spare capacity and perceived effort in information processing*. (Report 567). Suecia: Universidad de Estocolmo, Instituto de Psicología Aplicada.
- DORNIC, S. y ANDERSSON, O. (1980). *Difficulty and effort: A perceptual approach* (Report no. 566). Estocolmo, Suecia: Universidad de Estocolmo, Dpto. de Psicología.
- EBERTS, R. (1986). Development of mental models of display augmentation. *Proc. 16th IEEE Int. Conf.* 262-66.
- EBERTS, R.E. (1987). Internal models, tracking strategies and dual-task performance. *Human Factors*, 29(4), 407-419.
- EBERTS, R., SCHNEIDER, W. (1985). Internalizing the system dynamics for a second-order system. *Hum. Factors*, 27, 371-93.
- EGGEMEIER, F. T. (1981). Current issues in subjective assessment of workload. En R. Sugarman, Ed., *Proceedings 25ª Reunión anual de la Human Factors Society*. Santa Mónica, CA: Human Factors Society.

- EGGEMEIER, F. T. (1984). Workload metrics for system evaluation. *Proceedings of the Defense Research Group Panel VIII Workshop "Application of System Ergonomics to Weapon System Development"*, Shrivenham, Inglaterra, C/5-C/20.
- EGGEMEIER, F. T., CRABTREE, M. S. y LaPOINTE, P. A. (1983). The effect of delayed report on subjective ratings of mental workload. *Proceedings 27^a Reunión Anual de la Human Factors Society*, 139- 143.
- EGGEMEIER, F. T., CRABTREE, M. S., ZINGG, J. J., REID, G. B. y SHINGLEDECKER, C. A. (1982). Subjective workload assessment in a memory update task. *Proceedings 26^a Reunión Anual de la Human Factors Society*, 643-647.
- EINHORN, H. J. (1970). The use of nonlinear, noncompensatory models in decision making. *Psychological Bulletin*, 73, 221-230.
- EINHORN, H.J. y HOGARTH, R.M. (1978). Confidence in judgment: persistence of the illusion of validity. *Psychological Review*, 85, 395-416.
- ELLS, J.G. (1973). Analysis of temporal and attentional aspects of movement control. *Journal of Experimental Psychology*, 99, 10-21.
- ELLIS, S. R., MCGREEVY, M. W., HITCHCOCK, R. J. (1987). Perspective traffic display format and airline pilot traffic avoidance. *Hum. Factors*, 29, 371-82.
- ENSTROM, K. y ROUSE, W. (1977). Real time determination of how a human has allocated his attention between control and monitoring tasks. *IEEE Trans. Syst. Man Cybern.*, 7, 153-161.
- EPHRATH, A. R. y YOUNG, L. R. (1981). Monitoring vs. man-in-the-loop detection of aircraft control failures. En J. Rasmussen y W. B. Rouse (Eds.), *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- EYSENCK, H.J. y EYSENCK, S.B.G. (1968). *Eysenck Personality Inventory Manual*. San Diego, CA.: Educational and Industrial Testing Service.
- EYSENCK, H.J. y EYSENCK, S.B.G. (1987). *Cuestionario de Personalidad EPI (5^a Edición revisada)*. Madrid: TEA.
- EYSENCK, M.W. y EYSENCK, M.C. (1979). Processing depth, elaboration of encoding, memory stores, and expended processing capacity. *Jour. Exp. Psy.: Human Learning and Memory*, 5, 422-484.

- FAIRBANK, G., GUTTMAN, N., y MIRON, M. S. (1957). Effects of time comparison upon the comprehension of connected speech. *Journal of Speech and Hearing Disorders*, 22, 10-19.
- FECHNER, G.T. (1860). *Element der Psychophysik*. Leipzig: Breitkopf and Hartel.
- FEDERICO, P.A. (1982). Individual differences in cognitive characteristics and computer-managed mastery learning. *Journal of Computer-Based Instruction*, 9, 10-18.
- FEDERICO, P.A. (1983). Changes in the cognitive components of achievement as students proceed through computer-managed instruction. *Journal of Computer-Based Instruction*, 9, 156-168.
- FEDERICO, P.A. y LANDIS, D.B. (1980). *Relationship among selected measures of cognitive styles, abilities, and aptitudes*. (Tech. report NPRDC TP 80-23). San Diego, CA: Navy Personnel Research and Development Center.
- FERGUSON, R.L. (1980). A reliable plant is a safe plant: lessons learned from TMI. *Proc. Ann. Rel. & Maint. Sym.*, 77.
- FITTS, P. M., POSNER, I. M. (1967). *Human performance*. Belmont, Calif: Brooks Cole.
- FOGARTY, G. (1987). Timesharing in relation to broad ability domains. *Intelligence*, 11, 207-231.
- FOLEY, P., MORAY, N. (1987). Sensation, perception, and system design. En G. Salvendy (ed.), *Handbook of Human Factors*. NY: Wiley.
- FRACKER, L. y WICKENS, C. (1989). Resources, confusions, and compatibility in dual axis tracking: Displays, control, and dynamics. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 80-96.
- FRAISE, P. (1957). La période réfractaire psychologique. *Année Psychologique*, 57, 315-328.
- FRIEDMAN, A., POLSON, M. C. (1981). The hemispheres as independent resource systems: limited capacity processes and cerebral specialization. *J. Exp. Psychol.: Hum. Percept. Perform.*, 7, 1031-58.
- FRIEDMAN, A., POLSON, M. C., DAFOE, C. D., GASKILL, S. J. (1982). Dividing attention within and between hemispheres: testing a multiple resource approach to limited capacity information processing. *J. Exp. Psy.: Perc. Perm.*, 8, 625-50.

- FRJEDMAN, A., POLSON, M. C., GASKILL, S. y DAFOE, C. (1982). Competition for left hemisphere resources. *J. Exp. Psy.: Perc.and Perfor.*, 7, 1031-1051.
- FRWEIN, H.W. (1981). Selective effects of barbiturate and amphetamine on information processing and response execution. *Acta Psychologica*, 47, 105-115.
- GAI, E. G., Y GAI, E. G., Y CURRY, R. E. (1978). Perseveration effects in detection tasks with correlated decision intervals. *IEEE Transactions on systems, Man and Cybernetics*, SMC-8, 93-101.
- GALLWEY, T.J. (1982). Selection test for visual inspection on a multiple fault type task. *Ergonomics*, 25, 11, 1077-1092.
- GARDENIER, J.S. (1981). Ship navigational failure detection and diagnosis. En J. Rasmussen y W.B. Rouse (eds.), *Human detection and diagnosis of System failures*. NY: Plenum Press.
- GARNER, W.R. (1970). The stimulus in information processing. *American Psychologist*, 350-358.
- GARNER, W. R. (1974). *The processing of information and structure*. Hillsdale, NJ: Erlbaum.
- GARNER, W. R. (1978). Selective attention to attributes and stimuli. *J. Exp. Psychol.: Gen.*, 107, 287-308.
- GARNER, W. R. y FEFOLDY, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology*, 1, 225-241.
- GARTNER, W.B. y MURPHY, M.R. (1976). *Pilot Worload and Fatigue: A Critical Survey of Concepts and Assessment Techniques* (NASA TN D-8365) Washington, D.C.: National Aeronautical and Space Administration Ames Research Center.
- GARVEY, W. D. y TAYLOR, F. V. (1959). Interactions among operator variables, system dynamics, and task-induced stress. *J. Applied Psy.*, 43, 79-85.
- GENTNER, D., y STEVENS, A. L. (eds.). (1983). *Mental models*. Hillsdale, NJ: Erlbaum.
- GINSBURG, A.P. (1986). Spatial filtering and visual form perception. En K. Boff, L. Kaufman y J.P. Thomas (eds.), *Handbook of Perception and Human Performance*. New York: Wiley & Sons.

- GLASER, R. y PHILLIPS, J.C. (1954). *An analysis of proficiency for guided missile personnel: III. Patterns of troubleshooting behavior.* (Tech. Bulletin 55-16). Washington, DC: American Institute for Research.
- GOETTL, B., WICKENS, C. y KRAMER, A.F. (1991). Integrated displays and the perception of graphical data. *Ergonomics*, 34, 1047-1063.
- GOLDBECK, R.A., BERNSTEIN, B.B., HILLIX, W.A. y MELVIN, H.M. (1957). Application of the half-spilt technique to problem solving tasks. *J. Exp. Psychol.*, 53, 330-338.
- GOLDSMITH, T.E. y SCHVANEVELDT, R.W. (1984). Facilitating multicue judgments with integral information displays. En J. Thomas y M. Schneider (eds.), *Human Factors in Computer Systems*. Norwood, NJ: Ablex.
- GOODSTEIN, L. P. (1981). Discriminative display support for process operators. En J. Rasmussen y W.B. Rouse (eds.), *Human detection and diagnosis of System failures*. NY: Plenum Press.
- GOPHER, D. (1980). On the training of time-sharing skills: an attention viewpoint. En G. Goornick, M. Hazeltine y R. Durst (eds.), *Proceedings of the Annual Meeting of Human Factors Society*. Santa Monica, California.
- GOPHER, D. y NAVON, D. (1980). How is performance limited: Testing the notion of central capacity. *Acta Psychologica*, 46, 161-180.
- GOPHER, D., DONCHIN, E. (1986). Workload: an examination of the concept. En K. Boff, L. Kaufman y J.P. Thomas (eds.), *Handbook of Perception and Human Performance*. New York: Wiley & Sons.
- GOPHER, D. y KIMCHI, R. (1989). Engineering Psychology. *Ann. Rev. Psychol.*, 40, 431-455.
- GOPHER, D., SANDERS, A. F. (1984). S-Oh-R: Oh stages! Oh resources! En W. Prinz, A. F. Sanders (Eds.), *Cognition and motor processes*, 231-54, Berlin: Springer-Verlag.
- GOPHER, D. y BRAUNE, R. (1984). On the psychophysics of workload: why bother with subjective measures?. *Human Factors*, 26, 5, 519-532.
- GOPHER, D. y BRICKNER, M. (1980). On the training of time-sharing skills: an attention viewpoint. En G. Goornick, M. Hazeltine y R. Durst (eds.), *Proc. Annual Meeting of Human Factors Society*. Santa Monica, California.



- GOPHER, D., BRICKNER, M., y NAVON, D. (1982). Different difficulty manipulations interact differently with task emphasis: evidence for multiple resources. *Jour. Exp. Psy.: Perception and Performance*, 8, 146-157.
- GOTTSDANKER, R.M. y SENDERS, J.W. (1980). *On the Estimation of Mental Load*. (AFOSR-79-0122). Santa Barbara: Universidad de California.
- GRAY, W.C. (1978). Human factors, Oil companies. *International Marine Forum Safe Navigation Symposium*, Washington, D.C.
- GREENO, J. G. (1978). Natures of problem solving abilities. En W. K. Estes (Ed.). *Handbook of Learning and Cognitive Processes*, Vol. 5, (Erlbaum, Hillsdale, NJ).
- GREENO, J. G. y SIMON, H. A. (1984), Problem solving and reasoning. En R. C. Atkinson, R. Herstein, G. Lindzey y D. Luce (Ed.). *Stevens' Handbook of Experimental Psychology* (John Wiley, New York).
- GUGERTY, L. y OLSON, G.M. (1986). Debugging by skilled and novice programmers. *Proc. CHI86: Human Factors in Computing Systems* (ACM, New York).
- HACKER, W., PLATH, H.E., RICHTER, P. y ZIMMER, K. (1978). Internal representation of task structure and mental load of work: approaches and methods of assessment. *Ergonomics*, 21, 194-197.
- HALASZ, F. G. y MORAN, T. P. (1983), Mental models and problem solving in using a calculator. *Proceedings of the Chi'83 Conference on Human Factors in Computing Systems*. (Boston, MA), 212-216.
- HALE, S.L. y BILLMAYER, H.J. (1988). *Use of color CRTs in aircraft cockpits: a literature search*. (Tech. Note 3-88). Aberdeen Proving. Ground, MD: Human Engineering Lab.
- HALLSTEN, L. y BORG, G. (1975). *Six rating scales for perceived difficulty* (Report 58). Suecia: Universidad de Estocolmo, Instituto de Psicología Aplicada.
- HANCOCK, P.A. (1989). The effect of performance failure and task demand on the perception of mental workload. *Applied Ergonomics*, 20, 3, 197-205.

- HANISCH, K.A., KRAMER, A.F. y HULI, C. (1991). Cognitive representations, control, and understanding of complex systems: a field study focusing on components of users' mental models and expert/novices differences. *Ergonomics*, 34, 8, 1129-1145.
- HARRIS, S., OWENS, J. y NORTH, R. (1978). A system for the assesment of human performance in concurrent verbal and manual control task. *Behavior Research Methods and Instrumentation*, 10, 329-333.
- HART, S. (1975). Time estimation as a secondary task to measure workload. *Proceedings of the 11th Annual Conference on Manual Control*, NASA TM X-62, 464.
- HART, S. G. y STAVELAND, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. En P. A. Hancock y N. Meshkati (Eds.), *Human mental workload*. North-Holland, Amsterdam. Pp. 139-183.
- HART, S. G., CHILDRESS, M. E., y BORTOLUSSI, M. (1981). Defining the subjective experience of workload. *Proceedings of the twenty-Fifth Annual Meeting of the Human Factors Society*, 527-531.
- HAYES, J.R. (1964). Human data processing limits in decision making. En E. Bennett (ed.), *Information systems, science and engineering: Proceedings of the first international congress on the information systems science*. NY: McGraw-Hill.
- HELANDER, M. G. (1987). Design of visual displays. En G. Salvendy (ed.), *Cognitive Engineering in the Design of Human-Computer Interaction and Expert Systems*. Amsterdam: Elsevier.
- HELM, W. y DONNELL, M. L. (1979). *Mission operability assessment technique: a system evaluation methodology* (TP-79-31). California: Pacific Missile Test Center.
- HENNEMAN, R.L. y ROUSE, W.B. (1984). Measures of human problem solving performance in fault diagnosis tasks. *IEEE Trans. SMC*, 14, 99-112.
- HERRON, S. (1980). A case for early objective evaluation of candidate display. En G. Corrick, M. Hazeltine y R. Durst (eds.), *Proceedings, 24th Annual Meeting of the Human Factor Society*. Santa Monica, CA: Human Factors.

