

---

# FASE DE PONDERACIÓN DEL NASA-TLX: ¿UN PASO INNECESARIO EN LA APLICACIÓN DEL INSTRUMENTO?

# PHASE WEIGHTING OF NASA-TLX: AN UNNECESSARY STEP IN THE IMPLEMENTATION OF THE INSTRUMENT?

M<sup>a</sup> INMACULADA LÓPEZ NUÑEZ  
*Departamento de Psicología  
Facultad de Ciencias de la Salud  
Universidad Camilo José Cela*

SUSANA RUBIO VALDEHITA  
JESÚS MARTÍN GARCÍA  
LOURDES LUCEÑO MORENO  
*Laboratorio de Psicología  
del Trabajo y Estudios de Seguridad  
Universidad Complutense de Madrid*

e-mail: lluceno@psi.ucm.es

## RESUMEN

*El NASA-TLX es uno de los procedimientos subjetivos más ampliamente utilizado en la evaluación de la carga mental. La aplicación de este instrumento requiere una primera fase de ponderación de la importancia que cada dimensión de carga mental tiene para el sujeto, lo que parece ser un inconveniente ya que no sólo distingue un número excesivo de dimensiones, sino que además diversas investigaciones han puesto de manifiesto la dificultad de los sujetos para comprender esa fase, lo que dificulta su correcta realización. En este artículo, se pone de manifiesto el alto grado de correlación encontrado entre medidas ponderadas*

## ABSTRACT

*NASA-TLX is one of the subjective procedures more widely used in the evaluation of the mental workload. The application of this instrument needs the first phase of weighting of the importance that every dimension of mental load has for the subject, which seems to be a disadvantage since not only it distinguishes an excessive number of dimensions, but in addition diverse investigations have revealed the difficulty of the subjects to understand this phase, which impedes his correct accomplishment. In this article, there is revealed the high degree of correlation found between weighted measures and without weighting of mental load*

*y sin ponderar de carga mental obtenidas con el NASA-TLX, lo que sugiere que la fase de ponderación no es necesaria. Esto permitiría agilizar su aplicación, evitar la pérdida de casos y añadir una ventaja más al instrumento.*

#### **PALABRAS CLAVE**

*NASA-TLX, correlación, medidas ponderadas, carga mental*

*obtained with the NASA-TLX, which suggests that the phase of weighting is not necessary. This would allow to improve its application, to avoid the loss of cases and to add one more advantage to the instrument.*

#### **KEY WORDS**

*NASA-TLX, correlation, weighted measures, mental workload*

## **INTRODUCCIÓN**

Multitud de estudios (INSHT, 2004, 2007; Luceño, Rubio, Díaz y Martín, 2010; Rubio, Martín, y Díaz, 1995; Rubio, Díaz, Martín y Puente, 2004) ponen de manifiesto que el factor carga mental está creciendo como factor de riesgo psicosocial, lo que hace que, en la actualidad, la evaluación de la carga mental sea un aspecto central en la investigación de desarrollo de sistemas de trabajo que permitan obtener niveles más altos de confort, satisfacción, eficacia y seguridad y salud en el trabajo.

Existe un acuerdo general (Rubio, Luceño, Martín y Jaén 2007) en admitir que la carga mental es un concepto multidimensional y, por lo tanto, determinado por diferentes factores o dimensiones (Gopher y Donchin, 1986; O'Donnell y Eggemeier, 1986) y, 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: la primera englobaría todos los aspectos relativos a la presión temporal de la tarea, la segunda estaría formada por variables que hacen referencia a la cantidad de recursos de procesamiento que demanda la tarea y, la tercera, se relacionaría con aspectos de naturaleza más emocional (Eggemeier, Crabtree, Zingg, Reid, y Shingledecker, 1982; Hart y Staveland, 1988).

Se han propuesto múltiples métodos para la evaluación de la carga mental. Todos ellos pueden clasificarse dentro de alguno de los tres tipos siguientes: medidas del rendimiento, indicadores fisiológicos y técnicas subjetivas. Sin duda, los instrumentos subjetivos son los que han recibido una mayor atención, principalmente, por las ventajas que presentan para su aplicación en el contexto laboral. Los procedimientos subjetivos de evaluación de la carga mental destacan por poseer altos niveles de sensibilidad y validez, unos requisitos de implementación míni-

mos, buena aceptación por parte de los trabajadores y un grado de intrusión prácticamente nulo (Rubio y cols., 1999, 2004; Tsang y Wilson, 1997; Cañas y Waerns, 2001). Sin embargo, este tipo de procedimientos subjetivos no está exento de críticas. Entre ellas, las más importantes tienen que ver con su escaso poder predictivo del rendimiento y su limitado poder de diagnóstico (Eggemeier y Wilson, 1991; Moroney, Biers y Eggemeier, 1995; Tsang y Wilson, 1997).

