



Revista de Psicología del Trabajo y de las Organizaciones  
ISSN: 1576-5962  
revistas\_copm@cop.es  
Colegio Oficial de Psicólogos de Madrid  
España

Díaz Ramiro, Eva; Rubio Valdehita, Susana; Martín García, Jesús; Luceño Moreno, Lourdes  
Estudio Psicométrico del Índice de Carga Mental NASA-TLX con una Muestra de Trabajadores  
Españos  
Revista de Psicología del Trabajo y de las Organizaciones, vol. 26, núm. 3, diciembre, 2010, pp. 191-  
199  
Colegio Oficial de Psicólogos de Madrid  
Madrid, España

Disponible en: <http://www.redalyc.org/articulo.oa?id=231316502003>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org

# Estudio Psicométrico del Índice de Carga Mental NASA-TLX con una Muestra de Trabajadores Españoles

## Psychometric Study of NASA-TLX Mental Workload Index in a Sample of Spanish Workers

Eva Díaz Ramiro, Susana Rubio Valdehita, Jesús Martín García y Lourdes Luceño Moreno  
Universidad Complutense de Madrid

**Resumen.** El objetivo de este estudio fue evaluar las propiedades psicométricas del instrumento de evaluación de la carga mental de trabajo NASA-TLX en distintos grupos profesionales españoles. La muestra estuvo formada por 398 trabajadores que pertenecían a siete sectores profesionales diferentes. Todos los trabajadores evaluaron la carga mental percibida en sus puestos de trabajo, debida a cada una de las seis dimensiones de carga que distingue el NASA-TLX: esfuerzo, demanda mental, demanda física, demanda temporal, rendimiento y frustración. Los resultados mostraron una consistencia interna aceptable y una estructura factorial constituida por dos factores. Uno de los factores estaba formado únicamente por la dimensión "frustración", mientras que el otro factor englobaba a las cinco dimensiones restantes. Por último, se encontraron diferencias importantes en los perfiles de carga mental entre los grupos profesionales considerados, en todas las dimensiones de carga, excepto en rendimiento.

*Palabras clave:* carga mental, NASA-TLX, validez, fiabilidad, trabajadores.

**Abstract.** The objective of this study was to assess the psychometric properties of NASA-Task Load Index using several groups of Spanish workers. The sample consisted of 398 workers who belonged to seven different jobs. All workers assessed the perceived mental workload when perform their jobs, due to each of the six dimensions of mental workload that distinguishes the NASA-TLX: effort, mental demand, physical demand, temporal demand, performance and frustration. The results showed that NASA-Task Load Index had a acceptable internal consistency and a factorial structure formed by two components. The first of the two components was formed by all of the NASA-TLX dimensions except frustration. The second was formed only by frustration dimension. Significant differences in the profiles of mental workload experienced by workers on the basis of their jobs were found for all NASA-TLX dimensions except in performance.

*Keywords:* mental workload, NASA-TLX, validity, reliability, workers.

Los resultados de la VI Encuesta Nacional de Condiciones de Trabajo (INSHT, 2007) indican que en España, al igual que en el resto de países de nuestro entorno, la carga de trabajo mental está creciendo como uno de los más importantes factores de riesgo psicosocial. Históricamente en la definición de carga mental se han seguido dos perspectivas (Ferrer y Dalmau, 2004; Hacker, 1998): la primera, entendida como un factor dependiente únicamente de las exigencias de la tarea, a las que tiene que enfrentarse y adaptarse la persona, y la segunda como resultante de la interacción entre las exigencias de la tarea y las capacidades o recursos de la persona. Este segundo enfoque es el mayoritariamente aceptado y el que, actualmente goza de mayor apoyo (Young y Staton, 2001).

Los resultados presentados en este artículo forman parte de una investigación más amplia financiada por el Instituto de Seguridad e Higiene en el Trabajo (INSHT) (Referencia UCM/CARMEN). La correspondencia deberá ser enviada a la primera autora a la Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas, Buzón 218, Pozuelo de Alarcón, Madrid 28023. E-mail: [ediazram@psi.ucm.es](mailto:ediazram@psi.ucm.es)

En general, se admite que la carga mental es multidimensional y que por lo tanto está determinada por diferentes aspectos (O'Donnell y Eggemeier, 1986). En este sentido, la carga, fundamentalmente la subjetiva, se debe a tres grandes áreas. La primera englobaría los aspectos relativos a la presión temporal de la tarea (tiempo disponible, tiempo necesario). La segunda estaría formada por variables relacionadas con la cantidad de recursos de procesamiento que demanda la tarea y, por último, la tercera dimensión se relacionaría con aspectos de naturaleza más emocional: fatiga, frustración, estrés (Hart y Staveland, 1988).

