

CARMEN: UNA NUEVA HERRAMIENTA PARA LA EVALUACIÓN DE LA CARGA MENTAL EN EL TRABAJO

Susana Rubio-Valdehita^{1*}, Eva M^a Díaz-Ramiro^{}, M^a Inmaculada López-Núñez^{**} y M^a Dolores Vallengano-Pérez^{**}**

**Departamento de Psicología Diferencial y del Trabajo, Facultad de Psicología, Universidad Complutense de Madrid; **Universidad Complutense de Madrid, España*

RESUMEN. Antecedentes: Como señala la Norma ISO 10075, se hace necesario disponer de herramientas de evaluación de la carga mental de trabajo que posean adecuados índices de calidad psicométrica. Las técnicas disponibles actualmente carecen de un estudio psicométrico adecuado o incluyen procedimientos de difícil aplicación en el contexto real de trabajo. Por ello, el objetivo de esta investigación es presentar una nueva herramienta de evaluación de la carga mental de trabajo que cumpla con los estándares que señala la Norma ISO 10075. Método. Siguiendo los modelos de Hart y Staveland y de Xie y Salvendy se diseñó un cuestionario de 50 ítems (que hemos denominado Cuestionario CARMEN), que fue aplicado a 428 trabajadores de los sectores sanitarios, seguridad y administrativo. Se realizaron análisis de fiabilidad (consistencia interna) y de validez del instrumento (análisis factorial confirmatorio). Resultados. Se comparan diferentes versiones del cuestionario CARMEN, desde el punto de vista de sus índices de validez (ajuste al modelo factorial confirmatorio) y fiabilidad. Conclusiones. La versión reducida de 22 ítems y cinco dimensiones (demanda cognitiva, demanda temporal, demanda física, demanda emocional y exigencias de rendimiento) es la que muestra mejores indicadores de ajuste desde el punto de vista psicométrico. Algunas de las dimensiones inicialmente planteadas al construir la versión inicial del cuestionario CARMEN (entorno de trabajo y ajuste del puesto a las características individuales) desaparecen del modelo final, lo cual podría indicar que al evaluar la carga mental, este tipo de variables son subyacentes a la valoración del resto de las dimensiones de carga.

Palabras clave: Carga mental, cuestionario CARMEN, análisis factorial confirmatorio, psicología del trabajo, psicometría

ABSTRACT. Background: As stated in ISO 10075, it is necessary to have mental workload assessment tools that have adequate psychometric quality indexes. The currently available techniques lack adequate psychometric study or include procedures that are difficult to apply in the actual context of work. Therefore, the objective of this research is to present a new tool to evaluate the mental workload that complies with the standards set in ISO 10075. Method. Following the models of Hart and Staveland and of Xie and Salvendy,

¹ Correspondencia: Susana Rubio-Valdehita. Facultad de Psicología, Universidad Complutense de Madrid, Campus de Somosaguas, 28223, Pozuelo de Alarcón (Madrid). E-mail: srubiova@ucm.es

a questionnaire of 50 items (called CARMEN Questionnaire) was designed, which was applied to 428 workers in the health, safety and administrative sectors. We performed analyses of reliability (internal consistency) and instrument validity (confirmatory factor analysis). Results. Different versions of the CARMEN questionnaire are compared, from the point of view of their validity indices (adjustment to the confirmatory factorial model) and reliability. Conclusions. The reduced version of 22 items and five dimensions (cognitive demand, temporal demand, physical demand, emotional demand and performance requirements) is the one that shows the best indicators of adjustment from the psychometric point of view. Some of the dimensions initially proposed in constructing the initial version of the CARMEN questionnaire (working environment and job adjustment to individual characteristics) disappear from the final model, which may indicate that when assessing the mental load, these types of variables are underlying the evaluation of the rest of the mental workload dimensions.

Keywords: Mental workload, CARMEN questionnaire, confirmatory factor analysis, work psychology, psychometry.