- HESS, R.A. (1977). Prediction of pilot opinion ratings using an optimal pilot model. *Human Factors*, 19, 459-475.
- HICKS, R.E., MILLER, G.W. y GAIES, G. (1977). Concurrent processing demands and the experience of time in passing. *American Journal of Psychology*, 90, 431-446.
- HIGHLAND, R.W., NEWMAN, S.E. y WALLER, H.S. (1956). A descriptive study of electronic trouble shooting. *Air Force human engineering, personnel, and training research* (Tech. Report 56-8). Baltimore, MD: Air Research and Development Command.
- HOCHBERG, J. (1986). Representation of motion and space in video and cinematic displays. En K. Boff, L. Kaufman y J.P. Thomas (eds.), *Handbook of Perception and Human Performance*. New York: Wiley & Sons.
- HOLLNAGEL, E. (1988). Mental models and model mentality. En L. P. Goodstein, H. B. Andersen, y H. E. Olsen (eds.), *Tasks, Errors and Mental Models*. Londres: Taylor & Francis.
- HUMAN Factors Report on the Tenerife Accident*. (1979). Air Line Pilots Association. Washington, D.C.
- HUNT, E. (1980). Intelligence as an information processing concept. *British Journal of Psychology*, 71, 449-474.
- HUNT, E. (1987). A cognitive model of individual differences with an application to attention. En S.H. Irvine y S.E. Newstead (eds.), *Intelligence and Cognition: Contemporary Frames of Reference*. Dordrecht, Nijhoff.
- HUNT, E. y LANSMAN, M. (1982). Individual differences in attention. En R.J. Sternberg (ed.), *Advances in the Psychology of Human Intelligence*, vol. 1. Hillsdale, N.J.: Lawrence Erlbaum.
- HUNT, E. y FARR, S. (1984). Individual differences in attention. *Proc. Psychonomic Society* (abstract).
- HUNT, R. M., ROUSE, W. B. (1981). Problem-solving skills of maintenance trainees in diagnosing faults in simulated power plants. *Hum. Factors*. 23. 317-328.
- HUNT, R.M., HENNEMAN, R.L. y ROUSE, W.B. (1981). Characterizing the development of human expertise in simulated fault diagnosis tasks. *Proc. Intern. Confer. on Cybernetics and Society*. Atlanta, GA: IEEE SMC Society.

- HUNT, R. G., KRZYSTOFIAK, F. J., MEINDL, J. R. y YOUSRY A. M. (1989). Cognitive style and decision making. *Organizational Behavior and Human Decision Processes*, 44, 436-453.
- IEEE Spectrum. Special Issue on Three Mile Island. (1979). Report of the President' Commission on the Accident at Three Mile Island. Washington, D.C.
- ISREAL, J.B. (1980). *Structural Interference in Dual Task Performance*. Tesis Doctoral. Illinois.
- ISREAL, J. B., CHESNEY, G. L., WICKENS, C. D., y DONCHIN, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual task performance. *Psychophysiology*, 17, 57-70.
- ISREAL, J. B., WICKENS, C. D., CHESNEY, G. L. y DONCHIN, E. (1980). The event-related brain potential as an index of display-monitoring workload. *Human Factors*, 22, 211-224.
- JAGCINSKI, R. J. y MILLER, R. A. (1978), Describing the human operator's internal model of a dynamic system, *Human Factors*, 20, 425-433.
- JAHNS, D.W. (1973). *A concept of operator workload in manual vehicle operations*. Forschungsinstitut Anthropotechnik, Meckenheim, Bericht 14.
- JAMES, W. (1890). *The Principles of Psychology*. New York: Holt.
- JEFFRIES, R. y col. (1981). The processes involved in designing software. En J.R. Anderson (ed.), *Cognitive Skills and their Acquisition*. NY: Erlbaum.
- JENSEN, R, J. (1981). Prediction and quickening in prospective flight displays for curved landing and approaches. *Human Factors*, 23, 333-364.
- JEX, H.R. (1967). Two applications of the critical instability task to secondary task workload research. *IEEE Transactions on Human Factors in Electronics*, 8, 279-282.
- JEX, H.R. (1979). A proposed set of standardized subcritical tasks for tracking workload calibration. En N. Moray (ed.), *Mental Workload: Its Theory and Measurement*. NY: Plenum Press.
- JEX, H. R. y CLEMENT, W. F. (1979). Defining and measuring perceptual-motor workload in manual control tasks. En N. Moray, Ed., *Mental workload, its theory and measurement*. Nueva York: Plenum.

- JOHANSSON, G., MORAY, N., PEW, R., RASMUSSEN, J., SANDERS, A., Y WICKENS, C. (1979). Final report of experimental psychology group. En N. Moray (ed.), *Mental Workload: Its Theory and Measurement*. NY: Plenum Press.
- JOHNSON, E.J., PAYNE, J.W. SCHKADE, D.A. y BETTMAN, J.R. (1989). *Monitoring information processing and decision: the Mouselab System*. Manual versión 4.2.
- JOHNSON, W., GREENBERG, S. FISHER, R. y MARTIN, D. (1970). Divided attention: a vehicle for monitoring memory processes. *Jour. Exp. Psychol.*, 83, 164-171.
- JUBIS, R.M.T. (1990). Coding effects on performance in a process control task with uniparameter and multiparameter displays. *Human Factors*, 32(3), 287-297.
- JUBIS, R. M. T., Y TURNER, D. (1988). *The effectiveness of redundant color coding on search and identification in a process control task* (DCIEM Report No. 88-RR-9). North York, Ontario, Canada: Defence and Civil Institute of Environmental Medicine.
- KAHNEMAN, D. (1973). *Attention and Effort*. Englewood Cliffs, N. J.: Prentice Hall.
- KAHNEMAN, D. y BEATTY, J. (1966). Pupil diameter and load on memory. *Science*, 154, 1583-1585.
- KAHNEMAN, D., Y HENIK, A. (1981). Perceptual organization and attention. En M. Kubovy and J.R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, NJ: Erlbaum.
- KAHNEMAN, D., TREISMAN, A. (1983). Changing views of attention and automaticity. En R. Parasuraman, J. Beatty (Eds.), *Varieties of attention*. New York: Academic.
- KAHNEMAN, D. y PEAVLER, W.S. (1969). Insensitive effects and pupillarity changes in association learning. *Jour. Exp. Psy.*, 79, 312-318.
- KAHNEMAN, D., TVERSKY, B., SHAPIRO, D. y CRIDER, A. (1967). Pupillary, heart rate and skin resistance changes during a mental task. *Jour. Exp. Psy.*, 27, 187-196.

- KAHNEMAN, D., PEAVLER, W.S., y ONUSKA, L. (1968). Effects of verbalization on incentive on the pupil response to mental activity. *Canadian Journal of Psychology*, 22, 186-196.
- KAHNEMAN, D., BEATTY, J. y POLLACK, I. (1967). Perceptual deficit during a mental task. *Science*, 157, 218-219.
- KALSBECK, J. W. H. y SYKES, R. N. (1967). Objective measurement of mental load. *Acta Psychologica*, 27, 253-261.
- KANTOWITZ, B. H. (1987). Mental workload. En P. A. Hancock (ed.), *Human factors psychology*. North-Holland, Amsterdam. Pp. 81-121.
- KANTOWITZ, B. H. y KNIGHT, J. L. (1976). Testing tapping time-sharing: II Auditory secondary task. *Acta Psychologica*, 40, 343-362.
- KARLIN, L. y KASTENBAUM, R. (1968). Effects of number of alterations on the psychological refractory period. *Quarterly Journal of Experimental Psychology*, 20, 167-178.
- KASPERSON, C.J. (1978). An analysis of the relationship between information sources and creativity in scientists and engineers. *Human Commun. Res.*, 4, 2, 113-119.
- KATZ, J.G. (1980). *Pilot workload in the air transport environmet: measurement, theory and the influence of air traffic control*. (FTL- R-80-3). Cambridge, Mass: MIT.
- KAUFMAN, L. (1979). *Sight and Mind*. New York: Oxford.
- KAY, D. S. y BLACK, J. B. (1985). *Knowledge transformations during the acquisition of computer expertise*, CCT Report, 85-1, (Teachers College, Columbia University, Department of Communications, Computing, and Technology, New York).
- KEELE, S.W. (1973). *Attention and Human Performance*. Pacific Palisades, Cal.: Goodyear.
- KESSEL, C. y WICKENS, C. (1978). *Development and utilization of internal models in dynamic systems. A comparison of monitors and opertors as failure detectors*. University of Illinois Engineering Psychology Lab, EPL-78-2/AFOSR-78-5.

- KINSBOURNE, M. y HICKS, R. (1978). Functional cerebral space. En J. Requin (ed.), *Attention and Performance VIII*. Hillsdale, NJ: Erlbaum.
- KLAPP, S. T. (1987). Short-term memory limits in human performance. En P. A. Hancock (ed.), *Human Factors Psychology*. Amsterdam: Elsevier Science.
- KLEINMAN, D.L., BARON, S. y LEVISON, W.H. (1970). An optimal control model of human response. *Automatica*, 6, 357-383.
- KNOWLES, W. B. (1963). Operator loading tasks. *Human Factors*, 5, 151-161.
- KOELEGA, H.S., BRINKMAN, J.A., HENDRIKS, L. y VERBATEN, M.N. (1989). Processing demands, effort, and individual differences in four different vigilance tasks. *Human Factors*, 31, 1, 45-62.
- KOTTEMAN, J. y REMUS, W. (1990). When and how cognitive style impacts decision making. *Management Science*.
- KOUBEK, R. y SALVENDY, G. (1991). Cognitive performance of super-experts on computer program modification tasks. *Ergonomics*, 34, 8, 1095-1112.
- KRAMER, A., WICKENS, C. D., DONCHIN, E. (1983). An analysis of the processing demands of a complex perceptual-motor task. *Hum. Factors*, 25, 597-621.
- KRANTZ, D.H. y TVERSKY, A. (1971). Conjoint measurement analysis of composition rules in psychology. *Psychological Review*, 78, 151-169.
- KREBS, M.J. y WINGERT, J.W. (1976). *Use of the oculometer in pilot workload measurement*. (NASA CR-144951). Washington, D.C.: NASA.
- KRUSKAL, J. B. y WISH, M. (1978). *Multidimensional scaling. University Paper series on Qualitative Applications in the Social Sciences*, 07-011. (Sage, Beverley Hills and London).
- KVALSETH, T. (1980). A decision theoretic model of the sampling behavior of the human process monitor, experimental evaluation. *Human Factors*, 21, 671-686.
- LaBERGE, D., VAN GILDER, P., y YELLOT, S.A. (1971). A cueing technique in choice reaction time. *Jour. Exp. Psy.*, 87, 225-228.
- LANSMAN, M. y HUNT, E. (1983). Individual differences in secondary task performance. *Memory & Cognition*, 11.
- LANSMAN, M., POLTROCK, S.E. y HUNT, E. (1983). Individual differences in the ability to focus and divide attention. *Intelligence*, 7, 299-312.

- LARKIN, J.H. (1983). The role of problem representation in physics. En D. Gentner y A. Stevens (eds.), *Mental Models*. NY: Erlbaum.
- LARKIN, J.H., McDERMOTT, J., SIMON, D.P. y SIMON, H.A. (1980). Expert and novice performance solving physics problems. *Science*, 208, 1135-1342.
- LEBACQZ, J.V. y AIKEN, E.W.A. (1975). *A flight investigation of control, display and guidance requirements for decelerating descending. VTOL instrument (AK-5336-F1)* Buffalo, NY: Calspan Corporation.
- LePLAT, J. (1978). Factors determining workload. *Ergonomics*. 21, 143-149.
- LEVINE, J.M., SCHULMAN, D., BRAHLEK, R.E. y FLEISHMAN, E.A. (1980). *Trainability of abilities: training and transfer of spatial visualization*. (Tech. Report AFHRL-TR-78-47). Valencia, PA: Applied Science Associates.
- LIND, M. (1981). The use of flow models for automated plant diagnosis. En Rasmussen y Rouse (Eds.), *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- LINDGAARD, G. (1987). *Who Needs What Information About Computer Systems: Some Notes on Mental Models, Metaphors and Expertise* (Branch Paper 126), Clayton, Australia: Telecom Australia Research Laboratories.
- LINTON, P.M., JAHNS, D.W. y CHATELIER, P.R. (1977). Operator workload assessment model: an evaluation of a VF/VAA/STOL Systems. En *Methods to Assess Workload*. AGARD-CPP-216, A12/1-A12/11.
- LIHOU, D.A. (1981). Aiding process plant operators in fault finding and corrective action. En J. Rasmussen y W. Rouse (eds.), *Human Detection and Diagnosis of System Failures*. NY: Plenum Press.
- LOGAN, G. D. (1978). Attention in character classification: evidence for the automaticity of component stages. *Journal of Experimental Psychology: General*, 107, 32-63.
- LOGAN, G. D. (1979). On the use of concurrent memory load to measure attention and automaticity. *J. Exp. Psychol.: Hum. Percept. Perform.* 5, 198-207.
- LONG, J. (1976). Division of attention between non-verbal signals: all or none or shared processing. *Quart. Jour. Exp. Psy.*, 28, 47-69.
- LUDER, C.B. y BARBER, P. (1984). Redundant color coding on airborne CRT displays. *Human Factors*, 26, 19-32.