Por otro lado, entre todos los procedimientos propuestos para evaluar la carga mental de trabajo (técnicas subjetivas, indicadores fisiológicos y medidas del rendimiento en la tarea), son los subjetivos los que destacan por poseer elevados niveles de sensibilidad, requisitos de implementación mínimos, buena aceptación por los trabajadores, un grado de intrusión prácticamente nulo y niveles adecuados de validez. Estos procedimientos reflejan la opinión directa del sujeto acerca del nivel de carga mental experimentada duran-

te la realización de una tarea. En la práctica, los procedimientos subjetivos multidimensionales son los más utilizados, principalmente, porque permiten diferenciar entre las distintas causas de carga mental, manteniendo las características psicométricas anteriormente mencionadas (Tsang y Wilson., 1997). Todos ellos permiten obtener una puntuación de carga para cada una de las dimensiones que distinguen y una puntuación global de la carga mental de una tarea (O'Donnell y Eggemeier, 1986; Tsang y Wilson, 1997; Arquer y Nogareda, 1999). Entre los procedimientos subjetivos multidimensionales más frecuentemente utilizados y desarrollados, destaca el NASA-Task Load Index (TLX, Hart y Staveland, 1988). Su validez, comprobada en numerosas investigaciones, y su facilidad de uso hacen que sea el instrumento de evaluación de carga mental más ampliamente utilizado (Charlton, 2002; Marras y Karwowski, 2006; Rutledge, et al., 2009).

El desarrollo de la técnica TLX ha supuesto un programa muy extenso de investigación (Hart y Staveland, 1988), mediante el cual ha quedado demostrada su sensibilidad empleando una gran variedad de tareas (Noyes y Bruneau, 2007; Rubio, Díaz, Martín y Puente, 2004), principalmente en condiciones de laboratorio o de simulación. Otros estudios han analizado su validez convergente mediante su correlación con otros procedimientos para evaluar la carga mental. Por ejemplo, Recarte, Pérez, Conchillo y Nunes (2008) compararon los resultados obtenidos mediante el NASA-TLX con medidas fisiológicas como la tasa de parpadeo y el tamaño de la pupila, durante la conducción de un vehículo y bajo diferentes condiciones de tarea. Estos autores concluyeron la existencia de validez convergente entre los tres indicadores de carga mental, especialmente bajo condiciones de tarea simple. Sin embargo, Dalmau (2007) evaluó la carga mental percibida por trabajadores de hospital y de la industria farmacéutica mediante la aplicación de los métodos FPSICO, ISTAS21 y NASA-TLX, concluyendo la existencia de cierta relación entre las dimensiones y preguntas de los tres métodos, referidas a la carga mental, aunque los valores de los índices de correlación no fueron muy elevados y se encontraron correlaciones entre elementos, que conceptualmente resultaban de difícil explicación.

Otros estudios se han dedicado al análisis de la fiabilidad del NASA-TLX. Por ejemplo, Vidulich y Tsang (1987) encontraron valores medios test-retest de 0.42 y apreciaron que la sensibilidad y la fiabilidad del NASA-TLX eran inferiores a las de otras técnicas como SWORD. Sin embargo, Batiste y Bortolussi (1988) encontraron un índice de estabilidad temporal del NASA-TLX elevado ( $r=0.769$ ) y superior al de la técnica SWAT. Más recientemente, Xiao, Wang, Wang y Lan (2005) han estudiado las propiedades psicométricas del NASA-TLX en una muestra de trabajadores de diferentes sectores profesionales y han encontrado índices de fiabilidad por el método de las dos mitades

y mediante el coeficiente alpha de Cronbach superiores a 0.80. Además, los coeficientes de correlación entre los ítems y el total de la escala fueron superiores a 0.60 en todo los casos excepto para la dimensión rendimiento.

Por otro lado, aunque desde todos los sectores es ampliamente reconocida la necesidad de medir la carga mental que perciben los trabajadores, todavía son muy escasos los estudios en los que se hayan aplicado técnicas como el NASA-TLX en contextos laborales reales (Dalmau, 2007). La mayoría de las investigaciones se han desarrollado y aplicado principalmente en contextos de laboratorio y con muestras de estudiantes, donde los participantes debían evaluar la carga mental producida por tareas "artificiales", bastante diferentes a las que se realizan en los contextos reales (Rubio, Díaz y Martín, 2001). La aplicación en entornos laborales es muy limitada en general y, más en concreto, en España (Dalmau, 2007). Además, y como señalan Tsang y Wilson (1997), los estudios realizados hasta el momento han proporcionado datos parciales e información algo contradictoria sobre las características psicométricas de este instrumento, por lo que se hace necesaria la realización de más estudios sistemáticos al respecto. Por todo ello, nuestra investigación tiene por objetivo evaluar la carga de trabajo mental a través del NASA-TLX en población trabajadora española y analizar las principales características psicométricas de esta técnica en nuestro país.

## Método

### Participantes

La muestra total estuvo formada por 398 participantes, todos ellos trabajadores pertenecientes a 7 grupos profesionales diferentes. La distribución por grupos profesionales fue la siguiente: personal administrativo,  $n=62$  (15%), concejales,  $n=37$  (9,1%), personal de seguridad,  $n=159$  (40,9%), personal docente,  $n=29$  (7,2%), periodistas,  $n=12$  (3%), personal sanitario,  $n=83$  (20,8%), personal de mantenimiento,  $n=16$  (4%). La media de edad de todos los participantes fue de 39,75 años ( $DT=9,59$ ). Del total de la muestra, 244 (61,5%) eran hombres y 154 (38,5%) mujeres.