Introducción

Los factores de riesgo psicosocial son las condiciones presentes en una situación laboral, directamente relacionadas con la organización, el contenido del trabajo y la realización de la tarea, que pueden afectar al bienestar, la salud y el rendimiento de los trabajadores. Entre estos factores se incluye la carga mental como uno de los más importantes factores de riesgo psicosocial. La VI Encuesta Nacional de Condiciones de Trabajo refleja el aumento de la carga mental como factor de riesgo laboral, ya que el 67% de los trabajadores encuestados manifestó estar sometido a elevadas demandas de tipo cognitivo, y aproximadamente el 20% respondió que siempre o casi siempre tenía que realizar varias tareas al mismo tiempo durante su jornada laboral. La carga mental puede definirse como el nivel de recursos atencionales necesarios para equilibrar los criterios de ejecución objetivos y subjetivos. Los recursos atencionales son limitados, por lo cual cuando las demandas de la tarea exceden la capacidad del individuo se produce un descenso en el rendimiento laboral y un aumento en los sentimientos de insatisfacción, estrés y frustración de los trabajadores. Estar expuesto a condiciones de carga mental en el trabajo puede producir importantes problemas de salud en los trabajadores. La consecuencia más directa es el estrés laboral que, mantenido en el tiempo, puede dar lugar a problemas físicos, psicológicos y organizacionales (Gil-Monte, 2009).

La complejidad del concepto de carga mental ha llevado al desarrollo de gran variedad de técnicas de evaluación: indicadores de rendimiento, medidas fisiológicas y escalas y cuestionarios. Actualmente, son muchos los modelos teóricos de carga mental que consideran fundamental plantear un enfoque centrado en los aspectos subjetivos, ya que será la percepción subjetiva del trabajador la que determinará su satisfacción y bienestar (Szalma, 2008). Además, las técnicas fisiológicas y de rendimiento implican el uso de métodos intrusivos que impiden la evaluación fiable de la carga mental en situaciones reales de trabajo, llegando a ser rechazados por los trabajadores. Todo ello ha llevado al desarrollo de escalas subjetivas que en la mayoría de los casos han surgido en entornos específicos lo que dificulta, y en algunos casos imposibilita, su aplicación y utilización de forma generalizada en cualquier contexto. Investigadores de la Universidad de la Laguna han desarrollado la escala ESCAM que permite obtener puntuaciones en varias dimensiones de carga: demandas cognitivas, organización del tiempo, efectos

sobre el individuo, dificultades e interferencias, retroalimentación y apoyo de terceras personas, y responsabilidad sobre el trabajo realizado. Los resultados obtenidos (Rolo-González, Díaz y Hernández-Fernaud, 2009), señalan las propiedades psicométricas del instrumento; sin embargo, como sus autoras destacan, la utilidad de la escala ESCAM es todavía limitada.

A nivel internacional la escala Nasa-TLX (Hart y Staveland, 1988) es la que ha sido objeto de mayor investigación y uso. Distingue seis dimensiones de carga mental (esfuerzo, demanda mental, demanda física, demanda temporal, rendimiento y frustración) que son evaluadas de 0 a 100. Nasa-TLX ha demostrado ser una herramienta sensible, válida y fiable (Rubio, Díaz, Martín y Puente, 2004). Sin embargo, posee limitaciones de carácter práctico especialmente en lo referido a su formato de presentación, poco habitual en los ámbitos reales de trabajo, donde los trabajadores están mucho más acostumbrados a responder a cuestionarios conformados por una serie de ítems y en los que claramente se les pide que expresen su opinión acerca de la intensidad, importancia o frecuencia con la que se presenta en su puesto de trabajo una determinada condición referida en un ítem concreto, utilizando para ello una escala tipo Likert.

La abundancia de procedimientos propuestos es buen indicador de las limitaciones que entrañan todos ellos. Las principales dificultades que seguimos encontrando a la hora de evaluar la carga mental son que la mayoría de los instrumentos son específicos del dominio, muchos sólo son aplicables en puestos de producción, algunos son protocolos de actuación que no establecen el procedimiento de medida, en ocasiones son métodos intrusivos y su cumplimentación presenta cierta complejidad para algunos trabajadores.