- MADDOX, M.E. y BRICKERY, M.C. (1982). The integration of human factors methodology into Nuclear Power Plant Control Room reviews. *Proc. of Human Factor Society, 26th Annual Meeting*, 654-658.
- MANE, A. M., COLES, M. G. H., WICKENS, C. D., y DONCHIN, E. (1983). The use of the additive factors methodology in the analysis of skill. *Proceedings of the Twenty-Seventh Annual Meeting of the Human Factors Society*, 407-411.
- MARTIN, D. W. (1970). Residual processing capacity during verbal organization in memory. *Journal of Verbal Learning and Verbal Behavior*, 9, 391-397.
- MARTIN, M. (1980). Attention to words in different modalities: four channel presentation with physical and semantic selection. *Acta Psychologica*, 44, 99-115.
- MAY, J.G., KENNEDY, R.S., WILLIAMS, M.C., DUNLAP, W.P. y BRANNAN, J.R. (1990). Eye movement indices of mental workload. *Acta Psychologica*, 75, 75-89.
- MAYER, R. E. (1975). Different problem solving competencies established in learning computer programming with and without meaningful models. *J. Educ. Psychol.* 67, 725-734.
- McCLELLAND, J. L. (1979). On the time-relations of mental processes: An examination of processes in cascade. *Psychological Review*, 86, 287-330.
- McCLELLAND, J. L., RUMELHART, D. E. (1985). Distributed memory and the representation of general and specific information. *J. Exp. Psy.: Hum. Percept. Perform.*, 114, 159-188.
- McCLELLAND, J. L., RUMELHART, D. E., Eds. (1986). *Parallel distributed processing: Explorations in the Microstructure of Cognition. Vol. 2: Psychological and Biological Models*. Cambridge: MA: MIT Press.
- McCLOSKEY, J. (1983). Intuitive physics. En D. Gentner y A. Stevens (eds.), *Mental Models*. Hillsdale, NJ: Erlbaum.
- McDONNELL, J.D. (1968). *Pilot Rating Techniques for the Estimation and Evaluation of Landing Qualities*. (AFFDL-TR-68-76). Dayton, OH: Air Force Flight Dynamics Lab.
- McFARLAND, K. y ASHTON, R. (1978). The influence of concurrent task difficulty on manual performance. *Neurophysiologica*, 16, 735-741.

- MCGREEVY, M. W., ELLIS, S. R. (1986). The effect of perspective geometry on judged direction in spatial information instruments. *Hum. Factors*, 28, 439-56.
- McKEITHEN, K. B. y REITMAN, J. S. (1981). Knowledge organization and skill differences in computer programmers. *Cognitive Psychology*, 13, 307-325.
- McKENNEY, J.L y KEEN, P.G.W. (1974). How managers' minds work. *Harvard Business Rev.*, 79-90.
- McLEOD, P. (1972). A dual task response modality effect: support for multiprocessor models of attention. *Quar. Jour. Exp. Psy.*, 29, 651-667.
- McLEOD, P. D. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *Q. J. Exp. Psychol.*, 30, 83-90.
- McLEOD, P. D. (1978). Does probe RT measure central processing demand?. *Q. J. Exp. Psychol.*, 30, 83-90.
- MICHON, J.A. (1966). Tapping regularity as a measure of perceptual motor load. *Ergonomics*, 8, 401-412.
- MILLER, D.P. y SWAIN, A.D. (1987). Human error and human reliability. En G. Salvendy (ed.), *Handbook of Human Factors*. NY: Wiley.
- MIYATA, Y., y NORMAN, D. A. (1986). Psychological issues in support of multiple activities. En D. A. Norman y S. W. Draper (Ed.), *User-centered system design: New perspectives on human-computer interaction*. Hillsdale, NJ: Erlbaum.
- MOORE, J.V., SATZ, E. y HOEHN, A.J. (1955). Improving equipment maintenance by means of a preplanning technique. (Tech. Report. AFPTRC-TN-55-26). Lackland Air Force Base, TX: Air Force Personnel Training Research Center.
- MORAY, N. (1959). Attention and dichotic listening: affective cues and the influence of instruction. *Quarterly Journal of Experimental Psychology*, 11, 56-60.
- MORAY, N. (1967). Where is capacity limited?. A survey and a model. *Acta Psychologica*, 27, 84-92.
- MORAY, N. (ed.). (1979). *Mental workload: Its theory and measurement*, Nueva York: Plenum.
- MORAY, N. (1980). *Human information processing and supervisory control*. Man-machine Systems Laboratory Report. M.I.T.



- MORAY, N. (1981). The role of attention in the detection of errors and the diagnosis of failures in man-machine systems. En Rasmussen y Rouse (Eds.), *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- MORAY, N. (1982). Subjective mental workload. *Human Factors*, 24, 25-40.
- MORAY, N. (1987). Intelligent aids, mental models, and the theory of machines. *International Journal of Man-Machine Studies*, 27, 619-629.
- MORAY, N., FILTER, OSTRY, D. FAVREAN, D. y NAGY, V. (1976). Attention to pure tones. *Quarterly Journal of Experimental Psychology*, 38, 271-284.
- MORAY, N., JOHANSEN, J., PEW, R.W., RASMUSSEN, J. SANDERS, A.F. y WICKENS, C.D. (1979). Report of the experimental psychology group. En N. Moray (ed.), *Mental Workload: Its Theory and measurement*. NY: Plenum.
- MORRIS, N. M., ROUSE, W. B. (1985). Review and evaluation of empirical research in troubleshooting. *Hum. Factors*, 27, 503-30.
- MORRISON, D.L. y DUNCAN, K.D. (1988). The effect of scrolling, hierarchically paged displays and ability on fault diagnosis performance. *Ergonomics*, 31, 6, 889-904.
- MOSCOVITCH, M. y KLEIN, D. (1980). Material-specific perception for visual words and faces. *Jour. Exp. Psychol.: Perception and Performance*, 6, 590-603.
- MULAİK, S.A. (1972). *The foundations of factor analysis*. NY: McGraw-Hill.
- MULDER, G., MULDER, L. J. (1981). Information processing and cardiovascular control. *Psychophysiology*, 18, 392-401.
- MUNSON, R. C., HORST, R. L. (1986). Evidence for global processing of complex visual displays. *Proc. Hum. Factors Soc.*, pp 776-80.
- NAJJAR, L., PATTERSON, M. y CORSO, G. (1982). redundancy in coding of a visual display as assessed by a signal detection paradigm. *Proc. Hum. Fact. Society 26th Ann. Meeting*. Santa Monica: Human Factors Society.
- NATANI, K., GOMER, F. E. (1981). *Electrocortical activity and operator workload: A comparison of changes in the electroencephalogram and in event-related potentials*. MDC E242, McDonnell Douglas Corp., St. Louis.
- NAVON, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognit. Psychol.*, 9, 353-83.

- NAVON, D. (1981). The forest revisited: More on global precedence. *Psychol. Rev.*, 43, 1-32.
- NAVON, D. (1984). Resources-A theoretical soup stone. *Psychological Review*, 91, 2, 216-234.
- NAVON, D., GOPHER, D. (1979). On the economy of the human processing system. *Psychol. Rev.* 86, 214-55.
- NAVON, D. y GOPHER, D. (1980). Task difficulty, resources, and dual task performance. En R. Nickerson (ed.), *Attention and Performance VIII*. Hillsdale, NJ: Erlbaum.
- NOBLE, M. y SANDERS, A.F. (1980). Searching signals while engaged in compensatory tracking. *Human Factors*, 22, 89-102.
- NORMAN, D. A. (1968). Toward a theory of memory and attention. *Psychol. Rev.*, 75, 522-536.
- NORMAN, D. A. (1983). Design rules based on analyses of human error. *Communications of the ACM*, 26, 254-258.
- NORMAN, D. A. (1983b). Some observations on mental models. En D. Gentner y A. L. Stevens (Ed.). *Mental Models* (Erlbaum, hillsdale, NJ).
- NORMAN, D. A. (1986). Cognitive engineering. En D. A. Norman y S. Draper (eds.), *User-Centered System Design*. Hillsdale, NJ: Erlbaum.
- NORMAN, D. A., BOBROW, D. G. (1975). On data limited and resource limited processes. *Cognit. Psychol.*, 7, 44-64.
- NORMAN, D. A., BOBROW, D. G. (1979). Descriptions: An intermediate stage in memory retrieval. *Cognitive Psychology*, 11, 107-123.
- NORTH, R.A. (1977). *Task Components and Demands as Factors in Dual Task Performance*. (report: ARL-77-2/AFOSR-77-2). Aviation Research Lab.
- NORTH, R.A., y col. (1979). *Performance, physiological and oculometer evaluation of VTOL landing displays* (Report 3171). Langley, Va: NASA Langley Research Center.
- NOTESTINE, J. (1983). *Subjective workload assessment in a probability monitoring task and the effect of delayed ratings*. Unpublished master's thesis, Wright State University, Dayton, Ohio.

- NOTEFINE, J. (1984). Subjective workload assessment and effect of delayed ratings in a probability monitoring task. *Proceedings of the Human Factors Society*, 28, 685-690.
- O'DONNELL, R. y EGGEMEIER, F. T. (1986). Workload assessment methodology. En K. R. Boff, L. Kaufman y J. P. Thomas (Eds.), *Handbook of perception and human performance*, 42, 1-49. Nueva York: Wiley.
- OGDEN, G. D., LEVINE, J. M. y EISNER, E. J. (1979). Measurement of workload by secondary tasks. *Human Factors*, 21, 529-548.
- OSKAMP, S. (1965). Overconfidence in case-study judgments. *Journal of Consulting Psychology*, 29, 261-265.
- PARAMORE, B. y col. (1979). *Task performance problems in reports of collisions, ramblings and groundings in harbors and entrances*. CG-D-28-79, National Technical Information Service. NTIS AD A071658, Virginia.
- PARKS, D.L. (1979). Current workload method and emerging challenges. En N. Moray (ed.), *Mental Workload: Its Theory and Measurement*. NY: Plenum Press.
- PEAVLER, W.S. (1974). Individual differences in pupil size and performance. En M. Jansse (ed.), *Pupillary Dynamics and Behavior*. NY: Plenum Press.
- PEW, R.W. y BARON, S. (1978). The components of an information processing theory of skilled performance based on an optimal control perspective. En G.E. Stelmach (ed.), *Information Processing in Motor Control and Learning*. NY: Academic Press.
- PHILLIP, V., REICHE, D. y KIRCHNER, J. (1971). The use of subjective rating. *Ergonomics*, 14, 611-616.
- POMERANTZ, J. R. (1981). Perceptual organization in information processing. En R. Kubovy y J. R. Pomerantz (Eds.), *Perceptual Organization*. Erlbaum, Hillsdale, NJ.
- POSNER, M. I. (1978). *Chronometric explorations of the mind*. Hillsdale, NJ: Erlbaum.
- POSNER, M.I. y BOIES, S.J. (1971). Components of attention. *Psychological Review*, 78, 391-408.

- POSNER, M.I. y KEELE, S.W. (1969). Attention demands of movements. *Proc. 17Th Annual Congress of Applied Psychology*. Amsterdam: Zeilinger.
- POSNER, M.I. y McLEOD, P. (1982). Information processing models: in search of elementary operators. *Annual Review of Psychology*, 33, 477-514.
- POWRIE, S.E. (1989). Applying cognitive psychology to 3-D modelling in CAD. *Applied Ergonomics*, 20, 4, 252-254.
- PRIBRAM, K.H. y McGUINNESS, D. (1975). Arousal, activation and effort in the control of attention. *Psychol. Rev.*, 82, 116-149.
- PROCIUK, T.J. y BREEN, L.J. (1977). Internal-external locus of control and information seeking in a college academic situation. *J. Soc. Psychology*, 101, 309-310.
- RAHIMI, M. y WIERWILLE, W.W. (1982). Evaluation of the sensitivity and intrusion of workload estimation techniques in piloting tasks emphasizing mediational activity. *Proc. IEEE International Conference on Cybernetics and Society*, 593-597.
- RASMUSSEN, J. (1978). Notes on diagnostic strategies in process plant environment. RIS0-M-1983.
- RASMUSSEN, J. (1979). On the Structure of Knowledge- A Morphology of Mental Models in Man-Machine Context. (Riso Report S-1-69). Roskilde, Dinamarca: Riso National Laboratory.
- RASMUSSEN, J. (1980). *Some trends in man-machine-interface design for industrial process plants* (Tech. Rep. Riso-M- 2228). Riso, Dinamarca: Riso National Laboratory.
- RASMUSSEN, J. (1981). Models of mental strategies in process control. En J. Rasmussen y W. Rouse (eds.), *Human Detection and Diagnosis of System Failures*. NY: Plenum Press.
- RASMUSSEN, J. (1983). Skills, rules, and knowledge: Signals, signs, and symbols, and other distinction in human performance models. *IEEE Trans. Syst. Man Cybern.* 13:257-66.
- RASMUSSEN, J. (1986). *Information processing and human-machine interaction: An approach to cognitive engineering*. New York: North-Holland.