Respecto al contenido general de los distintos puestos, el personal administrativo tenía como funciones principales el manejo de bases de datos, archivo, redacción de escritos, informes y tareas de atención telefónica. Las funciones más importantes incluidas en el puesto de concejal eran la asistencia a plenos, actos políticos, entrevistas con colectivos sociales, municipios y barrios, así como la gestión y solución de problemas. El objetivo fundamental del personal de seguridad se desglosaba en tres tareas clave: seguridad, vigilancia y protección de clientes. El personal docente desempeñaba las funciones referidas a la formación: preparación de

seminarios, impartición de clases, exámenes, entrevisas y atención a alumnos. Las funciones desarrolladas por los periodistas eran: entrevistar, ruedas de prensa, tareas de investigación y redacción de noticias. El personal sanitario se ocupaba fundamentalmente de la atención al paciente en los diferentes servicios. Por último, el personal de mantenimiento atendía las demandas en los arreglos necesarios y realizaba revisiones periódicas dependiendo de la labor desempeñada (fontaneros, electricistas, jardinero).

### Instrumentos

Para evaluar la carga mental, se utilizó el instrumento NASA-TLX (Task Load Index). Este instrumento se basa en el supuesto de que la carga mental es un constructo hipotético que representa el costo en el que incurre el trabajador al tratar de alcanzar un nivel específico de rendimiento. De esta forma el nivel de carga mental surgiría de la interacción entre los requisitos de una tarea, las circunstancias bajo las que ésta es realizada y las habilidades, conductas y percepciones del operador (Hart et al., 1988). La técnica NASA-TLX distingue las seis dimensiones de carga mental siguientes:

- **Esfuerzo.** Grado de esfuerzo mental y físico que tiene que realizar el sujeto para obtener su nivel de rendimiento.
- **Demandas mentales.** Cantidad de actividad mental y perceptiva que requiere la tarea (p. e.: pensar, decidir, calcular, recordar, mirar, buscar, etc.).
- **Demandas físicas.** Cantidad de actividad física que requiere la tarea (p.e.: pulsar, empujar, girar, etc.).
- **Demandas temporales.** 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.
- **Nivel de Frustración.** Hasta qué punto el sujeto se siente inseguro, estresado, irritado, descontento, etc. durante la realización de la tarea.

Su aplicación requiere dos fases. El objetivo de la primera 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. 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 entre 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).

En la segunda fase, 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 debida a cada una de las seis dimensiones. Con los datos obtenidos en las dos fases se calcula un índice global de carga mental de la tarea, en función de la suma de las valoraciones de cada dimensión ponderada por su peso obtenido en la fase de comparaciones binarias y dividido entre 15. Para una mayor descripción del procedimiento se puede consultar la NTP (Nota Técnica de Prevención) número 544 del Instituto Nacional de Seguridad e Higiene en el Trabajo (Arquer y Nogareda, 1999).

### Procedimiento

Una vez que se definieron las empresas y los grupos profesionales interesados en colaborar, el equipo investigador se reunió con los responsables de cada uno de los trabajadores para informarles del objetivo de la investigación. Todos los participantes fueron informados de las características prácticas del estudio en el que iban a colaborar, haciendo especial hincapié en que se trataba de una participación totalmente voluntaria y anónima. Siguiendo el procedimiento de aplicación de la técnica NASA-TLX, se procedió en primer lugar a realizar la fase de comparaciones binarias entre las dimensiones de carga, y posteriormente a la valoración de la carga mental del puesto de trabajo en cada una de las dimensiones.

### Resultados

En la Tabla 1 se muestran las medias y desviaciones típicas de las puntuaciones de carga mental obtenidas por la muestra total y por cada uno de los grupos profesionales, tanto para el índice global de carga como para cada una de las dimensiones que distingue el NASA-TLX.

Debido a la ausencia de datos publicados que permitan interpretar los niveles de carga mental obtenidos con el NASA-TLX, y como una primera aproximación a la realización de una baremación, se calcularon los percentiles para cada una de las dimensiones y para el índice global de carga mental, con los datos obtenidos para el total de participantes. Los cortes de baremación fueron realizados por decatípos. Los resultados de este análisis aparecen en la tabla 2.

Los resultados muestran que docentes y periodistas son los que se sitúan en el percentil más alto (80), mientras que el personal de mantenimiento es el que se encuadra en el percentil más bajo (30).

Por dimensiones, las mayores puntuaciones en esfuerzo y demanda mental fueron las emitidas por el personal docente y el grupo que asignó menores valoraciones en estas dimensiones fue el personal de man-

Tabla 1. Medias y desviaciones típicas (entre paréntesis) de las evaluaciones de carga mental en cada una de las dimensiones del NASA-TLX, y puntuación global, para cada grupo profesional y la muestra total

	Esfuerzo	Demandas Mental	Demandas Física	Demandas Temporal	Rendimiento	Frustración	Índice Global
Personal administrativo	55.71 (23.79)	70.90 (22.60)	32.05 (25.16)	62.98 (25.02)	73.15 (19.69)	36.97 (23.23)	55.29 (15.99)
Concejales	62.14 (18.14)	72.84 (19.43)	37.16 (21.19)	75.65 (18.33)	73.43 (18.00)	33.92 (27.40)	59.18 (11.90)
Personal docente	75.10 (20.34)	79.79 (16.93)	59.38 (24.99)	79.93 (17.10)	69.41 (23.07)	41.41 (29.75)	67.50 (11.08)
Periodistas	58.25 (22.43)	75.33 (15.47)	31.67 (18.10)	76.67 (13.22)	77.83 (13.70)	34.00 (17.78)	67.50 (9.93)
Personal de seguridad	59.25 (23.38)	70.16 (26.53)	49.47 (27.68)	61.03 (25.37)	74.74 (23.69)	44.78 (29.54)	58.95 (16.64)
Personal sanitario	70.65 (20.72)	75.87 (19.23)	52.07 (30.41)	66.55 (21.43)	77.80 (16.41)	41.64 (27.26)	59.90 (14.05)
Personal de mantenimiento	51.88 (26.51)	54.06 (28.08)	51.06 (28.17)	45.19 (26.41)	82.75 (15.51)	23.19 (20.29)	51.35 (16.95)
Muestra Total	62.17 (23.12)	71.92 (23.42)	46.40 (28.07)	65.06 (24.16)	75.03 (20.67)	40.46 (27.67)	60.17 (15.48)