En la actualidad los procedimientos subjetivo multidimensionales son los más utilizados para la evaluación de la carga mental de las tareas en contextos aplicados. Permiten distinguir diversas dimensiones o factores determinantes de la carga mental, todos ellos, obtienen una puntuación para cada dimensión de carga mental y, además, incluyen la forma en la que se deben combinar estas puntuaciones para obtener una puntuación global de la carga mental de una tarea (Arquer y Nogareda, 1999; Rubio y Díaz, 1999; Tsang y Wilson, 1997).

Uno de los procedimientos subjetivos multidimensionales más frecuentemente utilizado es el NASA-Task Load Index (NASA-TLX). Esta técnica fue derivada de la Escala de Valoración Bipolar de la NASA (*NASA Bipolar Rating Scale*) que contempla diez subescalas (Hart y Staveland, 1988). A través de un cuidadoso trabajo de investigación, los autores seleccionaron seis de estas diez subescalas en función de criterios tales como: la sensibilidad de cada una de ellas, la independencia estadística entre las mismas, y la importancia subjetiva respecto al concepto individual de carga de los sujetos (Lysaght, Hill, Dick, Plamondon, Linton, Wierwillw, Zakland, Bittner y Wherry., 1989).

Este procedimiento se basa en el presupuesto de que la carga mental es un constructo hipotético que representa el costo en el que incurre el operador al tratar de alcanzar un nivel específico de rendimiento. De esta forma el nivel de carga mental surgiría de la interacción entre los requerimientos de una tarea, las circunstancias bajo las que ésta es realizada y las habilidades, conductas y percepciones del operador (Hart y Staveland, 1988).

Las seis subescalas del instrumento pueden ser agrupadas dentro de tres bloques: subescalas relacionadas con las características de la tarea (demanda mental, demanda física y demanda temporal), subescalas relacionadas con las características conductuales (esfuerzo y rendimiento) y subescalas relacionadas con las características individuales (nivel de frustración). La descripción de cada una de estas dimensiones es la siguiente:

- Demanda mental: Cantidad de actividad mental y perceptiva que requiere la tarea (p. e.: pensar, decidir, calcular, recordar, mirar, buscar, etc.)
- Demanda física: Cantidad de actividad física que requiere la tarea (p.e.: pulsar, empujar, girar, etc.)

- Demanda temporal: Nivel de presión temporal sentida. Razón entre el tiempo requerido y el disponible
- Rendimiento: Hasta qué punto el individuo se siente satisfecho con su nivel de rendimiento
- Esfuerzo: Grado de esfuerzo mental y físico que tiene que realizar el sujeto para obtener su nivel de rendimiento
- Nivel de Frustración: Hasta qué punto el sujeto se siente inseguro, estresado, irritado, descontento, etc. durante la realización de la tarea

La aplicación de esta técnica requiere dos fases: una de obtención de la importancia inicial que tiene cada dimensión de carga mental para cada individuo y otra de valoración.

## FASE 1. OBTENCIÓN DE PESOS INICIALES

El objetivo de esta fase es la obtención de la importancia que, a priori, da cada individuo a cada una de las seis dimensiones como fuente potencial de carga mental. Esta fase es previa a la realización de la tarea y permite obtener los valores por los que se van a ponderar las estimaciones de carga al calcular el índice global de carga mental de una determinada tarea o combinación de tareas.

Para la recogida de los datos necesarios se utiliza el procedimiento de comparaciones binarias, de manera que: en primer lugar se establecen las 15 comparaciones binarias de las seis dimensiones, en las que el sujeto debe elegir, de cada par, la que percibe como mayor fuente de carga.

Para cada dimensión se obtiene un peso que viene dado por el número de veces que ésta haya sido seleccionada en las comparaciones binarias. Este peso puede variar entre 0 (la dimensión no ha sido elegida en ninguna de las comparaciones) y 5 (la dimensión ha sido elegida en todas las comparaciones en las que aparecía).

## FASE 2. VALORACIÓN

Una vez realizada la tarea o tareas de interés, el sujeto tiene que estimar, en una escala de 0 a 100, dividida en intervalos de 5 unidades, la carga mental de la tarea debida a cada una de las seis dimensiones. Con los datos obtenidos en las dos fases se puede calcular un índice global de la carga mental de la tarea aplicando la siguiente fórmula:

6

$$IC = \sum_{i=1}^{15} piXi / 15$$

i-1

En la que:

*IC* es Índice de Carga, *pi* es el peso obtenido para cada dimensión en la fase de ponderación, y *Xi* es la puntuación obtenida por la dimensión en la fase de valoración.