En este trabajo se presentan los resultados del estudio psicométrico de una nueva herramienta de evaluación y diagnóstico de la carga mental, que hemos denominado CARMEN, y que pretende ser un cuestionario válido, fiable y de fácil aplicación a cualquier sector profesional.

Método

Participantes

Participaron 428 trabajadores pertenecientes a los sectores de sanidad, seguridad y administrativo. 232 eran mujeres (54,2%) y 196 hombres (45,8%). La Tabla 1 muestra la distribución de los participantes por sector y sexo. Todos tenían una antigüedad en el puesto superior a seis meses. La media de edad fue de 36,4 años (DT = 10,3) variando entre 25 y 60 años.

Tabla 1

Distribución de la muestra por sector profesional y sexo

	Mujer	Varón	Total
Sanitario	163	51	214
Administrativo	60	25	85
Seguridad	9	120	129

Materiales/Instrumentos

Todos los participantes contestaron el cuestionario CARMEN para la evaluación de la carga mental de sus puestos de trabajo. El cuestionario CARMEN está formado por 50 ítems referidos a las siguientes dimensiones: exigencias cognitivas o de procesamiento (p.e. Mi trabajo requiere mantener periodos de intensa concentración), exigencias temporales o de ritmo de trabajo (p.e. No puedo parar o detener mi trabajo cuando lo necesito), exigencias de rendimiento o ejecución (p.e. Mis errores pueden tener

consecuencias graves), demanda emocional o de salud física y emocional (p.e. Mi trabajo me afecta mucho emocionalmente), entorno de trabajo (p.e. Dispongo de las herramientas adecuadas para hacer bien mi trabajo) y características individuales (p.e. Las exigencias de mi trabajo se corresponden con mi formación académica). Cada uno de los ítems es evaluado en una escala Likert de 4 puntos (0 nunca; 1 pocas veces; 2 a menudo; 3 siempre).

Procedimiento

Previo a la aplicación de los instrumentos, el equipo investigador contactó con los responsables de los servicios de prevención de los centros de trabajo, los cuales, tras la autorización del estudio, se encargaron de organizar los momentos más adecuados para la recogida de los datos, así como de informar a la plantilla. Todos los participantes fueron informados del carácter anónimo y voluntario de su participación en el estudio y dieron su consentimiento para la utilización de los datos con fines de investigación.

Resultados

Se analizó la validez y la fiabilidad (consistencia interna) de tres versiones del cuestionario CARMEN: versión inicial de 50 ítems (CARMEN 1.0), versión reducida de 35 ítems (CARMEN 2.0) y reducida a 22 ítems (CARMEN 3.0). Para analizar la validez se utilizó el análisis factorial confirmatorio (AFC) mediante el programa AMOS. La fiabilidad fue analizada mediante el coeficiente alpha de Cronbach, calculado con el SPSS 22.0.

Se probaron diferentes modelos factoriales:

CARMEN 1.0: Modelo de 50 ítems caracterizado por seis factores de primer orden (demandas cognitivas, temporales, emocionales y físicas, rendimiento, entorno organizacional y ajuste individual) y un factor de segundo orden (carga mental).

CARMEN 2.0: Modelo unidimensional (carga mental) de 35 ítems.

CARMEN 2.0: Modelo de 35 ítems caracterizado por 4 factores (demandas cognitivas, demandas temporales, demandas emocionales y físicas y exigencias de rendimiento) de primer orden y un factor de segundo orden (carga mental).

CARMEN 3.0: Modelo de 22 ítems caracterizado por 5 factores (demandas cognitivas, demandas temporales, demandas emocionales, demandas físicas y exigencias de rendimiento) de primer orden y un factor de segundo orden (carga mental).