- RASMUSSEN, J. y ROUSE, W. B. (1981). *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- RASMUSSEN, J., JENSEN, A. (1974). Mental procedures in real life tasks: A case study of electronic troubleshooting. *Ergonomics*, 17, 193-207.
- RASMUSSEN, J., LIND, M. (1981). Coping with complexity. In *First European Annual Conference on Human Decision Making and Manual Control*, ed. H. G. Stassen. New York: Plenum.
- REASON, J. (1990). *Human Error*. Cambridge: Cambridge University Press.
- REID, G. B., SHINGLEDECKER, C. A., y EGGEMEIER, F. T. (1981). Application of conjoint measurement to workload scale development. En Sugarman (Ed.), *Proceedings of the Human Factors Society Annual Meeting*, pp. 522-526.
- REID, G. B., EGGEMEIER, F. T. y SHINGLEDECKER, C. A. (1982). Subjective workloadassessment technique. *Proceedings of the 1982 AIAA Workshop on Flight Testing to Identify Pilot Workload and Pilot Dynamics*, 281-288.
- REID, G. B., SHINGLEDECKER, C. A., NYGREN, T. E. y EGGEMEIER, F. T. (1981). Development of multidimensional subjective measures of workload. *Proceeding of the IEEE International Conference on Cybernetics and Society*, 403-406.
- REMYNTOON, R., WILLIAMS, D. (1986). On the selection and evaluation of visual display symbology: factors influencing search and identification time. *Hum. Factors*, 28, 407-20.
- RIGNEY, J.W. (1968). *An analysis of structure and errors in corrective maintenance work*. Universidad del Sur de California, LA.
- RILEY, M. S. (1986), Understanding users' interactions. En D. A. Norman y S. W. Draper (Ed.). *User Centered System Design: New Perspectives on Human-Computer Interaction* (Erlbaum Associates, Inc., Hillsdale, NJ), 157-169.
- ROHMERT, W. (1971). An international symposium on objective assessmentsment of workload in air traffic control task- introduction. *Ergonomics* ,14, 545-547.
- ROHMERT, W., LAURIG, W. (1972). Zusammenhang zwischen belastungsund beanspruchungsparametern bei fluglotsen in der flugverkehrskontrolle. *Arbeitsmedizin, Sozialmedizin, Arbeitshygiene*, 7, 252-257.

- ROLFE, J. M. (1971). Multiple task performance operator overloads. *Occupational Psychology*, 45, 125-132.
- ROLFE, J. M. (1973). The secondary task as a measure of mental load. En W.T. Singleton, J.G. Fox y D. Whitefield (eds.), *Measurement of man at work*. Londres: Taylor & Francis, págs. 135-148.
- ROLFE, J.M., LINDSAY, S.J.E. (1973). Flight deck environment and pilot workload: biological measures of workload. *Applied Ergonomics*, 4, 199-206.
- ROLLINS, R. y HENDRICKS, R. (1980). Processing of words presented simultaneously to eye and ear. *Jour. Exp. Psy.: Perception and Performance*, 6, 99-109.
- ROSCOE, S. N. (1968). Airborne displays for flight and navigation. *Human Factors*, 10, 321-332.
- ROSCOE, A. H. (1978). *Assessing pilot workload*. AGARD-A6-233.
- ROSE, A.M, FINGERMAN, P.W., WHEATON, G.R. EISNER, E. y KRAMER, G. (1974). *Methods for predicting job-ability requirements: II. Ability requirements as a function of changes in the characteristics of an electronic fault finding task*. (TR R74-6). Washington, DC: American Institute for Research.
- ROUSE, W.B. (1978). Human problem solving performance in a fault diagnosis task. *IEEE Trans. on Systems, Man and Cybernetics*, SMC-8, 258.
- ROUSE, W. B. (1979a). Problem-solving performance of maintenance trainees in a fault diagnosis task. *Hum. Factors* 21:195-203.
- ROUSE, W. B. (1979b). A model of human decision making in fault diagnosis tasks that include feedback and redundancy. *IEEE Trans. Syst. Man, Cybern.* 9:237-41.
- ROUSE, W. B. (1981). Experimental studies and mathematical models of human problem solving performance in fault diagnosis tasks. En Rasmussen y Rouse, (Eds.), *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- ROUSE, W. B. (1981b). Human-computer interactions in the control of dynamic system. *Comput. Surv.* 13:71-99.

- ROUSE, W.B., y HUNT, R.M. (1984). Human problem solving in fault diagnosis tasks. En W.B. Rouse (ed.), *Advances in man-machine systems research*, vol. 1. Greenwich, CT: JAI Press.
- ROUSE, W. B. y MORRIS, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, 100, 349-363.
- ROUSE, W. B., ROUSE, S. H. (1979). Measures of complexity of fault diagnosis tasks. *IEEE Trans. Syst. Man Cybern.* 9:720-27.
- ROUSE, W.B. y ROUSE, S. (1982). Cognitive style as a correlate of human problem solving performance in fault diagnosis tasks. *IEEE Trans. SMC*, 12, 649-652.
- ROUSE, W.B. y ROUSE, S. (1984). Human information seeking and design of information systems. *Information Processing and Management*, 20, 129-138.
- RUMMELHART, D. (1977). *Human information processing*. Nueva York: Wiley.
- RUMMELHART, D. y McCLELLAND, J.L. (1986). Parallel distributed processing. En D. Rummelhart y J.L. McClelland (eds.), *Exploration in the microstructure of cognition*, vol. 1.: *Foundations*. Cambridge, Mass: MIT Press.
- SALTZ, E. y MOORE, J.V. (1953). *A preliminary investigation of trouble shooting*. (Tech. Report 53-2). Lackland Air Force Base, TX: Human resources Research Center.
- SANDERS, A. (1979). Some remarks on mental load. En N. Moray (Ed.) *Mental workload: Its theory and measurement*. Nueva York: Plenum Press.
- SANDERS, A. F. (1980). Stage analysis of reaction processes. En G. E. Stelmach y J. Requin (eds.), *Tutorials in Motor Behavior*. Amsterdam: North Holland.
- SANDERS, A. F. (1983). Toward a model of stress and human performance. *Acta Psychologica*, 53, 61-97.
- SANDERS, A.F., WIJMEN, J. y VAN ARKEL, A. (1982). An additive factors analysis of the effects of sleep loss on reaction time. *Acta Psychologica*, 51, 41-59.
- SANDERSON, P. (1986). Designing "mental models" for complex worlds. *Proc. 16th IEEE Int. Conf.*, 267-71.

- SANDERSON, P.M., FLACH, J.M., BUTTIGIEG, M.A. Y CASEY, E.J. (1989). Object displays do not always support better integrated task performance. *Human Factors*, 31, 183-198.
- SAUPE, J.L. (1954). Troubleshooting electronic equipment: an empirical approach to the identification of certain requirements of a maintenance occupation. *Dissertation Abstracts International*, 14.
- SCHANK, R. C. (1982). *Dynamic Memory*. NY: Cambridge University Press.
- SCHMIDT, D.K.A. (1978). A queueing analysis of the air traffic controller's workload. *IEEE Transc. SMC*, 8, 492-498.
- SCHIFLETT, S. G. (1980). *Evaluation of pilot workload assessment device to test alternative display formats and control handling qualities*. Tech. Rep. SY-33R-80. Patuxent River, Md: Naval Air Test Center.
- SCHNEIDER, W. y DETWEILER, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. *Human Factors*, 30, 5, 539-566.
- SCHNEIDER, W. y FISK, A.D. (1982). Concurrent automatic and controlled visual search: can processing occur without resource cost. *Jour. Exp. Psy.: Lear., Mem., Cogn.*, 8, 261-278.
- SCHNEIDER, W. y SHIFFRIN, R.M. (1977). Controlled and automatic human information processing: Detection, search and attention. *Psychological Review*, 84, 1-66.
- SCHNEIDER, W., VIDULICH, M., YEH, y Y. Y. (1985). Training spatial skill for air traffic control. *Proc. Hum. Factors Soc.*, 10-14.
- SCHROEDER, R.G. y BENBASSAT, D. (1975). An experimental evaluation of the relationship of uncertainty to information used by decision makers. *Decision Sciences*, 6, 556-567.
- SCHULTZ, W.C., NEWELL, F.D. y WHITBECK, R.F. (1970). A study of relationship between aircraft system performance and pilot ratings. *Proc. 6th Annual NASA University Conference on Manual Control*, 339-340.
- SCHWARTZ, D. R., HOWELL, W. C.(1985). Optional stopping performance under graphic and numeric CRT formatting. *Hum. Factors*, 27, 433-44.

- SENDERS, J. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *IEEE Transactions on Human Factors in Electronics*, 5, 2-6.
- SENDERS, J. W. (1979). *Personal Communication*.
- SEPPALA, P. y SALVENDY, G. (1985). Impact of depth of menu hierarchy on performance effectiveness in a supervisory task: computerized flexible manufacturing system. *Human Factors*, 27, 6, 713-722.
- SHAFFER, H.L. (1971). Attention in transcription skill. *Quarterly Journal of Experimental Psychology*, 23, 107-112.
- SHAFFER, H.L. (1975). Multiple attention in continuous verbal tasks. En P.M.A. Rabbitt y S. Dornic (eds.), *Attention and Performance V*. Londres: Academic Press.
- SHANNON, C. E. y WEAVER, W. A. (1949). *A Mathematical Model of Communications*. Urbana, ILL: University of Illinois Pres.
- SHEPHERD, A., MARSHAL, E., TURNER, A., DUNCAN, K. (1977). Diagnosing plant failures from a control panel: A comparison of three training methods. *Ergonomics*, 20, 347-61.
- SHERIDAN, T. (1980). Mental workload: What is it? Why bother with it? *Human Factors Society Bulletin*, 23, 1-2.
- SHERIDAN, T. B. (1981). Understanding human error and aiding human diagnostic behavior in nuclear power plants. En J. Rasmussen y W. B. Rouse (Eds.), *Human Detection and Diagnosis of Systems Failures*. Nueva York: Plenum.
- SHERIDAN, T. B. y FERRELL, L. (1974). *Man-Machine Systems. Information Control and Decision Models of Human Performance*. Cambridge, Mass: MIT Press.
- SHERIDAN, T. B. y SIMPSON, R. W. (1979). *Toward the definition and measurement of the mental workload of Transport pilots*. FTR Rep. R-79-4. Cambridge, MA: MIT Flight Transp. Lab.
- SHERIDAN, T.B. Y SIMPSON, R. (1981). *Draft of report on mental workload scaling techniques*. U.S. Federal Aviation Administration.

- SHERIDAN, T.B. y STASSEN, H. (1979). Definitions, models and measures of human workload. en N. Moray (ed.), *Mental Workload: Its Theory and Measurement*. NY: Plenum Press.
- SHINGLEDECKER, C.A. (1983). Behavior and subjective workload metrics for operational environments. *Proc. AGARD (AMP) Symposium on Sustained Intensive Air Operations: Physiological and Performance Aspects*.
- SHINGLEDECKER, C.A. (1987). In flight workload assessment using embedded secondary radio communications tasks. En A.H. Roscoe (ed.), *The practical assessment of pilot workload, AGARDograph 282*, pp. 11-14.
- SHINGLEDECKER, C. A., CRABTREE, M. S., SIMONS, J. C., COURTRIGHT, J. F. y O'DONNELL, R. D. (1980). *Subsidiary radio communications tasks for workload scaling* (Report No. AFAMRI-TR-80- 126). Wright- Paterson Air Force Base, Ohio: U.S. Air Force Aerospace Medical Research Laboratory.
- SKELLY, J.J., REID, G.B. y WILSON, R. (1983). *B-52 full mission simulation: subjective and physiological workload applications*. 2th Aerospace Behavioral Engineering Technology Conference. Long Beach, California.
- SKIPPER, J.H., RIEGER, C.A. y WIERWILLE, W.W. (1986). Evaluation of decision tree rating scales for mental workload estimation. *Ergonomics*, 29, 585-599.
- SMITH, H.T. y CRABTREE, R. (1975). Interactive planning. *Int. J. Man-Machine Stud.*, 7, 213.
- SPELKE, E., HIRST, W., y NEISSER, U. (1976). Skills of divided attention. *Cognition*, 4, 215-230.
- SPERANDIO, J. C. (1978). The regulation of working methods as a function of workload among air traffic controllers. *Ergonomics*, 21, 195-202.
- SPICUZZA, R., PINCUS, A., y O'DONNELL, R.D. (1974). *Development of Performance Assessment Methodology for the Digital Avionics Information System*. Dayton, OH: Systems Research Lab, Inc.
- STERNBERG, S. (1969). The discovery of processing stages: Extensions of Donder's method. *Acta Psychologica*, 30, 276-315.
- SVERKO, B, JERNEIC, Z. y KULENOVIC, R. (1983). A contribution to the investigation of time-sharing ability. *Ergonomics*, 26, 151-160.