Tabla 2. Percentiles de las puntuaciones en cada una de las dimensiones y el índice global de carga mental

	Esfuerzo	D.Mental	D.Física	D.Temporal	Rendimiento	Frustración	Índice Global
<b>10</b>	30.00	40.00	10.00	30.00	50.00	2.90	40.15
<b>20</b>	40.00	50.00	20.00	50.00	60.00	10.00	48.30
<b>30</b>	50.00	68.00	30.00	51.00	70.00	20.00	53.45
<b>40</b>	60.00	70.00	36.00	60.00	75.00	30.00	57.50
<b>50</b>	67.50	80.00	49.00	70.00	80.00	40.00	61.66
<b>60</b>	70.00	80.00	50.00	79.00	81.00	50.00	65.23
<b>70</b>	79.00	89.00	60.00	80.00	89.00	54.00	68.50
<b>80</b>	80.00	90.00	70.00	89.00	90.00	69.00	73.16
<b>90</b>	90.00	99.10	90.00	91.00	100.0	80.00	78.20

tenimiento. En demanda física, de nuevo el personal docente mostró la mayor puntuación. El grupo de periodistas fue el que mostró menor demanda física. Respecto a la demanda temporal, los docentes dieron los máximos valores con respecto al resto, mientras que el personal de mantenimiento estaba en el percentil más bajo. Por el contrario, en rendimiento, el personal de mantenimiento obtuvo las mayores valoraciones y el personal docente las puntuaciones más bajas. La mayor frustración se dio en el sector de la seguridad y la menor en los técnicos de mantenimiento.

Según el puesto de trabajo, el personal docente fue el que obtuvo mayores puntuaciones en cuatro de las seis dimensiones y por otra parte, fue el grupo que valoró más bajo su rendimiento. El personal de mantenimiento dio las puntuaciones más bajas en esfuerzo, demanda mental, demanda temporal y nivel de frustración, sin embargo, en su valoración de rendimiento fue el grupo con puntuación más alta. En cuanto al personal sanitario, éste ocupó el segundo lugar con puntuaciones más altas en esfuerzo, demanda mental y demanda física. El personal administrativo obtuvo

bajas puntuaciones en esfuerzo, demanda mental, demanda física, demanda temporal y rendimiento. Los concejales mostraron valores altos en esfuerzo y demanda temporal, mientras que en frustración, demanda física y rendimiento mostraron bajas puntuaciones. Respecto a los periodistas, fue el grupo con valores más bajos en demanda física, obteniendo sus valores más altos en demanda temporal y rendimiento. El personal de seguridad fue el grupo con puntuaciones más altas en frustración, en el resto de las dimensiones se encontraron entre los percentiles 40 y 50.

### Consistencia interna

Con objeto de analizar la fiabilidad de la técnica NASA-TLX, entendida como consistencia interna, se calculó el coeficiente alpha de Cronbach para la escala total. Los resultados obtenidos en este análisis mostraron un índice de consistencia  $\alpha = 0.69$ . Por otra parte, el alpha de Cronbach para la escala total, si se elimina cada una de las dimensiones fue de  $\alpha = 0.58$ , si se eliminara la dimensión esfuerzo;  $\alpha = 0.62$ , si se eliminara

la dimensión demanda temporal;  $\alpha = 0.66$ , si se eliminara la dimensión demanda física;  $\alpha = 0.63$ , si se eliminara la dimensión demanda temporal;  $\alpha = 0.66$ , si se eliminara la dimensión rendimiento; y  $\alpha = 0.72$ , si se eliminara la dimensión frustración. Por tanto, parece que si se eliminara la dimensión frustración, la consistencia interna del NASA-TLX aumentaría ligeramente.

### **Estructura factorial**

Para estudiar la estructura factorial de la escala NASA-TLX se realizó un análisis factorial de componentes principales con rotación ortogonal Varimax. El coeficiente de adecuación muestral KMO fue igual a 0.714. El valor de la prueba de esfericidad de Bartlett resultó significativo,  $\chi^2(15)=476.28$ ,  $p \leq .001$ . Se obtuvieron dos factores principales, el primero de ellos formado por cinco dimensiones y el segundo integrado únicamente por la dimensión frustración.

La tabla 3 muestra la matriz de componentes rotados. Los dos factores obtenidos explicaban el 59.15% de la varianza total.

