

# Utilidad de los criterios compuestos y múltiples mediante eficacia discriminativa en selección de personal

Mariano García Izquierdo y Antonio León García Izquierdo  
Universidad de Murcia

En el presente estudio se analiza la utilidad de las decisiones en selección de personal comparando el criterio múltiple y compuesto mediante la eficacia discriminativa, en cursos de formación ocupacional de albañilería. Se ha constatado la validez de todos los predictores (aptitudes, personalidad y experiencia laboral) con el criterio (ERPF), mediante regresión logística y múltiple. Los resultados muestran que no hay un criterio completamente mejor que otro, sino que según los fines que se pretendan y las características de la decisión que se quiera tomar es más aconsejable uno u otro.

*Composite and multiple criteria: utility analysis with discriminative efficiency in personnel selection.* The present study examined the utility of decision making in personnel selection comparing multiple and composite criteria by discriminative efficiency, in a bricklaying training program. We have found valid predictors (aptitudes, personality and experience) with different forms of the criterion, the ERPF scale, in logistic and multiple regression analysis. Results seem indicate that there is not a better criterion, so it depends on different conditions and different aims to reach. We proposed criteria as combining measures but not opposed.

El problema del criterio (Guion, 1997) sigue vigente y es uno de los más importantes con los que se encuentra la Psicología del Trabajo y de las Organizaciones (Murphy y Cleveland, 1995; Viswesvaran y Ones, 2000). El criterio en la selección de personal se corresponde con la medida del rendimiento de un trabajador y es precisamente la variable que se quiere pronosticar. La evaluación del rendimiento sigue siendo una de las funciones cruciales en la gestión de personas por la importancia de estas decisiones y sus consecuencias y, de hecho, cuenta con notables contribuciones recientes (Carless y Roberts-Thompson, 2001; Carretta y Ree, 2000; García, Barbero, Ávila y García, 2003; Escudero, Balagué y García-Mas, 2002; Ilgen y Pulakos, 1999; Nijenhuis y van Der Flier, 2000; Sackett, 1998; Schmidt y Hunter, 1998, 2002; Salgado, 1998, 2003a, 2003b; Steele-Johnson, Osburn y Pieper, 2000; Tziner, Murphy y Cleveland, 2001).

A pesar de su importancia, el rendimiento laboral no ha sido definido todavía con suficiente claridad; las razones son diversas, algunas de ellas las señalan Campbell, Blake y Oswald (1996) con cierto tono irónico: es muy difícil dar una definición útil, no es conveniente para la organización, y sabemos lo que es cuando lo tenemos delante y tenemos que enfrentarnos a ello; por lo tanto, no merece la pena hacer el esfuerzo. La principal duda que se ha venido planteando es si evaluar resultados o desempeños, polémica iniciada por Dunnette (1963) a la que desde entonces se han sumado diferentes autores como Blum y Naylor (1968), Campbell

(1970, 1990), Guion (1979), Latham (1984, 1986), Drenth y Algera (1988) y Schuler, Farr y Smith (1993).

El criterio debe cumplir una serie mínima de requisitos: que sea fiable, relevante, discriminante y práctico. Quizás, el principal inconveniente lo hallamos en la fiabilidad, debido a que se suele construir en función de los propósitos que se persiguen en cada organización. Incluso es un dato que no es frecuente encontrar (Schmitt, Gooding, Noe y Kirsch, 1984) debido a la dificultad de construir criterios que se puedan utilizar en diferentes contextos.

En la revisión bibliográfica se produce un notable desacuerdo en cuanto al tipo de criterio que se debe utilizar: criterio múltiple, compuesto o sencillo. El criterio múltiple es aquel que consiste en la medición de distintas variables para evaluar el rendimiento de un trabajador; es el preferido por Drenth y Algera (1988), ya que opinan que un criterio general del tipo «buen trabajador» o «mal trabajador» no refleja todas las conductas que se refieren a ese criterio. Como ya adelantase Toops (1944), las razones por las que un trabajador puede ser calificado como bueno son diversas, por lo que un criterio múltiple refleja la variedad de comportamientos de un individuo en el trabajo.