- SWAIN, A.D. y GUTTMAN, H.E. (1983). *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*. NUREG/CR 1278. Albuquerque, NM: Sandia National Laboratories.
- TAYLOR, D.A. (1975). Finite fading memory filtering. *IEEE Trans. SMC.*, 5, 134, 137.
- TAYLOR, D.A. (1976). Stage analysis of reaction time. *Psy. Bull.*, 83, 161-191.
- TAYLOR, R. N. (1984). *Behavioral Decision Making*. Glenview, IL: Scott, Foresman.
- TAYLOR, F.M. (1987). Workload management in military cockpits presents insights. *Proc. on Symposium in Military Aircraft*. London: Royal Aeronautical Society.
- THACKRAY, R.I. y TOUCHSTONE, R.M. (1991). Effects of monitoring under high and low taskload on detection of flashing and coloured targets. *Ergonomics*, 34, 8, 1065-1081.
- THOMPSON, D. A. (1981). Commercial aircrew detection of system failures: State of the art and future trends. En J. Rasmussen y W. B. Rouse, Eds., *Human detection and diagnosis of system failures*. Nueva York: Plenum.
- THURSTONE, L.L. y col. (1981). *PMA. Aptitudes Mentales Primarias (5ª Edición)*. Madrid: TEA.
- TOWNSEND, J.T. (1974). Issues and models concerning the processing of a finite number of inputs. En B. Kantowitz (ed.), *Human Information Processing*. Md: Erlbaum.
- TREISMAN, A. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12, 242-248.
- TREISMAN, A. (1964). The effect of irrelevant material on the efficiency of selection listening. *American Journal of Psychology*, 72, 533-546.
- TREISMAN, A. (1965). Monitoring and storage of irrelevant images in selective attention. *Journal of Verbal Learning and Verbal Behavior*, 3, 449-459.
- TREISMAN, A. (1969). Strategies and models of selective attention. *Psy. Rev.*, 76, 282-299.
- TREISMAN, A. (1986). Properties, parts, and objects. En K. Boff, L. Kaufman y J.P. Thomas (eds.), *Handbook of Perception and Human Performance*. New York: Wiley & Sons.

- TREISMAN, A. y DAVIES, A. (1973). Divided attention to ear and eye. En S. Kornblum (ed.), *Attention and Performance IV*. NY: Academic Press.
- TREISMAN, A., y GELADE, G. (1980). A feature integration theory of attention. *Cognitive Psychology*, 12, 97-136.
- TREISMAN, A. SHMIDT, H. (1982). Illusory conjunctions in the perception of objects. *Cognit. Psychol.*, 14, 107-41.
- TRUMBO, D. y MILONE, F. (1971). Primary task performance as a function of encoding retention and recall in secondary task. *Journal of Experimental Psychology*, 91, 273- 278.
- TSANG, P. (1983). *The Structural Constrains and the Strategic Control of Attention Allocation*. Tesis Doctoral. Universidad de Illinois.
- TULGA, M.K. (1978). Dynamic decision making in multitask supervisory control: comparison of an optimal algorithm to human behavior. Cambridge, MA: MIT.
- TULVING, E. y LINDSAY, P.H. (1967). Identification of simultaneously presented simple visual and auditory stimuli. *Acta Psychologica*, 27, 101-109.
- TYLER, S., HERTAL, P., McCALUM, M. y ELLIS, H. (1979). Cognitive effort and memory. *Jou. Exp. Psy.: Human Learning and Memory*, 5, 607-617.
- UNDERWOOD, G. (1976). *Attention and Memory*. Oxford, England: Permagon.
- VELDHUYZEN, W. y STASSEN, H.G. (1977). The internal model concept: an application to modelling human control of large ships. *Human Factors*, 19, 367.
- VESSEY, I. (1985). Expertise in debugging computer programs. a process analysis. *International Journal of Man-Machine Studies*, 23, 459-494.
- VIDULICH, M.K. (1986). The cognitive psychology of subjective mental workload. En P.A. Hancock y N. Meshkati (eds.), *Human Mental Workload*. Amsterdam: Elsevier.
- VIDULICH, M.K. (1988). Speech responses and dual-task performance: better time-sharing or asymmetric transfer?. *Human Factors*, 30, 4, 517-529.
- VIDULICH, M. K., y WICKENS, C. D. (1981). *Time sharing, manual control and memory search: The effects of input and output modality competition priorities and control order* (Tech. Rep. EPL-81-41/ONR-81-4). University of Illinois, December.

- VIDULICH, M. A. y WICKENS, C. D. (1986). Causes of dissociation between subjective measures and performance: caveats for the use of subjective assessments. *Appl. Ergon.*, 17, 291-96.
- VIDULICH, M. A. y TSANG, P. S. (1986). Techniques of subjective workload assessment: A comparison of SWAT and the NASA-Bipolar methods. *Ergonomics*, 29, 1385-1398.
- VINGE, E. y PITKIN, e. (1972). Human operator for aural compensatory tracking. *IEEE Transactions on Systems, Man and Cybernetics*, 2, 504-512.
- VOLONIO, L. (1987). An Empirical Investigation of the Impacts of Information Load on Decision Quality and Decision Confidence in Multi-Criteria Decision Making. Tesis Doctoral no publicada. School of Management, State University of New York, Buffalo.
- WALDEN, R.S. y ROUSE, W.B. (1978). A queuing model of pilot decision making in a multi task flight management situation. *IEEE Trans. on SMC*, 8, 867.
- WARD, S.L. y POTURALSKI, R.J. (1982). Changes in maze-solving errors due to stress. *Proceedings of the 26th Annual Meeting of Human Factors Society*, 479-482.
- WEINER, E.L. (1977). Controller Flight into Terrain Accidents. Systems induced errors. *Human Factors*, 19, 171-181.
- WEISER, M. y SHERTZ, J. (1983). Programming problem representation in novice and expert programmers. *International Journal of man-machine Studies*, 19, 391-398.
- WELFORD, A. T. (1959). Evidence of a single-channel decision mechanism limiting performance in a serial reaction task. *Quarterly Journal of Experimental Psychology*, 2, 193-210.
- WELFORD, A. T. (1967). Single channel operation in the brain. *Acta Psychologica*, 27, 5-21.
- WELFORD, A. T. (1968). *Fundamentals of Skills*. London: Methuen & Co.
- WELFORD, A. T. (1978). Mental workload as a function of demand, capacity, strategy, and skill. *Ergonomics*, 21, 151-67.

- WEST, B., CLARK, J. A. (1974). Operator interaction with a computer controlled distillation column. En E. Edwards y F.P. Lees (eds.), *The Human Operator in Process Control*. London: Taylor & Francis.
- WHITAKER, L. (1979). Dual task interference as a function of cognitive processing load. *Acta Psychologica*, 43, 71-84.
- WHITFIELD, D., BALL, R. G., ORD, G. (1980). Some human factors aspects of computer-aiding concepts for air traffic controllers. *Hum. Factors*, 22, 569-80.
- WICKENS, C.D. (1976). The effects of divided attention on information processing in tracking. *Jour. Exp. Psy.: Percep. and Perform.*, 1, 1-13.
- WICKENS, C. D. (1979). Human workload measurement. En N. Moray (Ed.) *Mental workload: Its theory and measurement*. Nueva York: Plenum Press.
- WICKENS, C. D. (1980). The structure of attentional resources. En R. Nickerson (Ed.) *Attention and performance VIII*. Englewood Cliffs, N. J.: Erlbaum.
- WICKENS, C. D. (1981). Workload: In defense of the secondary task. *Personnel Training and Selection Bulletin*, 2, 119-123.
- WICKENS, C. D. (1983). Processing resources in attention. En R. Parasuraman, J. Beatty y R. Davies (eds.), *Varieties of Attention*. NY: John Wiley & Sons.
- WICKENS, C. D. (1984a). *Engineering psychology and human performance*. Columbus, OH: Charles Merrill.
- WICKENS, C. D. (1984b). Processing resources in attention. En R. Parasuraman, D. R. Davis. (Eds.), *Varieties of Attention*. Orlando, Fla: Academic.
- WICKENS, C. D. (1986a). Gain and energetics in information processing. En G.R. Hockey, A. Gaillard y M. Coles (eds.), *Energetics and Human Information Processing*. Dordrecht: Martinus Nijhoff.
- WICKENS, C. D. (1986b). *The object display: principles and a review of experimental findings*, Cognitive Psychophysiology Technical Report CPL-86-6. Department of Psychology, University of Illinois, Champaign, IL.
- WICKENS, C. D. (1987). Information processing, decision making, and cognition. En G. Salvendy (ed.), *Cognitive Engineering in the Design of Human-Computer Interaction and Expert Systems*. Amsterdam: Elsevier.

- WICKENS, C.D. y LIU, Y. (1988). Codes and modalities in multiple resources: a success and qualification. *Human Factors*, 30, 5, 599-616.
- WICKENS, C. y DERRICK, W. (1981). Workload measurement and multiple resources. *Proc. IEEE Conference on Cybernetics and Society*. NY: Institute of Electrical and Electronics Engineers, Inc.
- WICKENS, C. D. y KRAMER, A. (1985). Engineering psychology. *Annual Review of Psychology*, 36, 307-348.
- WICKENS, C. D. y YEH, Y. Y. (1982). The dissociation of subjective ratings and performance. *Proc. IEEE International Conference on Cybernetics and Society*. Seattle, Wash., 584-587.
- WICKENS, C. D. y YEH, Y. Y. (1983). The dissociation of subjective ratings and performance: A multiple resources approach. En L. Haugh y A. Pope, Eds., *Proceedings*, 27ª Conferencia Anual de la Human Factors Society. Santa Mónica, CA: Human Factors Society.
- WICKENS, C. y SANDRY, D.L. (1982). Task hemispheric integrity in dual task performance. *Acta Psychologica*, 52, 227-248.
- WICKENS, C. D., KRAMER, A., VANASSE, L., DONCHIN, E. (1983). Performance of concurrent task: A psychophysiological analysis of reciprocity of information processing resources. *Science*, 221, 1080-82.
- WICKENS, C. D., MOUNTFORD, S. J., SCHREINER, W. (1981). Multiple resources, task-hemispheric integrity, and individual differences in time-sharing. *Hum. Factors*, 23, 211-29.
- WICKENS, C. D., SANDRY, D. y VIDULICH, M. (1983). Compatibility and resource competition between modalities of input, central processing, and output: Testing a model of complex task performance. *Human Factors*, 25, 227-248.
- WICKENS, C. D. y KESSEL, C. (1981). Failure detection in dynamic systems. En Rasmussen y Rouse, (Eds.), *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- WICKENS, C. D. y GOPHER, D. (1977). Control theory measures of tracking as indices of attention allocation strategies. *Human Factors*, 19, 349-365.

- WICKENS, C., TSANG, P. y BENEL, R. (1979). The dynamics of resource allocation. *Proc. 23th Ann. Meeting Human Factors Society*. Santa Monica, California: Human Factors Press.
- WICKENS, C. D. Y BOLES, D. B. (1983). The limits of multiple resource theory: the role of task correlation/integration in optimal display formatting. *Engineering Psychology Laboratory Technical Report EPL-83-5/ONR-83-5*. University of Illinois, Champaign, IL.
- WICKENS, C. D. Y ANDRE, A. D. (1990). Proximity compatibility and information display: effects of color, space, and objectness on information integration. *Human Factors*, 32, 61-77.
- WICKENS, C. D., VIDULICH, M. Y SANDRY-GARZA, D. (1984). Principles of S-C-R compatibility with spatial and verbal tasks: the role of display control location and voice interactive display control interfacing. *Hum. Factors.*, 26, 533-43.
- WICKENS, C. D. y KESSEL, C. (1980). The processing resource demands of failure detection in dynamic systems. *Journal of Experimental Psychology: Human Perception and Performance*, 6, 564-577.
- WICKENS, C. D., KESSEL, C. (1979). The effects of participatory mode and task workload on the detection of dynamic system failures. *IEEE Trans. Syst. Man Cybern.*, 9, 24-34.
- WICKENS, C., GILL, R., KRAMER, A., ROSS, W., y DONCHIN, E. (1981). The processing demands of higher order manual control. En J. Lyman y A. Besczy (eds.), *17th Annual Conference on Manual Control*. JPL 81-3. Pasadane, Calif.: Jet Propulsion Lab.
- WICKENS, C., HYMAN, F, DELLINGER, J., TAYLOR, H. y MEADOR, M. (1985). The Sternberg memory search task as an index of pilot workload. *Proc. Symposium on Aviation Psychology*, 2, 287-294.
- WIERWILLE, W. y CASALI, J. (1983). A validated rating scale for global mental workload measurement applications. En L. Haugh y A. Pope, Eds., *Proceedings of the Human Factors Society*. Santa Mónica, CA: Human Factors Society.