El desarrollo de la técnica TLX ha supuesto un programa muy extenso de investigación en laboratorio (Hart y Staveland, 1988), mediante el cual ha quedado demostrada la sensibilidad de este instrumento con una gran variedad de tareas (Noyes y Bruneau, 2007). La técnica TLX ha sido utilizada en un gran número de investigaciones. Por ejemplo, Sawin y Scerbo (1995) emplearon esta técnica para estudiar los efectos del tipo de instrucciones y de la tendencia al aburrimiento en tareas de vigilancia. Alm y Nilsson (1995) examinaron el efecto que sobre el rendimiento y la carga mental tiene la utilización del teléfono móvil durante la conducción. Otro ejemplo de utilización de esta técnica es el estudio realizado por Dickinson, Winston y Ryan (1993) en el que se analizaba, entre otras cosas, el efecto que tiene el proceso de ponderación inicial sobre el índice global de carga mental.

Además NASA-TLX ha sido aplicado con éxito en diversos contextos multitarea (Xie y Salvendy, 2000) como, por ejemplo, en tareas de vuelo reales (Shively, Battiste, Matsumoto, Pepiton, Bortolossi, y Hart., 1987) y con simulador (Battiste y Bortolussi, 1988; Nataupsky y Abbott, 1987; Tsang y Johnson, 1989; Vidulich y Bortolussi, 1988), en situaciones de defensa aérea (Bittner, Byers, Hill, Zakland y Christl, 1989; Hill, Buers, Zakland y Christ, 1989), en la conducción de vehículos de control remoto (Byers, Bittner, Hill, Zakland, y Christ, 1988) o en la percepción de la calidad del servicio de páginas web de inserción laboral (Tong, Duffy, Cross, Tsung y Yen, 2003).

La validez de la técnica ha sido probada en numerosas investigaciones, demostrando resultados más favorables que otras técnicas como el SWAT o la Escala de Cooper-Harper modificada (Hancock y Warm, 1989, Schick, Tejen, Uckermann y Hann, 1989; Hill, Iavecchia, Byers, Bittner, Zackland y Christ, 1992), lo que unido a su facilidad de uso hace que sea actualmente el instrumento de evaluación de carga mental más ampliamente utilizado (Cooke, Gorman, Duran y Taylor, 2007; Morrow, Raquel, Schriver, Redenbo, Rozovski y Weiss, 2008; Recarte, Pérez, Conchillo y Nunes, 2008; Rutledge, Stucky, Dollarhide, Shively, Jain y Wolfson, 2009).

Sin embargo, uno de los inconvenientes que parece tener la técnica NASA-TLX es que distingue un número excesivo de dimensiones de carga mental y que

la fase de ponderación es innecesaria. Nygren (1991) opina que bastaría con distinguir un número menor de dimensiones, por ejemplo las tres que considera SWAT (Rubio, 1992) y, que una simple media aritmética sin ponderar de las valoraciones obtenidas para cada dimensión es un índice global de la carga mental de una tarea tan válido y sensible como el IC que propone calcular el TLX.

Por ello, en este artículo se presentan los resultados de una investigación que tenía entre sus objetivos comprobar las correlaciones entre medidas de carga mental ponderadas y sin ponderar obtenidas a través del instrumento NASA-TLX mediante su aplicación en dos situaciones (real y de laboratorio) y utilizando dos tipos de muestras (estudiantes y trabajadores) dentro de un estudio más amplio de generalizabilidad del instrumento (López, 2009).

## MÉTODO

### *Muestra*

La muestra estuvo formada por 364 participantes, todos ellos diestros, de los cuales 197 eran mujeres y 167 hombres. La media de edad fue de 29,06 años (desviación típica = 9,99). Del total de la muestra, 168 sujetos eran trabajadores y 196 estudiantes de Psicología de la Universidad Complutense de Madrid. El grupo de trabajadores pertenecía a 6 contextos laborales diferentes: administrativos, consultores, periodistas, policías municipales, políticos, militares y profesores. La razón por la que se seleccionaron tanto estudiantes como trabajadores es poder estudiar si se producen resultados diferentes en ambos tipos de muestras.

La distribución por puestos de trabajo fue la siguiente: 14 administrativos, 27 consultores, 37 políticos, 29 militares, 13 periodistas, 48 policías municipales, y 196 estudiantes.

Por otro lado, solamente el 8% de la muestra había sufrido baja médica en el último año, siendo el 20 % de ellas debidas a estrés y ansiedad.

El 81 % tenía estudios universitarios, el 16% había terminado la enseñanza secundaria obligatoria y sólo el 0,3% de la muestra total tenía estudios primarios.

Todos los sujetos participaron de forma voluntaria y anónima en esta investigación.

### *Diseño*

Se siguió un diseño factorial mixto con dos factores. Uno de ellos, la situación de evaluación, fue intrasujetos o de medidas repetidas. El factor entresujetos fue el tipo de participante (trabajador o estudiante). Mediante el NASA-TLX, todos

los sujetos debían estimar la carga mental de cada una de las tareas propias de su puesto de trabajo (para el grupo de trabajadores) o de su actividad académica (para el grupo de estudiantes), así como la de una serie de tareas experimentales que eran las mismas para ambos grupos. Se establecieron así las dos situaciones de evaluación: real (tareas laborales o académicas) y en laboratorio (tareas experimentales).