En esta línea, Muchinsky (1994) considera que el rendimiento es multidimensional, por lo que debemos incluir diversos aspectos pertinentes del trabajo cuando se elaboran criterios. Como argumenta Quijano (1993, p. 43), «el criterio múltiple agrupa conjuntos homogéneos de elementos y los correlaciona con predictores sueltos o conjuntos homogéneos de predictores, de forma que las puntuaciones de validación obtenidas son claramente interpretables, y pueden ser utilizadas de manera potencial en la formulación de una Teoría del Comportamiento en el trabajo».

Los autores que se decantan por un criterio múltiple basan su postura, principalmente, en dos puntos: el primero es que si las dimensiones del rendimiento en el trabajo son independientes,

no pueden sumarse para formar un compuesto (Ghiselli, 1956), excepto en el caso que exista una alta correlación entre ellos, lo que indica, en cierta medida, que existe unidimensionalidad en el criterio. Sin embargo, al agrupar todos los elementos del criterio en una puntuación se pierde la comprensión de los factores que contribuyen al éxito en el trabajo, ya que el desempeño es multidimensional e incluye numerosos aspectos (Muchinsky, 1994).

El criterio compuesto se suele elaborar midiendo por separado variables que podemos denominar criterios secundarios y combinarlos en un valor global. Éstos aparecen en numerosos estudios de validez, como los de Hunter y Schmidt (1990) y Murphy (1997), y pueden ser útiles aunque ofrezcan dificultades para la medición del rendimiento como, por ejemplo, la ponderación de las variables, donde no hay acuerdo en el método más adecuado. Las distintas alternativas las ha resumido Nagle (1953) en las siguientes: mediante el consenso de expertos, mediante la ponderación proporcional a la fiabilidad de cada criterio, mediante el cálculo de la proporción según la correlación media con otras variables relacionadas con el predictor, maximizando la predicción, eligiendo aquella con mayor correlación entre el predictor y el criterio compuesto según el análisis factorial, minimizando el coste (Brogden y Taylor, 1950), maximizando la diferencia entre individuos, mediante la evitación de la asignación de diferentes pesos o bien ponderando todos de igual modo y, finalmente, mediante la ponderación según el coste económico. Para Blum y Naylor (1968, p. 269), la ponderación ecuánime es la que menos errores comete «sobre todo cuando los ‘expertos’ disponibles para emitir juicios sobre la pertinencia no tienen la calidad deseable», lo que la hace apropiada para aquellos casos en que los evaluadores carecen de la formación suficiente sobre evaluación y medición en Psicología. Hecho que se da con frecuencia en el contexto aplicado de la selección de personal, donde, en muchos casos, la aplicación de la técnica queda relegada a prácticas propias del método intuitivo (Blasco, 1996). En definitiva, es el criterio más conservador el que menos riesgo conlleva.

Entre los defensores del uso del criterio compuesto encontramos, por ejemplo, a Campbell y otros (1993), Toops (1944) y Schmitt y Chan (1998), y entre los que defienden un criterio múltiple, por ejemplo, a Dunnette (1963b), Ghiselli (1956) y Guion (1961). Es decir, hallamos sólidos representantes de ambas tendencias, por lo que es difícil presentar la postura idónea.

Ahora bien, si lo que se pretende es disponer de una medida ágil que facilite la toma de decisiones, parece más apropiado un criterio único que ha sido el predominante en las investigaciones sobre selección de personal. Este uso viene avalado por su capacidad de explicar la mayor parte de las covarianzas de las puntuaciones verdaderas (Campbell, 1996). La base de este tipo de criterio descansa en la medida del éxito total de un trabajador en un puesto. Sin embargo, es claramente incompleto y menos versátil, ya que el criterio múltiple y el compuesto se pueden convertir con cierta facilidad en un criterio sencillo e incluirlo como ítem formando parte de un conjunto.