- WIERWILLE, W. W., RAHIMI, M. y CASALI, J. G. (1985). Evaluation of 16 measures of mental workload using a simulated flight task emphasizing mediational activity. *Human Factors*, 27, 489-502.
- WIERWILLE, W. W. y WILLIGES, R. C. (1978). *Survey and analysis of operator workload assessment techniques*. (Report No.S- 78-101). Blacksburg, VA: Systemetrics.
- WIERWILLE, W. W. y WILLIGES, B. H. (1980). *An annotated bibliography on operator mental workload assessment*. (Report SY- 27R-80). Patuxent River, MD: Naval Air Test Center.
- WIERWILLE, W. W., CONNOR, S. A. (1983). Evaluation of 20 workload measures using a psychomotor task in a moving-base aircraft simulator. *Hum. Factors*, 25, 1-16.
- WILKINSON, R.T. (1961). Comparison of paced, unpaced irregular and continuous display in watch keeping. *Ergonomics*, 4, 259-267.
- WILLIGES, R. C., WIERWILLE, W. W. (1979). Behavioral measures of aircrew mental workload. *Hum. Factors*, 21, 549-74.
- WILSON, J.R. y RUTHERFORD, A. (1989). Mental models: Theory and application in human factors. *Human Factors*, 31(6), 617-634.
- WITKIN, H.A. y col. (1982). *Test de figuras enmascaradas*. Madrid: TEA
- WOHL, J. G. (1981). System complexity, diagnostic behavior, and repair time: A predictive theory. En Rasmussen y Rouse (Eds.), *Human Detection and Diagnosis of System Failures*. Nueva York: Plenum.
- WOHL, J. (1982). Maintainability, prediction revisited: Diagnostic behavior, system complexity, and repair time. *IEEE Trans. Syst. Man Cybern.* 12:240-50.
- WOLFE, J. D. (1978). *Crew workload assessment: Development of a measure of operator workload* (Report No. AFDL-TR-78-165). Wright-Patterson Air Force Base, Ohio: Air Force Flight Dynamics Laboratory.
- WOODS, D. D. (1986). Paradigms for intelligent decision support. En Hollnagel, G. Mancini, y D. D. Woods (Ed.), *intelligent decision support in process environments*. New York: Springer-Verlang

- WOODS, D., WISE, J. y HANES, L. (1981). An evaluation of nuclear power plant safety parameter displays systems. En R. C. Sugarman (ed.), *Proceedings 25ª Reunión Anual de la Human Factors Society*. Santa Mónica, CA: Human Factors Society.
- WORKMAN, D., FISHER, D. L. (1987). Selection of display symbology: a new metric of similarity. *Proc. 31st Ann. Meet. Hum. Factors Soc.*, 510-13.
- WRIGHT, P. (1974). The harassed decision maker: time pressures, distractions and the use of evidence. *Journal of Applied Psychology*, 59, 555-561.
- YEH, Y. Y. y WICKENS, C. D. (1984). Why do performance and subjective workload measures dissociate? *Proceedings 28ª Reunión Anual de la Human Factors Society*. Santa Mónica, CA: Human Factors Society.
- YEH, Y. Y., WICKENS, C.D. (1988) Dissociation of performance and subjective measures of workload. *Hum. Factors*, 30, 111-20.
- YOUNG, R. M. (1983). Surrogates and mappings: Two kinds of conceptual models for interactive devices. En D. Gentner y A. L. Stevens (eds), *Mental Models*. Hillsdale, NJ: Erlbaum.
- ZAKAY, D. y WOOLER, S. (1984). Time pressure, training and decision effectiveness. *Ergonomics*, 27, 3, 273-284.
- ZWAGA, H.J.G. y DULJNHOUWER, F. (1984). The influence of a fixed background in a VDU display on the efficiency of colour and shape coding. En C.P. Gibson (ed.), *Proc. NATO Workshop on Color-Coded versus Monochrome Electronic Displays*. England: Royal Aircraft Establishment.

APENDICE 1

PRETEST

NOMBRE: _____

INSTRUCCIONES

Imagina que te encuentras en un edificio cuyo sistema de conducción de agua se compone de los siguientes elementos:

- . llave de paso general,
- . depósito de agua,
- . bomba eléctrica, y
- . tubería distribuidora del agua a los distintos pisos.

Tu tarea consistirá en enumerar todas las posibles causas de cada uno de los problemas que aparecen a continuación, y que pueden darse en este sistema, debidas al mal funcionamiento de alguno de sus elementos. Una vez que las hayas enumerado, debes valorar (de 0 a 10) la probabilidad de ocurrencia de cada una de dichas causas, entendiendo que 0 significa la probabilidad nula y 10 la máxima, siendo 5 la probabilidad intermedia.

Al final de la prueba aparecerá un espacio en blanco en el cual podrás sugerir otros problemas, distintos a los anteriores, y que, a tu parecer, puedan darse en este tipo de sistemas. Para cada uno de ellos deberás enumerar sus posibles causas, así como la probabilidad de ocurrencia de cada una de ellas, del mismo modo que en los anteriores.

Si tienes alguna duda, este es el momento de plantearla.

MUCHAS GRACIAS POR TU COLABORACION.

1.- NO HAY AGUA EN NINGUN PISO DEL EDIFICIO.

Causas:

Probabilidad:

2.- EL AGUA SALE SUCIA.

Causas:

Probabilidad:

3.- EL TERCER PISO TIENE HUMEDAD EN EL TECHO.

Causas:

Probabilidad:

4.- EL AGUA TIENE MAL OLOR.

Causas:

Probabilidad:

5.- EL AGUA TARDA BASTANTE TIEMPO EN SALIR.

Causas:

Probabilidad:

6.- EL VECINO DEL ULTIMO PISO TIENE MANCHAS DE HUMEDAD EN EL TECHO DE SU VIVIENDA.

Causas:

Probabilidad:

7.- EL SUELO DEL PISO MAS BAJO SE HA LEVANTADO.

Causas:

Probabilidad:

8.- SOLO HAY AGUA EN LOS PISOS BAJOS PERO NO EN LOS MAS ALTOS.

Causas:

Probabilidad:

9.- SALE MAS CANTIDAD DE AGUA DE LO HABITUAL.

Causas:

Probabilidad:

10.- EL AGUA SALE MAS CALIENTE DE LO NORMAL.

Causas:

Probabilidad:

11.- EL SEGUNDO PISO TIENE HUMEDAD EN LA PARED.

Causas:

Probabilidad:

12.- EL AGUA TIENE UN SABOR EXTRAÑO.

Causas:

Probabilidad:

13.- EL FLUJO DE AGUA ES DISCONTINUO.

Causas:

Probabilidad:

14.- SE HA ENCONTRADO HUMEDAD CERCA DE LA ENTRADA GENERAL
AL EDIFICIO.

Causas:

Probabilidad:

15.- OTROS.

DEFINICION DE LAS DIMENSIONES

1-DEMANDA MENTAL:

Nivel de actividad mental (recordar, decidir, memorizar, calcular, etc...) y perceptiva (mirar, buscar, distinguir, etc...) que requiere la tarea. ¿Se trata de una tarea fácil o difícil, simple o compleja, exigente o relajada, respecto a esta dimensión?.

2-DEMANDA TEMPORAL:

Presión temporal sentida debida al tiempo disponible para realizar cada ensayo de la tarea. ¿Es mucho tiempo, poco, suficiente?. ¿En que medida el tiempo disponible para realizar la tarea te hace sentirte inseguro/a o molesto/a?.

3-DEMANDA FISICA:

Nivel de actividad física que requiere la tarea (tirar, empujar, pulsar, mover, girar, deslizar, etc...) para alcanzar los objetivos o metas establecidas por el investigador (o por ti mismo). ¿Se trata de una tarea fácil o difícil, aburrida o amena, demasiado lenta o muy rápida, descansada o laboriosa?. ¿Estas satisfecho/a con tu nivel de rendimiento?. ¿Hasta que punto te sientes molesto/a o agradado/a con la tarea?.

Compara los siguientes pares de dimensiones, y señala cual de los dos elementos del par (A o B) te parece, en principio, más importante para la tarea experimental.

- | <u>A</u> | <u>B</u> | |
|----------------|-----------------|-------|
| 1) DEM. MENTAL | - DEM. TEMPORAL | |
| 2) DEM. FISICA | - DEM. TEMPORAL | |
| 3) DEM. MENTAL | - DEM. FISICA | |

EJEMPLO DE ESCALA

VALORA LA DEMANDA MENTAL DE LA TAREA EN LA SIGUIENTE ESCALA:

63.88

0 10 20 30 40 50 60 70 80 90 100

Enter this box and click a mouse button when you're ready

Figura 7. Ejemplo del tipo de escala utilizada para valorar las demandas de la tarea.

APENDICE 2

INSTRUCCIONES

Imagina que eres un técnico de control y mantenimiento de los sistemas de conducción de agua de diversos edificios. En este experimento, tu tarea consistirá principalmente en DIAGNOSTICAR, en función de los datos que aparezcan para cada ensayo, EL FALLO que se produce en el sistema, y elegir una de las cuatro posibles soluciones a dicho problema (que aparecerán en la parte inferior de la pantalla).

En cada uno de los ensayos aparecerá, en primer lugar, una breve descripción del problema. En la siguiente pantalla aparecerá, en forma de matriz, el valor que tiene cada componente del sistema en una serie de indicadores. En cada caso, los componentes del sistema serán: DEPOSITO de agua, LLAVE DE PASO general, BOMBA eléctrica y TUBERIA distribuidora del agua a los distintos pisos. Sus indicadores serán: PRESION DE ENTRADA y PRESION DE SALIDA (valores de 0 a 9), NIVEL O FLUJO de agua (en %), y NUMERO DE REPARACIONES que ha tenido el componente.

La bomba funcionará adecuadamente cuando aumente la presión del agua en dos unidades como máximo. Los niveles adecuados de agua son 90-95%

NO TODOS LOS EDIFICIOS TIENEN EL MISMO NUMERO DE COMPONENTES NI EL MISMO NUMERO DE VARIABLES.

Una vez elegida una de las alternativas deberás responder, en algunos casos, a 3 preguntas que aparecerán consecutivamente. Si tienes alguna duda este es el momento de plantearla. Procura trabajar con precisión y recuerda que no hay tiempo limite para realizar la tarea.

EJEMPLOS DE DISPLAYS

	PRES. ENTRADA	PRES. SALIDA	NIVEL/FLUJO	REPARACIONES
DEPOSITO				
LLAVE				
BOMBA		+ ?		

Choose one: Nueva llave Nueva bomba Nuevo deposito No hay fallo

Figura 1. Ejemplo de un display del experimento 1, matriz normal 3x4.

	DEPOSITO	LLAVE	BOMBA
PRES. ENTRADA			
PRES. SALIDA			+ ?
NIVEL/FLUJO			
REPARACIONES			

Choose one: Nueva llave Nueva bomba Nuevo deposito No hay fallo

Figura 2. Ejemplo de un display del experimento 1, matriz traspuesta 3x4.

TABLA 3

F. Variac.	TT		ND		NA	
	F(1,16)	Signif.	F(1,16)	Signif.	F(1,16)	Signif.
M	0.39	0.540	0.85	0.371	0.20	0.660
F	0.30	0.568	0.39	0.539	0.29	0.598
M x F	0.01	0.934	3.84	0.068	0.30	1.000

Tabla 1. Resultados del ANOVA para los factores entresujetos, para cada una de las variables dependientes. M=Tipo de matriz; F=Tipo de feedback; TT=Tiempo de tarea; ND=Número de datos recogidos; NA=Número de aciertos. Se observa que ninguno de los efectos resultó significativo.

F. Variac.	g.l	TT		ND		NA	
		F	Signif.	F	Signif.	F	Signif.
CP	3,48	23.59	0.000	9.79	0.000	1.40	0.255
NC	1,16	18.65	0.001	11.87	0.004	8.17	0.011
PM	1,16	15.21	0.001	11.78	0.000	16.15	0.001
CPxNC	3,48	16.97	0.000	0.77	0.515	6.77	0.001
CPxPM	3,48	7.94	0.003	2.26	0.093	0.76	0.523
NCxPM	1,16	2.92	0.107	0.04	0.852	3.98	0.063
CPxNCxPM	3,48	17.15	0.000	0.70	0.555	6.34	0.797

Tabla 2. Tabla de ANOVA resumida de los factores intrasujeto, para cada variable dependiente. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media; TT=Tiempo total de tarea; ND=Número de datos observados; NA=Número de aciertos.

F.Variac.	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
	F(1,16)	Signif.	F(1,16)	Signif.	F(1,16)	Signif.	F(1,16)	Signif.
M	0.40	0.536	0.15	0.701	0.41	0.530	0.89	0.360
F	0.81	0.382	0.89	0.360	0.02	0.884	0.75	0.399
M x F	0.54	0.472	0.02	0.898	2.22	0.155	0.06	0.810

Tabla 3. Resultados del ANOVA para los factores entresujetos, para cada una de las dimensiones de carga y para el índice global de carga. M=Tipo de matriz; F=Tipo de feedback. Se observa que ninguno de los efectos resultó significativo.

F. Variac.	g.l	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
		F	Signif.	F	Signif.	F	Signif.	F	Signif.
CP	3,48	2.90	0.044	3.64	0.019	3.86	0.015	2.39	0.080
NC	1,16	4.17	0.058	1.03	0.326	1.51	0.236	2.47	0.136
PM	1,16	13.90	0.002	13.32	0.002	11.64	0.004	20.34	0.000
CPxNC	3,48	1.67	0.187	3.91	0.014	4.19	0.010	2.33	0.086
CPxPM	3,48	0.69	0.562	0.30	0.825	1.41	0.251	0.50	0.682
NCxPM	1,16	1.49	0.239	1.31	0.269	0.57	0.462	1.21	0.287
CPxNCxPM	3,48	0.82	0.490	0.81	0.493	1.44	0.243	0.76	0.519

Tabla 4. Tabla de ANOVA resumida de los factores intrasujeto, para cada dimensión de la carga subjetiva y para el índice global de carga. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media.

R Múltiple	.26115				
R Cuadrado	.07903				
R Cuadrado Ajust.	.07029				
Error típico	14.98363				
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regresión	3	6068.03399	2029.34466		
Residual	316	70944.93350	224.50928		
F =	9.03502	Signif F =		.0000	
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
R	-.70331	.16747	-.24092	-4.230	.0000
NEU	.77158	.23217	.18109	3.323	.0015
PO	.62752	.24079	.14831	2.636	.0096
--- Variables no en la Ecuación ---					
Variable	Beta In	Partial	Min Toler	T	Sig T
EXT	.07679	.07292	.82098	1.298	.1953
SIN	-.09422	-.09111	.86120	-1.624	.1054

Tabla 5. Resultados del análisis de regresión realizado para la variable criterio Tiempo de Tarea. R= Razonamiento; PO= Figuras Ocultas; NEU= Neuroticismo; EXT= Extroversión; SIN= Sinceridad.

R Múltiple	.16935				
R Cuadrado	.03162				
R Cuadrado Ajust.	.02732				
Error típico	3.90188				
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regresión	2	159.40199	79.70099		
Residual	317	4582.01928	14.45432		
F =	5.81199	Signif F =		.0044	
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
R	-.12150	.04110	-.18170	-3.126	.0019
PO	.12468	.06102	.11878	2.043	.0419
--- Variables no en la Ecuación ---					
Variable	Beta In	Partial	Min Toler	T	Sig T
NEU	-.04414	-.04449	.88557	-.792	.4291
EXT	.06547	.03345	.85025	1.489	.1376
SIN	.04180	.04196	.87862	.747	.4559

Tabla 6. Resultados del análisis de regresión realizado para la variable criterio Número de Datos. R= Razonamiento; PO= Figuras Ocultas; NEU= Neuroticismo; EXT= Extroversión; SIN= Sinceridad.