## *Tareas e Instrumentos*

### *Tareas experimentales*

Se utilizaron dos tipos de tareas experimentales: una tarea de búsqueda en la memoria y una tarea de seguimiento manual o tracking. Ambas tareas fueron aplicadas mediante un programa de ordenador y fueron realizadas tanto por separado (versiones simples) como conjuntamente (versiones duales). La respuesta de ambas tareas era manual (apretando las teclas correspondientes del teclado del ordenador, para la tarea de memoria, y manejando el ratón para la tarea de seguimiento). La tarea de memoria se realizó siempre respondiendo con la mano izquierda, mientras que la de seguimiento se realizó con la derecha. En las versiones duales, en las que el sujeto debía realizar las dos tareas al mismo tiempo, se instruyó a los sujetos para que atendiesen por igual a ambas tareas. A continuación se describe de forma más detallada cada una de estas tareas.

- *Tarea de búsqueda en la memoria.* Los sujetos debían memorizar un conjunto de letras consonantes al comienzo de cada ensayo. Se establecieron dos niveles de dificultad objetiva según el número de letras a memorizar: la versión más fácil requería memorizar 2 letras (m2) y la más difícil 4 (m4). La tarea de los sujetos consistía en responder si la letra que se le presentaba coincidía o no con una de las que había memorizado previamente. Todas las letras presentadas al sujeto, tanto las que debía memorizar como la que aparecía en cada ensayo, eran elegidas de forma aleatoria por el programa de ordenador. Los sujetos respondían con la mano izquierda presionando la tecla F1 en caso de que su respuesta fuera afirmativa y F2 en el caso contrario. Los sujetos debían realizar esta tarea de la forma más rápida y precisa.
- *Tarea de seguimiento.* Esta tarea consistía en mantener el cursor dentro de un camino móvil de referencia, utilizando el ratón. Como parámetro de dificultad objetiva se estableció el ancho de la vía: siendo el nivel más estrecho el de mayor dificultad (w2) y el nivel más ancho el

de menor dificultad (w1). Los sujetos realizaron esta tarea con la mano derecha.

- *Tareas duales.* A partir de la combinación de las tareas anteriores, se emplearon cuatro dos duales (m2w1, m2w2, m4w1 y m4w2). La tarea de seguimiento aparecía en el lado derecho de la pantalla y la tarea de memoria en el lado izquierdo, lo suficientemente cercanas para minimizar el coste de concurrencia. Los sujetos recibieron la instrucción de prestar la misma atención a las dos tareas y hacer lo mejor posible ambas. Para asegurar que los sujetos atendían a las dos tareas, el programa incluía un mensaje de aviso visual que aparecía en la pantalla cuando la respuesta se demoraba más de 10 segundos en la tarea de memoria o cuando el cursor se desviaba del camino de referencia en la tarea de seguimiento.

### *Tareas reales*

Las diferentes tareas reales fueron obtenidas a partir del análisis de puestos realizado mediante entrevistas con los trabajadores/estudiantes y sus jefes inmediatos/profesores. Todos los sujetos evaluaron mediante el NASA, la carga mental de sus puestos de trabajo, en el caso de la muestra de trabajadores, o de su actividad académica, en el caso de los estudiantes.

### *Aparatos*

Para la presentación y recogida de los datos relativos a las tareas experimentales, se utilizó un ordenador portátil Toshiba AMD Turion 64x2 1.60GHz, 2046 MB RAM. Monitor de 15 pulgadas.

### *Procedimiento*

La recogida de los datos necesarios para esta investigación se realizó en tres sesiones. En la primera de ellas se realizaron las entrevistas necesarias para confeccionar el análisis de los puestos.

En las dos sesiones siguientes se recogieron las estimaciones de carga mental tanto del puesto de trabajo como de las tareas experimentales. El orden de recogida entre ambas situaciones se contrabalanceó de manera que se eliminaran los posibles efectos del orden. En este sentido, la mitad de la muestra de estudiantes y de trabajadores estimó primero la carga mental de su actividad real, mientras que la otra mitad de la muestra realizó antes la sesión en laboratorio.