En los dos últimos modelos, se introdujeron modificaciones según los índices de modificación, los de bondad de ajuste y las cargas factoriales. Considerando los resultados, el último modelo mostró el mejor ajuste (ver Tabla 2). En todos los casos se obtuvieron KMO superiores a 0,80 y niveles de significación inferiores a 0,001. Además, se utilizó el coeficiente de Mardia (Bollen, 1989), considerando normalidad multivariante cuando es menor que $p(p+2)$, donde p es el número de variables observadas. Se examinó la magnitud χ^2 indicando un mejor ajuste aquel en el que ésta fue más pequeña; la discrepancia dividida por sus grados de libertad (CMIN / DF [debe ser inferior a 3]); el Error Cuadrático Medio de Aproximación (RMSEA<0,08); Razón de Parsimonia (PRATIO); Índice de Bondad de Ajuste (GFI); Índice Ajustado de Bondad de Ajuste (AGFI). Los valores de estos índices deben ser cercanos a 0,90 para considerar un buen ajuste. Se calcularon el Criterio de Información (AIC) y el Índice Esperado de Validación Cruzada (ECVI), indicando valores bajos un mejor ajuste (Tabachnick y Fidell, 2013).

Tabla 2

Índices de bondad de ajuste del análisis factorial confirmatorio de los cuatro modelos

<i>Índice</i>	<i>CARMEN 1.0</i> <i>50 items</i> <i>6 factores</i> <i>Modelo 1</i>	<i>CARMEN 2.0</i> <i>35 items</i> <i>Unifactorial</i> <i>Modelo 2</i>	<i>CARMEN 2.0</i> <i>35 items</i> <i>4 factores</i> <i>Modelo 3</i>	<i>CARMEN 3.0</i> <i>22 items</i> <i>5 factores</i> <i>Modelo 4</i>
Coefficiente de	275,73	138,40	138,40	67,68
Mardia				
χ^2	3439,66 (p= ,000)	1266,49 (p = ,000)	1039,95 (p = ,000)	349,01 (p = ,000)
CMIN/DF	2,94	2,26	1,88	1,74
RMSEA	0,067	0,054	0,045	0,042
RMSEA (CI-90%)	0,065-0,070	0,050 – 0-058	0,041 – 0,050	0,034 – 0,049
PRATIO	0,95	0,94	0,93	0,87
AIC	3651,66	1406,49	1195,95	453,01
ECVI	8,55	3,29	2,80	1,06
ECVI (CI-90%)	8,15-8,97	3,06-3,54	2,60 – 3,02	0,95 -1,19
GFI	0,72	0,83	0,86	0,93
AGFI	0,70	0,81	0,84	0,91

Los resultados indican que aunque varios modelos pueden considerarse con buen ajuste, es el último el que mantiene indicadores muy superiores al resto. En la Figura 1 se muestra el resultado del AFC para este modelo.