Tabla 3. Estructura factorial de NASA-TLX. Matriz factorial rotada

	Componente	
	1	2
Esfuerzo	.829	.162
Rendimiento	.721	-.168
D. Mental	.688	.245
D. Física	.605	.154
D. Temporal	.491	.563
Frustración	-.022	.897

### **Análisis de perfiles de carga mental en función del grupo profesional**

Para estudiar si existen diferencias significativas entre los perfiles de carga mental de los distintos grupos profesionales, se realizó un análisis de varianza de medidas repetidas, considerando como factor entre sujetos el grupo profesional y como factor de medidas repetidas el perfil de carga, con los seis niveles correspondientes a las dimensiones de carga que establece el NASA-TLX. Para comprobar la homogeneidad multivariada de las varianzas, se calculó el estadístico  $M$  de Box, el cual resultó significativo ( $F(126, 14405)=1.68$ ;  $M=234.255$ ;  $p \leq .001$ ), por lo que se utilizó la lambda de Wilks con criterio de Pillai (Tabanick y Fidell, 1989). Resultaron significativos tanto los efectos principales de ambos factores ( $F(6, 391)=4.09$ ,  $p \leq .001$ ,  $\eta^2 = .06$ , para el grupo profesional;  $F(5, 387)=88.65$ ,  $p \leq .001$ ,  $\eta^2 = .53$ , para el perfil de carga mental), como el efecto de interacción ( $F(30, 1955)=4.12$ ,  $p \leq .001$ ,  $\eta^2 = .06$ ).

Las comparaciones a posteriori entre los grupos

mostraron diferencias estadísticamente significativas en todas las dimensiones de carga, excepto para rendimiento. Los resultados de este análisis para las cinco dimensiones en las que se encontraron algunas comparaciones significativas (resaltadas en negrita) aparecen en la tabla 4.

Destacan las diferencias estadísticamente significativas entre administrativos y profesores, en concreto en las dimensiones demanda física, esfuerzo y demanda temporal. Las diferencias negativas indican valores más altos de carga para los profesores. En estas dimensiones también resultan significativas las comparaciones entre el grupo de profesores y el personal de seguridad, puntuando más alto los profesores. Además el grupo de profesores obtuvo diferencias estadísticamente significativas al compararse con el personal de mantenimiento. En general, de las seis dimensiones incluidas en el NASA-TLX, la frustración fue la variable que mostró menos diferencias, mientras que demanda temporal fue la que obtuvo el mayor número de diferencias estadísticamente significativas de entre todas las comparaciones de los grupos.

### **Discusión**

En países como Estados Unidos, se realizan desde hace algunos años, estudios para comparar distintos métodos de evaluación de la carga mental, sobre todo de tipo subjetivo. En España todavía no existe suficiente investigación al respecto, por lo que una de las principales dificultades que presenta la evaluación de la carga mental es el escaso número de estudios en los que se han aplicado este tipo de técnicas en contextos reales de trabajo. La mayoría de los instrumentos existentes se han desarrollado en situaciones de laboratorio, donde los participantes debían evaluar la carga mental producida por tareas “artificiales”, bastante diferentes a las que se realizan en los contextos reales. Este ha sido el principal motivo por el que el objetivo del presente estudio ha sido la evaluación la carga mental a través del NASA-TLX en población trabajadora y el análisis de las propiedades psicométricas de dicho instrumento.

En general, los resultados indican que las mayores valoraciones de carga mental fueron las asignadas por el personal docente y los periodistas, mientras que el personal de mantenimiento fue el grupo profesional con los valores más bajos.

Debido a la inexistencia de datos normativos, en el presente estudio se calcularon, para el total de la muestra, los percentiles en cada una de las dimensiones y en el índice global de carga mental. Estos datos representan una primera aproximación en la baremación de los resultados de carga mental, que pueden ser de ayuda en la interpretación de resultados obtenidos en otros estudios.

Los resultados de este estudio han mostrado que el NASA-TLX es un instrumento cuya consistencia inter-

Tabla 4. Resultado de las comparaciones múltiples a posteriori entre los perfiles de carga mental de los diferentes grupos profesionales