La sesión de laboratorio se desarrolló de la siguiente forma. Al comienzo, el investigador explicaba a los sujetos el funcionamiento del programa y las tareas que iban a tener que realizar, aunque todos los sujetos disponían de estas instrucciones en la pantalla del ordenador. Si en algún momento durante la ejecución de las tareas el sujeto tenía alguna duda podía preguntar al investigador para su resolución. En primer lugar, el sujeto debía completar la fase de comparación de pares que incluye en instrumento NASA-TLX. A continuación, aparecían las dos versiones simples de la tarea de memoria (el sujeto tenía que memorizar 2 o 4 letras) y a continuación las dos versiones simples de la tarea de tracking (el sujeto debía mantener el cursor dentro de un camino ancho o de un camino más estrecho). Seguidamente se presentaban las dos tareas duales resultantes de la combinación de las dos tareas simples más fáciles y las dos más difíciles (m2w1 y m4w2). Inmediatamente después de la realización de cada una de estas tareas, el sujeto debía evaluar la carga mental de las mismas valorando en una escala de 0 a 100 cada una de las dimensiones de carga que distingue el instrumento NASA-TLX. Las cuatro tareas simples fueron consideradas como de entrenamiento y familiarización con el procedimiento. La muestra de estudiantes realizó esta sesión en una cabina del Laboratorio de la Facultad de Psicología de la Universidad Complutense de Madrid. La muestra de trabajadores realizó las tareas experimentales en un despacho situado en su centro de trabajo en el que se mantuvieron las mismas condiciones que en un laboratorio (ajeno a interrupciones, iluminación adecuada, sin ruidos externos, etc.). El orden de presentación de las tareas duales se contrabalanceo, de manera que la mitad de los estudiantes y de los trabajadores realizaron en primer lugar la tarea fácil (m2w1), mientras que el resto de participantes realizó primero la tarea difícil (m4w2).

La sesión en la que los sujetos evaluaban la carga mental de sus tareas laborales o académicas se llevó a cabo mediante la aplicación de una hoja de recogida de datos en la que aparecía cada una de las tareas correspondientes y las escalas de valoración para cada una de las dimensiones de carga. En esta hoja se incluyeron las definiciones de cada una de estas dimensiones, de manera que fueran accesibles al sujeto en todo momento.

## RESULTADOS

Todos los análisis estadísticos fueron realizados con el SPSS 15.0. A continuación se presentan los resultados obtenidos según la situación de evaluación, para la muestra total y para cada uno de los grupos. En todos los casos, se calcu-

laron los coeficientes de correlación de Pearson entre las evaluaciones de carga mental ponderadas (aplicando la fórmula del NASA-TLX) y las medias aritméticas simples (sin ponderar) de las seis dimensiones de carga, es decir, sin tener en cuenta la fase de ponderación.

### *Tareas experimentales*

En la tabla 1 aparecen las correlaciones entre todas las tareas experimentales para la muestra total. En todos los casos, se obtuvieron índices de correlación cercanos a la unidad.

Tabla 1. *Matriz de correlaciones de las valoraciones subjetivas de carga mental ponderada y sin ponderar de las tareas experimentales para la muestra total.*

	CARGA MEDIA TAREA m2	CARGA MEDIA TAREA m4	CARGA MEDIA TAREA w1	CARGA MEDIA TAREA w2	CARGA MEDIA TAREA m2w1	CARGA MEDIA TAREA m2w2	CARGA MEDIA TAREA m4w1	CARGA MEDIA TAREA m4w2
Carga ponderada tarea m2	,930(**)							
Carga ponderada tarea m4		,934(**)						
Carga ponderada tarea w1			,938(**)					
Carga ponderada tarea w2				,960(**)				
Carga ponderada tarea m2w1					,964(**)			
Carga ponderada tarea m2w2						,965(**)		
Carga ponderada tarea m4w1							,961(**)	
Carga ponderada tarea m4w2								,959(**)

(\*\*) La correlación es significativa al nivel 0,01 (bilateral).

En la tabla 2 aparecen las correlaciones entre todas las tareas experimentales para la muestra de trabajadores y en la tabla 3 para la de estudiantes. En todos los casos, se obtuvieron índices de correlación muy elevados y similares entre ambos tipos de muestras.

Tabla 2. *Matriz de correlaciones de las valoraciones subjetivas de carga mental ponderada y sin ponderar de las tareas experimentales para el grupo de trabajadores.*

	CARGA MEDIA TAREA m2	CARGA MEDIA TAREA m4	CARGA MEDIA TAREA w1	CARGA MEDIA TAREA w2	CARGA MEDIA TAREA m2w1	CARGA MEDIA TAREA m2w2	CARGA MEDIA TAREA m4w1	CARGA MEDIA TAREA m4w2
Carga ponderada tarea m2	,927(**)							
Carga ponderada tarea m4		,926(**)						
Carga ponderada tarea w1			,949(**)					
Carga ponderada tarea w2				,959(**)				
Carga ponderada tarea m2w1					,962(**)			
Carga ponderada tarea m2w2						,964(**)		
Carga ponderada tarea m4w1							,960(**)	
Carga ponderada tarea m4w2								,957(**)

(\*\*) La correlación es significativa al nivel 0,01 (bilateral).

Tabla 3. Matriz de correlaciones de las valoraciones subjetivas de carga mental ponderada y sin ponderar de las tareas experimentales para el grupo de estudiantes.