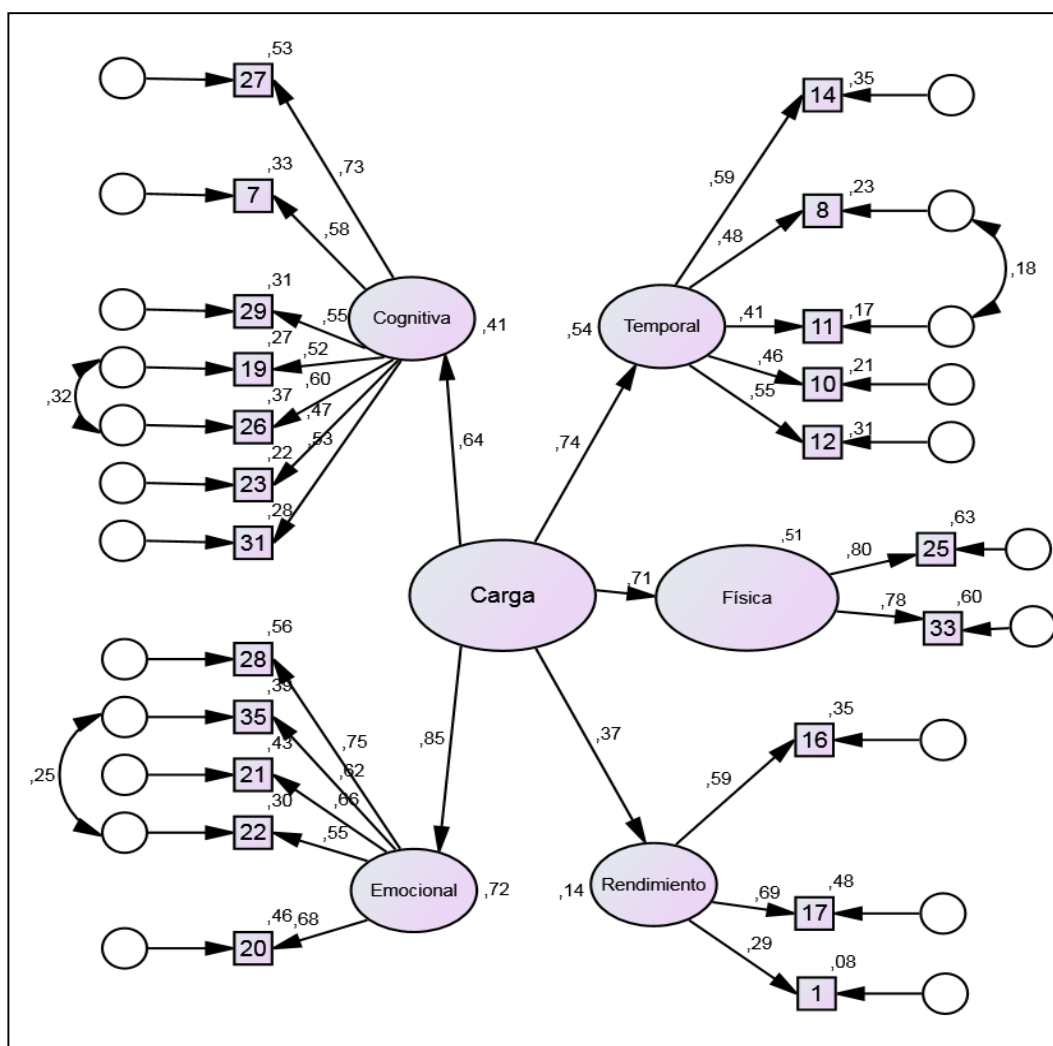


Figura 1. Modelo de cuatro factores de primer orden y uno de segundo orden (carga) de la versión de 22 ítems (CARMEN 3.0).

Para evaluar la consistencia interna, se calculó el coeficiente alfa de Cronbach en los dos modelos evaluados que mostraron mejores índices de ajuste en el AFC (ver Tabla 3).

Tabla 3
Descriptivo y fiabilidad de los modelos 3 y 4

Item	Media	DT	CARMEN 2.0		CARMEN 3.0	
			R ítem-total corregida	α si se elimina el elemento	R ítem- total corregida	α si se elimina el elemento
1	2,65	,49	,40	,906	,32	,849
2	2,36	,61	,43	,906		
3	2,27	,70	,54	,904		
4	1,82	,77	,56	,904		
5	2,33	,69	,43	,906		
6	1,83	,79	,58	,903		
7	2,38	,68	,53	,904	,46	,844
8	1,67	,79	,45	,905	,43	,845
9	1,43	,78	,52	,904		
10	1,61	,75	,39	,906	,38	,847
11	1,42	,73	,36	,907	,36	,848
12	1,88	,79	,45	,906	,43	,845
13	1,99	,67	,53	,904		
14	1,57	,73	,44	,906	,45	,844
15	1,26	,81	,55	,904		
16	2,53	,59	,31	,907	,25	,851
17	2,32	,62	,40	,906	,32	,849
18	2,31	,74	,43	,906		
19	1,86	,74	,52	,904	,48	,843
20	1,32	,70	,40	,906	,43	,845
21	1,09	,62	,44	,906	,48	,843
22	,75	,75	,35	,907	,39	,846
23	2,12	,63	,52	,905	,45	,844
24	2,40	,66	,45	,906		
25	1,13	,68	,32	,907	,36	,847
26	1,50	,77	,41	,906	,42	,845
27	2,11	,74	,55	,904	,50	,842
28	1,15	,69	,40	,906	,43	,845
29	1,65	,78	,50	,905	,48	,843
30	1,56	,80	,40	,906		
31	1,98	,78	,54	,904	,50	,842
32	1,71	,77	,45	,906		
33	1,27	,71	,33	,907	,36	,847
34	2,11	,73	,24	,908		
35	,77	,73	,37	,907	,40	,846