(I)	(J)	Esfuerzo		D. Mental		D. Física		D. Temporal		Frustración	
		Diferencia de medias (I-J)	Sig.								
Administrativo	Concejal	-6.42	.168	-1.93	.687	-5.11	.362	<b>-12.66</b>	<b>.009</b>	3.04	.592
	Profesor	<b>-19.39</b>	<b>.000</b>	-8.89	.088	<b>-27.33</b>	<b>.000</b>	<b>-16.94</b>	<b>.001</b>	-4.44	.471
	Periodista	-2.54	.719	-4.43	.544	.38	.964	-13.68	.062	2.96	.731
	Seguridad	-3.53	.292	.74	.829	<b>-17.42</b>	<b>.000</b>	1.95	.574	-7.81	.058
	Sanitario	<b>-14.94</b>	<b>.000</b>	-4.96	.201	<b>-20.02</b>	<b>.000</b>	-3.57	.360	-4.67	.310
	Mantenimiento	3.83	.541	<b>16.84</b>	<b>.010</b>	<b>-19.01</b>	<b>.012</b>	<b>17.79</b>	<b>.006</b>	13.78	.074
Concejal	Profesor	<b>-12.96</b>	<b>.020</b>	-6.95	.226	<b>-22.21</b>	<b>.001</b>	-4.28	.457	-7.49	.271
	Periodista	3.88	.601	-2.49	.745	5.49	.540	-1.01	.895	-.08	.993
	Seguridad	2.89	.479	2.68	.525	<b>-12.31</b>	<b>.013</b>	<b>14.61</b>	<b>.001</b>	<b>-10.86</b>	<b>.030</b>
	Sanitario	-8.51	.055	-3.03	.508	<b>-14.91</b>	<b>.005</b>	<b>9.09</b>	<b>.048</b>	-7.72	.155
Profesores	Mantenimiento	10.26	.126	<b>18.77</b>	<b>.007</b>	-13.90	.086	<b>30.46</b>	<b>.000</b>	10.73	.191
	Periodista	<b>16.85</b>	<b>.029</b>	4.46	.574	<b>27.71</b>	<b>.003</b>	3.26	.682	7.41	.431
	Seguridad	<b>15.85</b>	<b>.000</b>	<b>9.63</b>	<b>.040</b>	9.90	.070	<b>18.90</b>	<b>.000</b>	-3.36	.543
	Sanitario	4.45	.357	3.92	.431	7.30	.210	<b>13.37</b>	<b>.008</b>	-.22	.970
Periodistas	Mantenimiento	<b>23.22</b>	<b>.001</b>	<b>25.73</b>	<b>.000</b>	8.31	.323	34.74	<b>.000</b>	<b>18.22</b>	<b>.033</b>
	Seguridad	-.99	.882	5.17	.455	<b>-17.80</b>	<b>.028</b>	<b>15.63</b>	<b>.025</b>	-10.78	.189
	Sanitario	-12.40	.073	-.53	.940	<b>-20.40</b>	<b>.015</b>	10.11	.159	-7.63	.367
Seguridad	Mantenimiento	6.37	.456	<b>21.27</b>	<b>.016</b>	-19.39	.061	<b>31.47</b>	<b>.000</b>	10.81	.302
	Sanitario	<b>-11.40</b>	<b>.000</b>	-5.71	.069	-2.60	.477	-5.52	.079	3.14	.398
	Mantenimiento	7.37	.210	<b>16.09</b>	<b>.008</b>	-1.59	.822	<b>15.84</b>	<b>.010</b>	<b>21.59</b>	<b>.003</b>
Sanitarios	Mantenimiento	<b>18.77</b>	<b>.002</b>	<b>21.80</b>	<b>.001</b>	1.01	.891	<b>21.36</b>	<b>.001</b>	<b>18.45</b>	<b>.014</b>

na es aceptable, aunque ésta aumentaría ligeramente si se eliminase la dimensión “frustración”. En este sentido, nuestros resultados son acordes con los obtenidos en investigaciones previas, aunque el índice encontrado en este caso es algo inferior al obtenido por Xiao et al. (2005).

En cuanto a la estructura factorial, los resultados mostraron una configuración bifactorial. El primer factor estuvo formado por las dimensiones esfuerzo, demanda mental, demanda física, demanda temporal y rendimiento. El segundo factor estuvo formado únicamente por la dimensión frustración. Este resultado, unido al anterior, pone de manifiesto, que los trabajadores perciben la frustración que les genera su puesto de trabajo como un aspecto algo diferente al resto de dimensiones. Desde nuestro punto de vista, este resultado puede deberse simplemente a la escasa variabilidad que existe en los niveles de frustración entre los grupos profesionales considerados, que se pone de manifiesto en las numerosas comparaciones entre los grupos que no resultaron significativas en frustración (ver tabla 4). Llaman la atención las valoraciones tan bajas en frustración que proporciona la mayoría de trabajadores. Ninguno de los grupos profesionales llega, como media, a la puntuación 50 (en una escala de 100) en esta dimensión. Posiblemente, cierto grado de deseabilidad social pueda estar afectando a las puntuaciones obtenidas en esta dimensión. Consideramos que en futuras investigaciones sería muy interesante estudiar y controlar el posible efecto de la deseabilidad social

sobre las valoraciones subjetivas de las fuentes de carga mental del trabajo.

Por otro lado, la existencia de diferencias significativas entre los perfiles de carga mental de los distintos grupos profesionales avala su validez como instrumento de evaluación de carga mental. En este sentido, el NASA-TLX se ha mostrado eficaz para discriminar entre la carga mental experimentada por grupos profesionales que se enfrentan en su quehacer cotidiano a niveles de demandas muy diferentes, como es el caso de los profesores comparados con el personal de mantenimiento o con administrativos, por ejemplo. Estos resultados son similares a los obtenidos por Rolo, Díaz y Hernández (2009) con la escala ESCAM.

Por último, es necesario señalar que el presente estudio representa una primera aproximación en la evaluación de la carga mental de trabajo mediante la técnica NASA-TLX en diferentes profesiones en España. Nuestro trabajo ha puesto de manifiesto la eficacia del NASA-TLX como instrumento de evaluación de la carga mental en entornos laborales españoles. La evaluación mediante el NASA-TLX no sólo proporciona información acerca del nivel global de carga, sino que permite detectar de manera válida y fiable, las fuentes de carga. El conocimiento de estas fuentes de carga permitirá a los responsables de las empresas modificar las funciones de los puestos con el fin de reducir su nivel de carga y por tanto mejorar la satisfacción y la salud de los trabajadores.

## Extended Summary

Mental workload is growing as a factor of psychosocial risk, which means that, at present, the assessment of the mental workload is a key aspect of the research and development of working systems that allow obtain higher levels of comfort, satisfaction, efficacy and safety and health at work.