	CARGA MEDIA TAREA m2	CARGA MEDIA TAREA m4	CARGA MEDIA TAREA w1	CARGA MEDIA TAREA w2	CARGA MEDIA TAREA m2w1	CARGA MEDIA TAREA m2w2	CARGA MEDIA TAREA m4w1	CARGA MEDIA TAREA m4w2
Carga ponderada tarea m2	,934(**)							
Carga ponderada tarea 2		,945(**)						
Carga ponderada tarea w1			,922(**)					
Carga ponderada tarea w2				,961(**)				
Carga ponderada tarea m2w1					,965(**)			
Carga ponderada tarea m2w2						,967(**)		
Carga ponderada tarea m4w1							,962(**)	
Carga ponderada tarea m4w2								,963(**)

(\*\*) La correlación es significativa al nivel 0,01 (bilateral).

### Tareas reales

A continuación se calcularon las correlaciones entre todas las valoraciones de carga mental ponderadas y sin ponderar de los puestos de trabajo y las tareas académicas, para la muestra total y para ambos grupos de participantes (trabajadores y estudiantes). En todos los casos, se obtuvieron índices de correlación elevados y estadísticamente significativos. Los índices de correlación obtenidos fueron los siguientes: 0,865,  $p < ,000$ , para la muestra total; 0,855,  $p < ,000$ , para la muestra de trabajadores; y 0,876,  $p < ,000$ , para el grupo de estudiantes.

## DISCUSIÓN Y CONCLUSIONES

Uno de los procedimientos subjetivos multidimensionales más frecuentemente utilizado para evaluar la carga mental es el NASA-Task Load Index (NASA-TLX). El desarrollo de esta técnica ha supuesto un programa muy extenso de investigación en laboratorio (Hart y Staveland, 1988), mediante el cual ha quedado demostrada la sensibilidad de este instrumento con una gran variedad de tareas (Noyes y Bruneau, 2007). La técnica TLX ha sido utilizada en un gran número de investigaciones (Alm y Nilsson, 1995; Bittner et al., 1989; Hill et al., 1989; Nataupsky y Abbott, 1987; Sawin y Scerbo, 1995; Tong et al., 2003; Tsang y Johnson, 1989; Vidulich y Bortolussi, 1988; Xie y Salvendy, 2000).

Sin embargo, uno de los inconvenientes que parece que tiene la técnica NASA-TLX, y que se apuntaba al principio de este artículo, es que parece que distingue un número excesivo de dimensiones de carga mental y que la fase de ponderación es innecesaria (Dickinson, Winston y Ryan, 1993; Nygren, 1991). Por ello, el objetivo de esta investigación era comprobar el grado de correlación entre medidas de carga mental ponderadas y sin ponderar obtenidas a través del instrumento NASA-TLX mediante su aplicación en dos situaciones (real y de laboratorio) y utilizando dos tipos de muestras (estudiantes y trabajadores).

Los datos ponen de manifiesto el alto grado de correlación encontrado entre ambas medidas, tanto para las tareas experimentales como reales del puesto de trabajo y para los dos tipos de muestras utilizados (estudiantes y trabajadores), además, el hecho de que algunos sujetos tuvieran dificultad, en su comprensión, en realizar la fase de comparación de pares, nos sirvieron para utilizar las medidas sin ponderar en el análisis de los datos para otros estudios sobre generalizabilidad que hemos llevado a cabo.

Estos resultados parecen confirmar, de acuerdo a la bibliografía existente, que la fase de ponderación del NASA-TLX no es necesaria (Dickinson, Winston y Ryan, 1993; DiDomenico y Nussbaum, 2008; Hendy, Liao y Milgram, 1997; Moroney, Biers y Eggemeier, 1995, Nygren, 1991; ). Moroney y cols. (1995) señalaron un nivel de correlación de 0,94 entre los valores ponderados y no ponderados del NASA-TLX, Byers, Bittner y Hill, (1989) aumentaron este nivel hasta 0,98. En este estudio se han encontrado correlaciones de 0,87 para tareas experimentales, y de 0,85 para las tareas del puesto de trabajo, donde el rango variaba desde 0,68 para las tareas del grupo de periodistas y 1 (correlación perfecta) para las tareas del grupo de militares JES.

Las altas correlaciones encontradas y el hecho de que muchos sujetos puedan tener dificultades en la comprensión de la fase de comparación de pares

(Bohórquez, 2009) nos llevan a pensar que la fase de ponderación podría no ser necesaria en la aplicación del instrumento, lo que permitiría agilizar su aplicación, hacerlo menos tedioso para el sujeto y, evitaría la pérdida de casos. En la actualidad, continuamos ampliando las muestras para comprobar si dichos resultados se confirman y se mantienen las altas correlaciones entre medidas ponderadas y sin ponderar obtenidas a través del NASA-TLX, lo que unido a su facilidad de uso y los buenos resultados que, parece se van obteniendo en torno a su uso en entornos laborales, harían de este instrumento una excelente herramienta para la evaluación y modificación de las fuentes de carga mental.