Para CARMEN 2.0 el α de Cronbach, fue: escala total ($\alpha = ,908$); demanda cognitiva ($\alpha = ,873$), demandas emocionales y físicas ($\alpha = ,826$), demandas temporales ($\alpha = ,801$), rendimiento ($\alpha = ,745$). En el modelo CARMEN 3.0 los valores de α fueron: escala total ($\alpha = ,851$), demandas cognitivas ($\alpha = ,813$), demandas emocionales ($\alpha =$

,788), demandas temporales ($\alpha = ,764$), demandas físicas ($\alpha = ,800$), rendimiento ($\alpha = ,671$). Para esta última dimensión si se elimina el ítem 1, α aumenta a ,700.

Conclusiones

En este estudio se han analizado las propiedades psicométricas de un cuestionario diseñado para medir la carga mental en el trabajo (Cuestionario CARMEN). Estas propiedades fueron evaluadas en una amplia muestra de trabajadores de diferentes sectores de actividad. En primer lugar se analizó la estructura factorial de cuatro modelos jerárquicos según varias versiones del cuestionario derivadas de la consideración de diferentes dimensiones y de una reducción significativa en el número de ítems. El último de los modelos fue el que mostró mejores indicadores de ajuste. Este modelo resultó conformado por cinco dimensiones y solamente 22 elementos.

La fiabilidad fue analizada como consistencia interna. Aunque los índices de fiabilidad fueron menores con la versión final CARMEN 3.0, en todos los casos se obtuvieron coeficientes α cercanos o muy superiores a 0,70, manteniendo por tanto índices de fiabilidad aceptables.

Entre las limitaciones del estudio cabe señalar la conveniencia de refutar esta estructura factorial en muestras de otros sectores profesionales diferentes a los incluidos en este trabajo, así como analizar la validez convergente con otras técnicas y la validez de criterio empleando indicadores externos de rendimiento y salud. Uno de los problemas más importantes asociados con la evaluación y medida de la carga mental es que todavía no se dispone de unidades de medida estandarizadas que permitan el diagnóstico de la situación (Annett, 2002). Esta falta de baremos dificulta, o hace prácticamente imposible, que los responsables de los Servicios de Prevención puedan diseñar estrategias de diagnóstico, prevención e intervención adecuadas y eficaces. Sin duda, la obtención de muestras más amplias nos permitirá la elaboración de baremos que sirvan para realizar una interpretación adecuada de las puntuaciones obtenidas. Todos estos análisis serán objeto de futuras investigaciones con el cuestionario CARMEN.

Referencias

- Annett, J. (2002). Subjective rating scales: science or art? *Ergonomics*, 45(14), 966-987.
- Bollen, K. A. (1989). *Structural equations with latent variables*. Nueva York: Wiley.
- Gil-Monte, P. R. (2009). Algunas razones para considerar los riesgos psicosociales en el trabajo y sus consecuencias en la salud pública. *Revista Española de Salud Pública*, 83(2), 169-173.
- 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.
- Rolo-González, G., Díaz, D., y Hernández-Fernaund, E. (2009). Desarrollo de una Escala Subjetiva de Carga Mental de Trabajo (ESCAM). *Revista de Psicología del Trabajo y de las Organizaciones*, 25(1), 29-37.
- Rubio, S., Diaz, E., Martin, J., y Puente, J. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology: An International Review*, 53(1), 61-86.
- Szalma, J. L. (2008). Individual differences in stress reaction. En P.A. Hancock y J.L. Szalma (Eds.), *Performance Under Stress* (pp. 323-357). Hampshire: Ashgate.
- Tabachnick, B. y Fidell, L. (2013). *Using multivariate statistics*. Boston: Pearson.