R Múltiple		.29504			
R Cuadrado		.08705			
R Cuadrado Adjust.		.08418			
Error típico		.63055			
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regresión	1	12.05483	12.05483		
Residual	318	126.43267	.39759		
F =	30.31998		Signif F =	.0000	
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
R	.03652	6.63210E-03	.29504	5.506	.0000
--- Variables no en la Ecuación ---					
Variable	Beta In	Partial	Min Toler	T	Sig T
FO	.01701	.01691	.90212	.301	.7636
NEU	-.09294	-.09650	.98411	-1.726	.0853
EXT	-.04365	-.04516	.97743	-.805	.4215
SIN	-3.272E-02	-.00338	.97614	-1.060	.9520

Tabla 7. Resultados del análisis de regresión realizado para la variable criterio Número de Acuerdos. R=Razonamiento; FO= Figuras Ocultas ; NEU= Neuroticismo; EXT= Extroversión; SIN= Sinceridad.

APENDICE 3

INSTRUCCIONES

Imagina que eres un técnico de control y mantenimiento de los sistemas de conducción de agua de diversos edificios. En este experimento, tu tarea consistirá principalmente en DIAGNOSTICAR, en función de los datos que aparezcan para cada ensayo, EL FALLO que se produce en el sistema, y elegir una de las cuatro posibles soluciones a dicho problema (que aparecerán en la parte inferior de la pantalla).

En cada uno de los ensayos aparecerá, en primer lugar, una breve descripción del problema. En la siguiente pantalla aparecerá, en forma de matriz, el valor que tiene cada componente del sistema en una serie de indicadores. En cada caso, los componentes del sistema serán: DEPOSITO de agua, LLAVE DE PASO general, BOMBA eléctrica y TUBERIA distribuidora del agua a los distintos pisos. Sus indicadores serán: PRESION DE ENTRADA y PRESION DE SALIDA (valores de 0 a 9), NIVEL O FLUJO de agua (en %), y NUMERO DE REPARACIONES que ha tenido el componente.

La bomba funcionará adecuadamente cuando aumente la presión del agua en dos unidades como máximo. Los niveles adecuados de agua son 90-95%.

NO TODOS LOS EDIFICIOS TIENEN EL MISMO NUMERO DE COMPONENTES NI EL MISMO NUMERO DE VARIABLES.

Una vez elegida una de las alternativas deberás responder, en algunos casos, a 3 preguntas que aparecerán consecutivamente. Si tienes alguna duda este es el momento de plantearla. Procura trabajar con precisión. Dispones de 30 seg. para realizar cada ensayo.

EJEMPLOS DE DISPLAYS

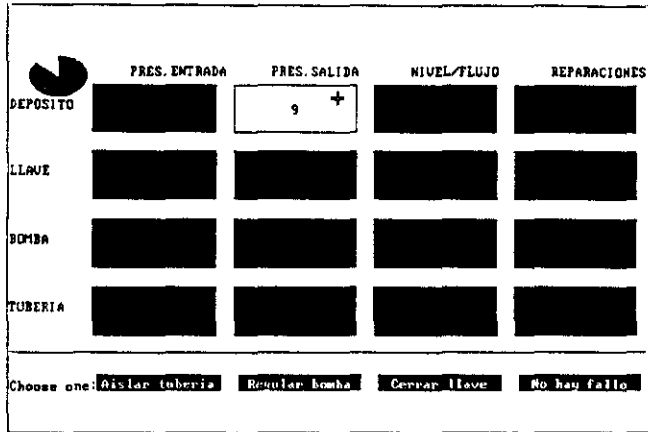


Figura 1. Ejemplo de un display del experimento 2, matriz normal 4x4.

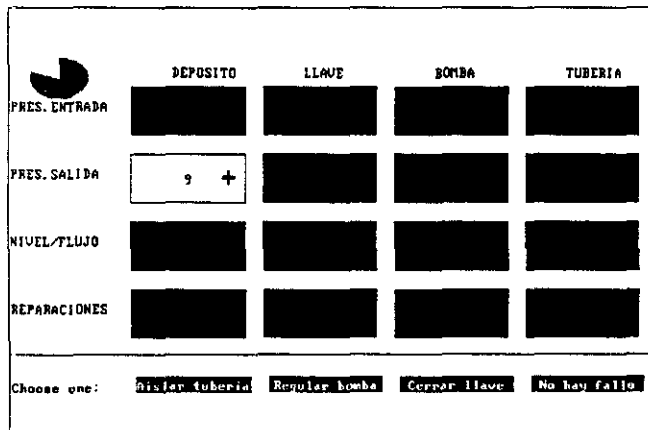


Figura 2. Ejemplo de un display del experimento 2, matriz traspuesta 4x4.

TABLAS

F. Variac.	g.l	TT		ND		NA	
		F	Signif.	F	Signif.	F	Signif.
PT	1,32	27.23	0.000	1.43	0.241	10.21	0.003
PTXCP	3,96	6.86	0.000	1.35	0.263	0.66	0.577
PTXNC	1,32	11.78	0.002	1.16	0.290	0.34	0.562
PTXPM	1,32	23.13	0.000	15.39	0.000	0.17	0.680
PTXCPXNC	3,96	4.68	0.004	0.12	0.949	1.00	0.397
PTXCPXPM	3,96	1.16	0.330	0.67	0.572	1.51	0.218
PTXNCXPM	1,32	0.06	0.812	5.52	0.025	5.49	0.026
PTXCPXNCXPM	3,96	8.02	0.000	0.44	0.724	0.40	0.751

Tabla 1. Tabla de ANOVA resumida de los efectos simple y de interacción de la presión temporal para cada variable dependiente. PT=Presión Temporal; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media; TT=Tiempo de Tarea; ND=Número de Datos; NA=Número de Aciertos.

F. Variac.	g.l	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
		F	Signif.	F	Signif.	F	Signif.	F	Signif.
PT	1,32	3.23	0.082	2.67	0.112	0.15	0.701	3.12	0.087
PTXCP	3,96	0.52	0.673	0.32	0.511	1.10	0.353	0.27	0.848
PTXNC	1,32	2.22	0.146	0.04	0.841	0.31	0.582	0.37	0.550
PTXPM	1,32	0.18	0.675	2.79	0.104	0.41	0.527	1.92	0.175
PTXCPXNC	3,96	0.17	0.917	1.96	0.123	1.01	0.392	0.41	0.747
PTXCPXPM	3,96	0.32	0.808	0.84	0.476	0.41	0.744	0.40	0.756
PTXNCXPM	1,32	0.12	0.733	0.27	0.613	0.91	0.348	0.64	0.428
PTXCPXNCXPM	3,96	0.46	0.707	1.76	0.160	0.27	0.850	0.45	0.715

Tabla 2. Tabla de ANOVA resumida de los efectos simple y de interacción de la presión temporal para cada dimensión de la carga subjetiva y para el índice global de carga. PT=Presión Temporal; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media.

F. Variac.	TT		ND		NA	
	F(1,16)	Signif.	F(1,16)	Signif.	F(1,16)	Signif.
M	4.50	0.052	0.70	0.415	3.29	0.089
F	0.26	0.615	1.03	0.325	0.05	0.819
M x F	1.30	0.272	0.61	0.915	1.35	0.262

Tabla 3. Resultados del ANOVA para los factores entresujetos. M es el tipo de matriz y F el tipo de feedback. TT=Tiempo de Tarea; ND=Número de Datos; NA=Número de Aciertos.

F. Variac.	g.l	TT		ND		NA	
		F	Signif.	F	Signif.	F	Signif.
CP	3,48	7.08	0.000	6.61	0.001	1.41	0.252
NC	1,16	1.62	0.221	8.28	0.011	1.13	0.304
PM	1,16	2.89	0.108	36.46	0.000	25.98	0.000
CPxNC	3,48	10.92	0.000	0.42	0.738	3.11	0.035
CPxPM	3,48	4.65	0.036	1.45	0.241	1.37	0.262
NCxPM	1,16	8.45	0.010	8.95	0.003	1.73	0.207
CPxNCxPM	3,48	3.63	0.019	2.38	0.081	0.80	0.502

Tabla 4. Tabla de ANOVA resumida de los factores intrasujeto, para cada variable dependiente. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media; TT=Tiempo de Tarea; ND=Número de Datos; NA=Número de Aciertos.

F.Variac.	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
	F(1,16)	Signif.	F(1,16)	Signif.	F(1,16)	Signif.	F(1,16)	Signif.
M	0.44	0.515	0.14	0.712	0.21	0.649	0.21	0.653
F	0.06	0.815	0.10	0.759	1.13	0.304	0.00	0.952
M x F	0.26	0.616	0.01	0.920	0.12	0.732	0.06	0.811

Tabla 5. Resultados del ANOVA para los factores entresujetos, para cada una de las dimensiones de carga y para el índice global de carga. M=Tipo de matriz; F=Tipo de feedback. Se observa que ninguno de los efectos resultó significativo.

F. Variac.	g.l	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
		F	Signif.	F	Signif.	F	Signif.	F	Signif.
CP	3,48	0.94	0.428	6.33	0.001	0.89	0.452	3.23	0.030
NC	1,16	0.01	0.914	3.76	0.070	0.28	0.605	1.41	0.252
PM	1,16	7.56	0.014	1.642	0.218	6.63	0.020	4.08	0.061
CPxNC	3,48	1.74	0.171	2.74	0.054	1.02	0.391	2.12	0.110
CPxPM	3,48	0.64	0.593	2.47	0.073	1.26	0.300	1.52	0.220
NCxPM	1,16	1.47	0.243	1.36	0.260	0.49	0.496	1.82	0.196
CPxNCxPM	3,48	0.20	0.893	6.53	0.001	1.39	0.257	1.88	0.146

Tabla 6. Tabla de ANOVA resumida de los factores intrasujeto, para cada dimensión de la carga subjetiva y para el índice global de carga. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media.

R Múltiple	.22254				
R Cuadrado	.04570				
R Cuadrado Adjust.	.04371				
Error típico	8.09422				
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regression	2	1066.28972	543.14486		
Residual	317	20768.69366	65.51638		
F =	8.29021	Signif F =	.0003		
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
EXT	.41539	.13129	.17359	3.164	.0017
R	-.28711	.10395	-.15155	-2.762	.0061
--- Variables no en la Ecuación ---					
Variable	Beta In	Parcial	Min Toler	T	Sig T
FO	.03168	.02955	.82681	.526	.5995
NEU	-4.2E-04	-.00040	.84352	-.007	.9843
SIN	.05531	.05649	.99063	1.006	.3153

Tabla 7. Resultados del análisis de regresión realizado para la variable criterio Tiempo de Tarea. R= Razonamiento; FO= Figuras Ocultas; NEU= Neuroticismo; EXT= Extroversión; SIN= Sinceridad.

R Múltiple	.22254				
R Cuadrado	.04570				
R Cuadrado Adjust.	.04371				
Error típico	8.09422				
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regression	1	67.93488	67.93488		
Residual	318	4062.81199	12.77614		
F =	5.91733	Signif F =	.0218		
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
R	-.10563	.104581	-.11624	-2.306	.0218
--- Variables no en la Ecuación ---					
Variable	Beta In	Parcial	Min Toler	T	Sig T
FO	-.01195	-.01136	.86792	-.202	.8399
NEU	-.01988	-.01892	.89072	-.337	.7364
EXT	-2.9E-03	-.00292	.99563	-.052	.9586
SIN	.04596	.04624	.99533	.824	.4105

Tabla 8. Resultados del análisis de regresión realizado para la variable criterio Número de datos. R= Razonamiento; FO= Figuras Ocultas; NEU= Neuroticismo; EXT= Extroversión; SIN= Sinceridad

APENDICE 4

INSTRUCCIONES

Imagina que eres un técnico de control y mantenimiento de los sistemas de conducción de agua de diversos edificios. En este experimento, tu tarea consistirá en DIAGNOSTICAR, en función de los datos que aparezcan para cada ensayo, EL FALLO que se produce en el sistema, y elegir una de las cuatro posibles soluciones a dicho problema (que aparecerán en la parte inferior de la pantalla).

En cada ensayo aparecerá, en primer lugar, una breve descripción del problema. En la siguiente pantalla aparecerá, en forma de matriz, el valor que tiene cada componente del sistema en una serie de indicadores. En cada caso, los componentes del sistema serán: DEPOSITO de agua, LLAVE DE PASO general, BOMBA eléctrica y TUBERIA distribuidora del agua a los distintos pisos. Sus indicadores serán: PRESION DE ENTRADA y PRESION DE SALIDA (nula, baja, media, alta y excesiva) y NIVEL O FLUJO de agua (nulo, bajo, medio, alto y excesivo).

La bomba funcionará adecuadamente cuando aumente la presión del agua en un nivel como máximo. Los niveles adecuados de agua son altos.

NO TODOS LOS EDIFICIOS TIENEN EL MISMO NUMERO DE COMPONENTES.

Una vez elegida una de las alternativas deberás responder a 3 preguntas que aparecieran consecutivamente. Si tienes alguna duda este es el momento de plantearla. Procura trabajar con precisión y recuerda que dispones de 30 segundos para realizar cada ensayo.