Of the various procedures proposed for assessing mental workload (subjective techniques, physiological indicators and measures of task performance) are subjectively distinguished by having high levels of sensitivity, minimum implementation requirements, good acceptance by workers, a degree of intrusion with the task practically nil, and adequate levels of validity. These procedures reflect the subject's direct opinion about the mental workload experienced while performing a task. In practice, multidimensional subjective procedures are the most used, mainly because they allow differentiating between the various causes of mental workload, keeping above the psychometric characteristics (Tsang & Wilson, 1997). All of them can score a load for each of the dimensions that distinguish and an overall score of the mental workload of a task (O'Donnell & Eggemeier, 1986, Tsang & Wilson, 1997; Arquer & Nogareda, 1999).

Among the procedures most frequently used subjective multidimensional and more developed countries, highlights the NASA-TLX (Task Load Index, and Staveland Hart, 1988). Currently, in Spain, has introduced a new subjective scale (ESCAM), but is still working on its generalizability and construct validity (Rolo, Diaz & Hernandez, 2009).

Although from all sectors is widely recognized the need to measure mental workload received by workers, are still very few studies in which validated techniques have been applied, such as NASA-TLX in real work settings (Dalmau, 2007). Most research has been developed and applied mainly in laboratory settings and with samples of students, where participants had to assess the burden caused by mental tasks "artificial", quite different from those made in real contexts (Rubio et al., 2001). The application in industrial environments is very limited in general and specifically in Spain (Dalmau, 2007). Therefore, our research is to evaluate the mental workload by NASA-TLX in the Spanish working population and to analyze if confirmed some of the major psychometric properties of this technique in our country.

### Method

#### Participants

The sample was formed by 398 workers. The workers were distributed in 7 different professional groups. The 15% of total sample were administrative, 9.1%

councillors, 40.9 security personnel, 7.2% teachers, 3% journalists, 20.8 % sanitary personnel and 4% were maintenance. The average age of all participants was 39.75 years (SD = 9.59). 244 (61.5%) were male and 154 (38.5%) women.

#### Instruments

The assessment tools were the NASA-TLX and a short questionnaire of demographic data, where it is collected data relating to the age, sex, work experience and expertise in the workplace. To assess the mental workload was used the NASA-TLX. Participants should evaluate the mental workload of each of their job functions. The NASA-TLX technique distinguishes six dimensions of mental workload as follows:

- **Effort.** Degree of mental and physical effort that has to make the subject to get your performance level.
- **Mental demand.** Number of mental and perceptual activity required for the task (eg thinking, deciding, calculating, remembering, looking, searching, etc.).
- **Physical demand.** Amount of physical activity required by the task (eg, push, push, turn, etc.).
- **Temporary demand.** Level of time pressure felt. The ratio of the time required and the available.
- **Performance.** To what extent the individual is satisfied with its performance level.
- **Frustration level.** The extent to which the subject feels insecure, stressed, angry, unhappy, etc. while performing the task.

#### Procedure

The subject completed the NASA-TLX together with a questionnaire of demographic data. All subjects should evaluate the degree in each of the functions of his job demanded each of the six dimensions of mental workload that distinguishes the NASA-TLX.

### Results

Means and standard deviations of mental workload scores obtained by the total sample and for each of the professional groups, both for the overall rate of charge to each of the dimensions distinguish the NASA-TLX. With the data obtained for all participants percentiles were calculated for each of the dimensions and the overall rate of mental workload.

### Internal consistency

In order to analyze the reliability of NASA-TLX technique, defined as internal consistency, we calculated the Cronbach alpha coefficient for the total scale. The results of this analysis showed a consistency index  $\alpha = 0.69$ . Furthermore, the Cronbach alpha for the total scale, removing each of the dimensions was  $\alpha = 0.58$ , if you eliminate the stress dimension,  $\alpha = 0.62$ , removing the temporary demand dimension,  $\alpha = 0.66$ , if you eliminate the physical demand dimension,  $\alpha = 0.63$ , removing the temporary demand dimension,  $\alpha = 0.66$ , if you eliminate the performance dimension, and  $\alpha = 0.72$ , if you eliminate the frustration scale. Therefore, it seems that if eliminate the frustration scale, internal consistency of the NASA-TLX increase slightly.

### Factor structure

To study the structure of the NASA-TLX scale was performed principal components factor analysis with Varimax orthogonal rotation. The KMO sampling adequacy ratio was equal to 0.714. The value of the Bartlett test of sphericity was significant,  $\chi^2 (15) = 476.28$ ,  $p \leq .001$ . We obtained two main factors, the first of which consists of five dimensions and the second dimension consisting of only frustration. The two factors obtained explained the 59.15% of the total variance.

### Analysis of mental workload profiles based on the professional group

To study significant differences between the profiles of mental workload of the professional groups concerned, an analysis of variance for repeated measures between subjects factor considered as professional group and repeated measures factor load profile with the six levels corresponding to the load size which sets the NASA-TLX. To test the multivariate homogeneity of variances, we calculated the Box M statistic, which was significant ( $F (126, 14405) = 1.68$ ,  $M = 234.255$ ,  $p \leq .001$ ), so we used the Wilks' lambda with criterion Pillai (Tabanick and Fidell, 1989). Both were significant main effects of both factors ( $F (6, 391) = 4.09$ ,  $p \leq .001$ ,  $\eta^2 = .06$ , to the professional group,  $F (5, 387) = 88.65$ ,  $p \leq .001$ ,  $\eta^2 = .53$  for the mental load profile) and the interaction effect ( $F (30, 1955) = 4.12$ ,  $p \leq .001$ ,  $\eta^2 = .06$ ). Comparisons between groups were statistically significant in all dimensions of mental workload, except for performance.