## REFERENCIAS BIBLIOGRÁFICAS

- Alm, H. y Nilsson, L. (1995). The effects of a mobile telephone task on driver behavior in a car following situation. *Accident analysis and prevention*, 27, 707-715.
- Arquer, I., y Nogareda, C. (1999). *Estimación de la carga mental de trabajo: el método NASA TLX*. NTP 544. Madrid: INSHT.
- Battiste, V. y Bortolussi, M. (1988). Transport pilot workload: A comparison of two subjective techniques. En *Proceedings of the Human Factors Society Thirty-Second Annual Meeting*, 150-154, Santa Monica, CA: Human Factors Society.
- Bittner, A. V., Byers, J. C., Hill, S. G., Zaklad, A. L. y Christ, R.E. (1989). Generic workload ratings of a mobile air defense system (LOS-FH). En *Proceedings of the Human Factors Society Thirty-Third Annual Meeting*, 1.476-1.480, Santa Monica, CA: Human Factors Society.
- Bohórquez, M. R., Garrido, M. A., Bueno, M. R., Lorenzo, M. (2009). Aproximación a la Evaluación de la Carga Mental en Jugadores Semiprofesionales de Fútbol. *Efdeportes*, 139.
- Byers, J.C., Bittner, A.C., Hill, S.G. (1989). Traditional and raw Task Load Index (TLX) correlations: Are paired comparisons necessary? En A. Mital (ed.) *Advances in industrial ergonomics and safety*, 1, 481-485.
- Byers, J.C., Bittner, A.C., Hill, S.G., Zakland, A.L. y Christ, R.E. (1988). Workload assessment of a remotely piloted vehicle (RPV) system. En *Proceedings of the Human Factors Society Thirty-Second Annual Meeting*, 1.145-1.149, Santa Monica, CA: Human Factors Society.
- Cañas, J. J. y Waerns, Y. (2001). *Ergonomía cognitiva. Aspectos psicológicos de la interacción de las personas con la tecnología de la información*. Madrid: Editorial Médica Panamericana.

- Cooke, N. J., Gorman, J. C., Duran, J. L., & Taylor, A. R. (2007). Team cognition in experienced command-and-control teams. *Journal of Experimental Psychology: Applied*, 13(3), 146-157.
- Dickinson, J.B., Winston, D. y Ryan, L.A. (1993). Order effects and the weighting process in workload assessment. *Applied Ergonomics*, 24, 357-361.
- EGGEMIER, F. Y WILSON, G. (1991). Performance-based and subjective assessment of workload. En D.L. Damos (ed.), *Multiple-Task Performance*. Londres: Taylor & Francis.
- 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 of the human factors society 26<sup>th</sup> annual meeting*, 643-647. Santa Monica, C.A.: Human factors and ergonomics society.
- Gopher, D. y 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.
- Hancock, P.A. y Warm, J. S. (1989). A dynamic model of stress and sustained attention. *Human Factors*, 31, 519-537.
- 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.
- Hendy, K., Liao, J. y Milgram, P. (1997). Combining time and intensity effects in assessing operator information processing load. *Human Factors*, 39, 30-47.
- Hill, S.G., Byers, J. C., Zaklad, A. L. y Chirst, R. E. (1989). Subjective workload assessment during 48 continuous hours of LOS-F-H operations. En *Proceedings of the Human Factors Society Thirty-Third Annual Meeting*, 1.129-1.133, Santa Monica, CA: Human Factors Society.
- Hill, S.G., Iavecchia, H.P., Byers, J.C., Bittner, A.C., Zackland, A.L. y Christ, R.E. (1992). Comparison of four subjective workload rating scales. *Human Factors*, 34, 429-439.
- Instituto Nacional de Seguridad e Higiene en el Trabajo (2004): *V Encuesta Nacional de Condiciones de Trabajo*. Madrid: INSHT.
- Instituto Nacional de Seguridad e Higiene en el Trabajo (2007): *VI Encuesta Nacional de Condiciones de Trabajo*. Madrid: INSHT.
- López Núñez, M. I. (2009). Generalización al ámbito laboral de dos instrumentos de medida subjetiva de la carga mental, Tesis doctoral (UCM) (en prensa).
- Luceño, L. Rubio, S. Díaz, E y Martín, J. (2010). Percepción de riesgos psicosociales en distintas muestras de trabajadores españoles. *EduPsykhé*, 9, 1, 119-141.