EJEMPLO DE DISPLAY

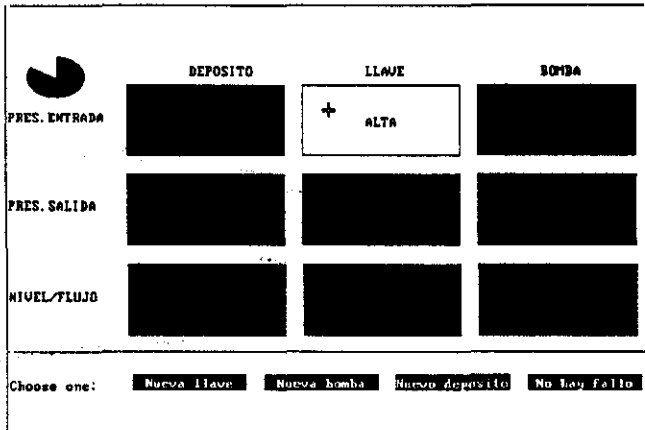


Figura 1. Ejemplo de un display del experimento 3, matriz 3x3.

TABLES

F. Variac.	g.l	TT		ND		NA	
		F	Signif.	F	Signif.	F	Signif.
TI	1,28	0.97	0.334	0.01	0.543	0.05	0.828
TIxCP	1,28	0.10	0.752	0.47	0.499	0.15	0.699
TIxNC	1,28	1.12	0.298	0.18	0.674	0.77	0.388
TIxPM	1,28	2.04	0.165	0.84	0.368	0.57	0.457
TIxCPxNC	1,28	0.46	0.503	0.01	0.917	0.07	0.790
TIxCPxPM	1,28	0.18	0.672	2.91	0.099	1.34	0.256
TIxNCxPM	1,28	0.00	0.954	0.01	0.910	0.33	0.572
TIxCPxNCxPM	1,28	0.07	0.791	2.74	0.109	1.68	0.205

Tabla 1. Tabla de ANOVA resumida de los efectos del tipo de información para cada variable dependiente. TI=Tipo de información; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media; TT=Tiempo de Tarea; ND=Número de Datos; NA=Número de Aciertos.

F. Variac.	g.l	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
		F	Signif.	F	Signif.	F	Signif.	F	Signif.
TI	1,28	1.62	0.214	4.37	0.046	3.08	0.090	4.57	0.041
TIxCP	1,28	0.45	0.509	1.56	0.222	0.05	0.954	2.01	0.166
TIxNC	1,28	4.78	0.037	0.05	0.825	0.02	0.901	0.97	0.334
TIxPM	1,28	0.99	0.329	0.85	0.363	1.02	0.322	0.91	0.347
TIxCPxNC	1,28	0.04	0.835	1.62	0.198	0.04	0.851	0.24	0.629
TIxCPxPM	1,28	0.05	0.818	0.45	0.508	0.02	0.899	0.43	0.515
TIxNCxPM	1,28	0.10	0.756	1.16	0.291	0.02	0.894	0.49	0.489
TIxCPxNCxPM	1,28	0.00	0.966	0.02	0.890	0.09	0.597	0.04	0.848

Tabla 2. Tabla de ANOVA resumida de los efectos del tipo de información, para cada dimensión de la carga subjetiva y para el índice global de carga. TI=Tipo de información; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media.

F. Variac.	g.l	TT		ND		NA	
		F	Signif.	F	Signif.	F	Signif.
CP	1,14	4.59	0.042	30.92	0.000	1.35	0.265
NC	1,14	0.33	0.576	2.25	0.156	0.38	0.547
PM	1,14	1.72	0.211	13.99	0.002	2.22	0.094
CPXNC	1,14	15.15	0.002	0.34	0.571	4.04	0.064
CPXPM	1,14	7.00	0.019	1.44	0.250	0.24	0.629
NCXPM	1,14	14.57	0.002	0.57	0.462	0.80	0.388
CPXNCXPM	1,14	25.10	0.000	0.19	0.668	0.20	0.663

Tabla 3. Tabla de ANOVA resumida de los factores intrasujeto, para cada variable dependiente. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media; TT=Tiempo de Tarea; ND=Número de Datos; NA=Número de Aciertos.

F. Variac.	g.l	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
		F	Signif.	F	Signif.	F	Signif.	F	Signif.
CP	1,14	0.30	0.591	4.54	0.051	1.25	0.301	1.80	0.202
NC	1,14	2.45	0.137	0.01	0.944	1.76	0.206	0.71	0.414
PM	1,14	13.01	0.003	2.88	0.172	0.00	0.953	7.35	0.017
CPXNC	1,14	0.25	0.624	5.70	0.032	0.36	0.557	1.72	0.210
CPXPM	1,14	1.13	0.306	0.28	0.602	0.41	0.533	0.00	0.998
NCXPM	1,14	0.10	0.751	0.98	0.338	3.13	0.099	0.26	0.618
CPXNCXPM	1,14	2.40	0.144	5.38	0.036	8.86	0.010	14.27	0.002

Tabla 4. Tabla de ANOVA resumida de los factores intrasujeto, para cada dimensión de la carga subjetiva y para el índice global de carga. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media.

R Múltiple		.47790			
R Cuadrado		.22839			
R Cuadrado Adjust.		.21520			
Error típico		7.03690			
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regresión	2	1714.83422	857.41711		
Residual	117	5793.60217	49.51797		
F =	17.31527	Signif F =	.0000		
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
Nº ERRORES	.29481	.07154	.34666	4.121	.0001
TIEMPO	.03078	.01033	.25075	2.961	.0035

Tabla 5. Resultados del análisis de regresión realizado para la variable criterio Tiempo de Tarea.

R Múltiple		.29988			
R Cuadrado		.08993			
R Cuadrado Adjust.		.08222			
Error típico		3.67508			
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regresión	1	157.48780	157.48780		
Residual	118	1593.73587	13.50624		
F =	11.66038	Signif F =	.0009		
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
TIEMPO	-.01778	5.205E-03	-.29988	-3.415	.0009
--- Variables no en la Ecuación ---					
Variable	Beta In	Partial	Min Toler	T	Sig T
Nº ERRORES	-.06650	-.06729	.93199	-.730	.4671

Tabla 6. Resultados del análisis de regresión realizado para la variable criterio Número de Datos.

R Múltiple	.23145				
R Cuadrado	.08357				
R Cuadrado Adjust.	.04555				
Error típico	.67491				
Análisis de Varianza					
	DF	Suma de Cuadrados	Media Cuadrática		
Regresión	1	3.04232	3.04232		
Residual	118	53.74936	.45550		
F =	6.67903	Signif F =		.0110	
--- Variables en la Ecuación ---					
Variable	B	SE B	Beta	T	Sig T
TIEMPO	-2.47E-03	9.56E-04	-.23145	-2.584	.0110
--- Variables no en la Ecuación ---					
Variable	Beta In	Partial	Min Toler	T	Sig T
N* ERRORES	-.02793	-.02771	.93199	-.100	.7648

Tabla 7. Resultados del análisis de regresión realizado para la variable criterio Número de Aciertos.

APENDICE 5

INSTRUCCIONES

Imagina que eres un técnico de control y mantenimiento de los sistemas de conducción de agua de diversos edificios. En este experimento, tu tarea consistirá en DIAGNOSTICAR, en función de los datos que aparezcan para cada ensayo, EL FALLO que se produce en el sistema, y elegir una de las cuatro posibles soluciones a dicho problema (que aparecerán en la parte inferior de la pantalla).

En cada ensayo aparecerá, en primer lugar, una breve descripción del problema. En la siguiente pantalla aparecerá, en forma de matriz, el valor que tiene cada componente del sistema en una serie de indicadores. En cada caso, los componentes del sistema serán: DEPOSITO de agua, LLAVE DE PASO general, BOMBA eléctrica y TUBERIA distribuidora del agua a los distintos pisos. Sus indicadores serán: PRESION DE ENTRADA, PRESION DE SALIDA y NIVEL O FLUJO de agua.

La bomba funcionará adecuadamente cuando aumente la presión del agua, pero sin que este aumento sea demasiado grande (no debe aumentar en más de 4X). Los niveles adecuados de agua son altos.

NO TODOS LOS EDIFICIOS TIENEN EL MISMO NUMERO DE COMPONENTES.

Una vez elegida una de las alternativas deberás responder a 3 preguntas que aparezcan consecutivamente. Si tienes alguna duda este es el momento de plantearla. Procura trabajar con precisión y recuerda que dispones de 30 segundos para realizar cada ensayo.

Además, y en el momento en el que el experimentador te indique, tendrás que atender a otra tarea que el experimentador te explicará a continuación.

EJEMPLO DE DISPLAY

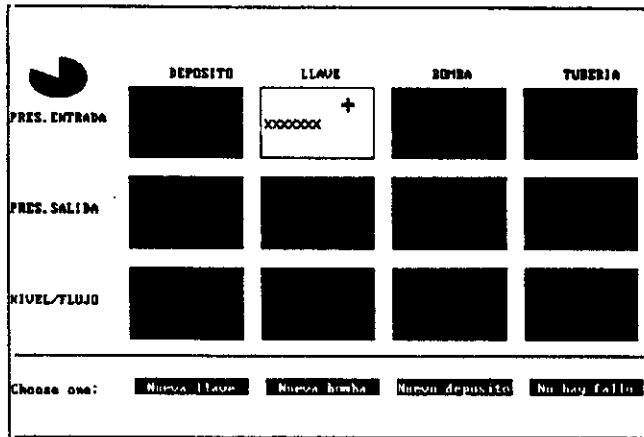


Figura 1. Ejemplo de un display del experimento 4, matriz 3x4.

TABLAS

F. Variac.	g.l	TT		ND		NA	
		F	Signif.	F	Signif.	F	Signif.
TI	2,39	0.78	0.464	1.50	0.236	0.54	0.588
TIxCP	2,39	0.21	0.812	0.51	0.605	0.17	0.841
TIxNC	2,39	0.48	0.625	0.14	0.872	1.32	0.279
TIxPM	2,39	0.82	0.450	1.77	0.184	0.75	0.479
TIxCPxNC	2,39	0.94	0.398	0.78	0.463	0.78	0.465
TIxCPxPM	2,39	1.09	0.347	1.94	0.157	0.89	0.421
TIxNCxPM	2,39	0.34	0.714	0.08	0.922	4.25	0.021
TIxCPxNCxPM	2,39	0.20	0.816	0.78	0.467	3.07	0.058

Tabla 1. Tabla de ANOVA resumida de los efectos del tipo de información para cada variable dependiente. TI=Tipo de información; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media; TT=Tiempo de Tarea; ND=Número de Datos; NA=Número de Aciertos.

F. Variac.	g.l	Dem. Mental		Dem. Temporal		Dem. Física		Indice Carga	
		F	Signif.	F	Signif.	F	Signif.	F	Signif.
TI	2,39	1.42	0.255	2.54	0.091	1.49	0.239	3.24	0.050
TIxCP	2,39	0.15	0.863	0.82	0.449	0.00	0.996	0.93	0.403
TIxNC	2,39	6.00	0.005	1.70	0.196	0.19	0.821	2.68	0.081
TIxPM	2,39	0.54	0.589	0.33	0.720	0.05	0.950	0.10	0.819
TIxCPxNC	2,39	0.51	0.603	1.29	0.288	0.39	0.662	0.94	0.398
TIxCPxPM	2,39	0.08	0.923	0.57	0.573	0.27	0.765	0.47	0.629
TIxNCxPM	2,39	0.50	0.611	2.59	0.088	0.07	0.932	1.01	0.373
TIxCPxNCxPM	2,39	0.54	0.584	0.71	0.496	2.02	0.146	1.19	0.314

Tabla 2. Tabla de ANOVA resumida de los efectos del tipo de información, para cada dimensión de la carga subjetiva y para el índice global de carga. TI=Tipo de Información; CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media.

F. Variac.	g.l	TT		ND		NA	
		F	Signif.	F	Signif.	F	Signif.
CP	1,13	5.00	0.043	4.97	0.044	0.48	0.500
NC	1,13	0.27	0.613	1.33	0.270	2.46	0.141
PM	1,13	0.14	0.719	2.12	0.169	12.13	0.004
CPxNC	1,13	21.38	0.000	1.00	0.337	0.09	0.775
CPxPM	1,13	9.27	0.009	1.19	0.295	0.32	0.583
NCxPM	1,13	1.80	0.203	0.13	0.720	9.16	0.010
CPxNCxPM	1,13	8.30	0.013	0.49	0.494	4.79	0.047

Tabla 3. Tabla de ANOVA resumida de los factores intrasujeto, para cada variable dependiente. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media; TT=Tiempo de Tarea; ND=Número de Datos; NA=Número de Aciertos.

F. Variac.	g.l	Dem. Mental		Dem. Temporal		Dem. Física		Índice Carga	
		F	Signif.	F	Signif.	F	Signif.	F	Signif.
CP	1,13	4.95	0.044	5.04	0.043	1.25	0.283	6.53	0.024
NC	1,13	11.20	0.005	6.62	0.023	0.05	0.825	8.89	0.011
PM	1,13	1.23	0.288	0.07	0.796	0.42	0.531	0.66	0.370
CPxNC	1,13	2.14	0.167	1.15	0.304	0.65	0.434	1.94	0.187
CPxPM	1,13	0.31	0.589	0.05	0.835	4.06	0.065	0.03	0.875
NCxPM	1,13	3.78	0.074	0.89	0.362	0.90	0.361	0.91	0.357
CPxNCxPM	1,13	0.06	0.814	0.26	0.620	2.37	0.148	0.00	0.946

Tabla 4. Tabla de ANOVA resumida de los factores intrasujeto, para cada dimensión de la carga subjetiva y para el índice global de carga. CP=Complejidad perceptiva; NC=Número de causas; PM=Probabilidad media.