### Discussion

In countries like the United States are performed for several years, studies to compare different methods of

assessing mental workload, mostly subjective type. In Spain there is not yet enough research about it, so that one of the main difficulties in the assessment of mental workload is the small number of studies that have applied these techniques work in real contexts. The objective of this study was to evaluate the mental workload by NASA-TLX in the working population and analysis of the psychometric properties of the instrument.

Results indicate that higher ratings of mental workload were assessed by teachers and journalists, while the maintenance staff was professional group with the lowest values. The percentiles in each of the dimensions and the overall rate of mental workload were calculated for the total sample. The results of this study have shown that the NASA-TLX is an instrument whose internal consistency is acceptable, although it increase slightly if the scale were eliminated "frustration". In this sense, our results are consistent with those obtained in previous research (Xiao et al., 2005).

Regarding the factorial structure, the results showed a bivariate setting. The first factor consisted of the dimensions effort, mental demand, physical demand, temporal demand and performance. The second factor was formed solely by the dimension frustration. On the other hand, the existence of significant differences between mental load profiles of different professional groups endorses its validity as an evaluation of mental workload. In this sense, the NASA-TLX has been effective in discriminating between mental workload experienced by professional groups face in their daily work at very different levels of demand. These results are similar to those obtained by Rolo, Diaz and Hernandez (2009) with scale.

Finally, this study represents a first approach to the assessment of mental workload by NASA-TLX technique in different professions in Spain. Knowledge of these sources will charge those responsible for business change job functions in order to reduce the level of burden and thus improve customer satisfaction and health of workers.

### Referencias

- 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.
- Charlton, S.G. (2002). Measurement of cognitive states in test and evaluation. En, S.G. Charlton y F.G.O'Brien (eds.), *Handbook of Human Factors Testing and Evaluation*. Londres: Lawrence Erlbaum Associates.

- Dalmau, I. (2007). *Evaluación de la carga mental en tareas de control: técnicas subjetivas y medidas de exigencia*. Tesis doctoral. Barcelona: UPC.
- Ferrer, R. y Dalmau, I. (2004). Revisión del concepto de carga mental: evaluación, consecuencias y proceso de normalización. *Anuario de Psicología*, 35, 521-545.
- Hacker, W. (1998). Carga mental de trabajo. En Stellman (Ed.), *Enciclopedia de Salud y Seguridad en el Trabajo*. Madrid: Ministerio de Trabajo y Asuntos Sociales 29, 44-46.
- 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* (pp. 139-183). North-Holland, Amsterdam..
- Instituto Nacional de Seguridad e Higiene en el Trabajo (2007): *VI Encuesta Nacional de Condiciones de Trabajo*. Madrid: INSHT.
- Marras, W.S. y Karwowski, W. (2006) (eds.). *Fundamentals and Assessment Tools for Occupational Ergonomics*. Florida: Taylor & Francis.
- Noyes, J. M., y Bruneau, D. P. J. (2007). A self-analysis of the NASA-TLX workload measure. *Ergonomics*, 50, 514-519.
- 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*, Nueva York: Wiley.
- Recarte, M.A., Pérez, E., Conchillo, A. y Nunes, L. M. (2008). Mental workload and visual impairment: differences between Pupil, Blink, and Subjective Rating. *The Spanish Journal of Psychology*, 11, 374-385.
- Rolo, G., Díaz, D., Hernández, E. (2009). Desarrollo de una escala Subjetiva de la Carga Mental de Trabajo (ESCAM). *Revista de Psicología del Trabajo y de las Organizaciones*, 25, 29-37.
- Rubio, S., Díaz, E. y Martín, J. (2001). Aspectos metodológicos de la evaluación subjetiva de la carga mental en el trabajo. *Archivos de Prevención de Riesgos Laborales*, 4, 160-168.
- 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.
- Rutledge, T., Stucky, E., Dollarhide, A., Shively, M., Jain, S. y Wolfson, T., (2009). A real-time assessment of work stress in physicians and nurses. *Health Psychology*, 28, 194-200.
- Tabachnick, B.G. y Fidell, L.S. (U1989). *Using Multivariate Statistics* (2<sup>a</sup> ed.). Nueva York: Harper Collins.
- Tsang, G. y Wilson, G.F. (1997). Mental workload. En G. Salvendy (ed.), *Handbook of Human Factors and Ergonomics*. Nueva York: Wiley & Sons.
- Vidulich, M. A. y Tsang, P. S. (1987). Absolute magnitude estimation and relative judgement approaches to subjective workload assessment. En *Proceedings of the Human Factors Society Thirty-First Annual Meeting*, 1057-1061, Santa Monica, CA: Human Factors Society.
- Xiao, Y.M., Wang, Z.M., Wang, M.Z. y Lan, Y.J. (2005). The appraisal of reliability and validity of subjective workload assessment and NASA-task load index. *Chinese Journal of Industrial Hygiene and Occupational Diseases*, 23, 178-181.
- Young, M. S. y Stanton, N. A. (2001). Mental Workload Theory, Measurement and Application, en Karwowski (Ed.) *International Encyclopedia of Ergonomics and Human Factors* (pp 507-509). London, New York: Taylor and Francis.

Manuscrito recibido: 23/11/2009

Revisión recibida: 13/10/2010

Aceptado: 08/11/2010