- Lysaght, R. J.; Hill, S. G.; Dick, A. O.; Plamondon, B. D.; Linton, P. M.; Wierwille, W. W.; Zaklad, A. L.; Bittner, A. C. y Wherry, R. J. (1989). *Operator workload: Comprehensive review and evaluation of operator workload methodologies* (Technical Report 851). Fort bliss, Tx: U.S. Army Research Institute, Field Unit.
- Moroney, W. F., Biers, D. W. y Eggemeier, F. T. (1995). Some measurement and methodological considerations in the application of subjective workload measurement techniques. *The international journal of aviation psychology*, 5, 87-106.
- Morrow, D., Raquel, L., Schriver, A., Redenbo, S., Rozovski, D. y Weiss, G. (2008). External support for collaborative problem solving in a simulated provider/patient medication scheduling task. *Journal of Experimental Psychology: Applied*, 14(3), 288-297.
- Nataupsky, M. y Abbott, T.S. (1987). Comparison of workload measures on computer-generated primary flight displays. En *Proceedings of the Human Factors Society Thirty-First Annual Meeting*, 548-552, Santa Monica, CA: Human Factors Society.
- Noyes, J. M. y Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics*, 50(4), 514-519.
- Nygren, T. E. (1991). Psychometric properties of subjective workload techniques: implications for their use in the assessment of perceived mental workload. *Human Factors*, 33, 1, 17-33.
- O'donell, 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, 149. Nueva York: Wiley.
- Recarte, M.A., Pérez, E., Conchillo, A., Nunes, L. M. (2008). Mental workload and visual impairment: differences between Pupil, Blink, and Subjective Rating. *The Spanish Journal of Psychology*, 11, 374-385.
- Rubio, S. (1992). *Evaluación y medida de la carga mental en una tarea de diagnóstico de fallos*. Madrid: UCM.
- Rubio, S. y Díaz, E. (1999). La medida de la carga mental de trabajo I: índices basados en el rendimiento. *Boletín Digital de Factores Humanos* (20).
- Rubio, S. y Díaz, E. (1999). La medida de la carga mental de trabajo II: Procedimientos subjetivos. *Boletín Digital de Factores Humanos* (21).
- Rubio, S., Martín, J. y Díaz, E. (1995). Carga mental y estrés: dos conceptos relacionados. *Ansiedad y Estrés*, 1 (2-3), 131-139.

- Rubio, S., Díaz, E. M.; Martín, J. y Puente, J.M. (1999). Comparing the psychometric properties of three subjective workload assessment techniques. En, P. Mondelo, M. Mattila y W. Karwowski (Eds.), *Proceedings of the International Conference on Computer-Aided Ergonomics and Safety*. Barcelona: UPC.
- Rubio, S., Díaz, E., Martín, J. y Puente, J.M. (2004). Evaluation of subjective mental workload: a comparison of SWAT, NASA-TLX and Workload Profile methods. *Applied Psychology: An International Review*, 53, 61-86.
- Rubio, S., Luceño, L., Martín, J. y Jaén, M. (2007). Modelos y procedimientos de evaluación de la carga mental de trabajo. *Edupsykhé*, 6, 1, 85-109.
- Rutledge, T., Stucky, E., Dollarhide, A., Shively, M., Jain, S., Wolfson, T., et al (2009). A real-time assessment of work stress in physicians and nurses. *Health Psychology*, 28(2), 194-200.
- Sawin, S.A. y Scerbo, M.W. (1995). Effects of instruction type and boredom proneness in vigilance: implications for boredom and workload. *Human Factors*, 37, 752-765.
- Schick, F.V., Tejen, U., Uckermann, R. y Hann, R.L. (1989). Validation of the Subjective Workload Assessment Technique in a simulated flight task. *DFVLR-Forschungsbericht*, 89-101.
- Shively, R., Battiste, V., Matsumoto, J., Pepiton, D., Bortolossi, M. y Hart, S.G. (1987). In flight evaluation of pilot workload measures for rotorcraft research. En *Proceeding of the Fourth Symposium on Aviation Psychology*, 637-643, Columbus, OH: Department of Aviation, The Ohio State University.
- Tong, J.P.C., Duffy, V.G., Cross, G.W., Tsung, F. y Yen, B. (2003). Evaluating the service quality of online recruitment websites: Comparing perceived overall service quality to measures of mental workload and performance time. *Human factors and ergonomics manufacturing*, 3-21.
- Tsang, G. y Wilson, G.F. (1997). Mental workload. En G. Salvendy (ed.), *Handbook of Human Factors and Ergonomics*. Nueva York: Wiley & Sons.
- Tsang, P.S. y Johnson, W.W. (1989). Cognitive demand in automation. *Aviation, Space, and Environmental Medicine*, 60, 130-135.
- Vidulich, M. A. y Bortolussi, M. R. (1988). A dissociation of objective and subjective workload measures in assessing the impact of speech controls in advanced helicopters. En *Proceedings of the Human Factors Society Thirty-Second Annual Meeting*, 1.471-1.475, Santa Monica, CA: Human Factors Society.
- Xie, B. y Salvendy, G. (2000). Prediction of mental workload in single and multiple tasks environments. *International journal of cognitive ergonomics*, 4(3), 213-242.