Dada esta situación, en el presente estudio queremos analizar la utilidad de los diferentes tipos de criterios comparando criterios múltiples y compuestos mediante regresión múltiple y regresión logística en un proceso de selección de personal mediante la eficacia discriminativa (Wiggins, 1973), técnica que nos permite clasificar los resultados obtenidos según el número de decisiones correctas. Para ello se comparan los valores predichos por las

ecuaciones de regresión en los criterios con los obtenidos expresando en porcentajes los éxitos y los fracasos.

## Método

### *Participantes*

La muestra inicial constaba de noventa alumnos de seis cursos de formación. Debido a diversos motivos (acceder a un puesto de trabajo, enfermedad o abandono), la muestra final quedó reducida a sesenta y seis alumnos, que fueron los que finalizaron las acciones formativas de albañilería pertenecientes a los planes de Formación e Inserción Profesional del INEM. La incorporación de los alumnos al curso se produce después de superar una entrevista semiestructurada con los responsables de la formación.

La media de edad es 23,3 años (desviación típica de 5 años y 2 meses, la moda es 18,6 años y un rango entre 16 y 37,4 años). Todos los sujetos son varones, el ochenta por ciento tiene una formación académica equivalente al graduado escolar y el resto posee un nivel superior al graduado escolar.

### *Variables e instrumentos*

Las variables aplicadas como predictores son las siguientes: inteligencia general, aptitud espacial, aptitud numérica, personalidad y los datos biográficos. Para medir la inteligencia general (factor g) se utilizó el test B-53 de Bonnardel en la adaptación española (Calles, 1975), cuya consistencia interna en nuestro estudio es de 0,88. La aptitud espacial se midió con los subtests 5 y 6 (factores 6,7 y 8; S<sub>1</sub> y S<sub>2</sub>) de la Batería Factorial de Aptitudes de Manzione en su adaptación española. El subtest S<sub>1</sub> que mide la rapidez perceptiva y rapidez de *gestalt* y está compuesto por dos pruebas: figuras idénticas y ladrillos-cubos. El factor S<sub>2</sub>, desplazamiento, mide la aptitud espacial para figuras rotadas. La consistencia interna obtenida es 0,80.

Utilizamos la versión Octagón de la serie OPQ (*Occupational Personality Questionnaire*) de Saville, Holdsworth, Nyfield, Cramp y Mabey (1985), en su adaptación española de 1989, que contiene ocho variables: «Autoafirmativo», «Empatía», «Gregario social», «Abstracto», «Metódico», «Relajado», «Autocontrolado» y «Vigoroso», y presentó una consistencia interna de 0,71. En todas las pruebas se respetaron los tiempos indicados en los manuales. Una descripción más detallada se puede encontrar en García-Izquierdo (2001) y en García-Izquierdo y García-Izquierdo (2002).

Los criterios se han construido con la escala que hemos diseñado y denominado «Evaluación del Rendimiento Profesional-Académico» (ERPF). Su contenido recoge una evaluación del rendimiento de los sujetos en el curso formativo. Para su elaboración partimos de entrevistas individuales y de grupo con los formadores, con el objetivo de obtener los aspectos más relevantes. Posteriormente, se redactaron los ítems y se comprobó la comprensión y diferenciación de los mismos. Se cumplimentó durante entrevistas individuales con cada formador. La escala final consta de catorce ítems, de tipo Likert, y que se valoran de 0 a 10, que tratan de la calidad de las tareas realizadas, rapidez en la ejecución de tareas, cuidado de herramientas, constancia en el trabajo, habilidad en el manejo de herramientas, atención durante las explicaciones, puntualidad, interés en aprender, autonomía en el trabajo, trato con compañeros, trato con superiores, cumplimiento de instrucciones y comprensión de instrucciones.

*Procedimiento*

El reclutamiento de los candidatos se realizó contando con la colaboración de la oficina principal del INEM en Oviedo. A los sujetos convocados se les presentó un vídeo donde se recogen diversos aspectos relacionados con la formación y con los objetivos de la organización, así como los diferentes cursos y sus características, apoyados en una descripción de las actividades más frecuentes en cada uno de ellos. Posteriormente, cumplimentaron una hoja de datos biográficos y, finalmente, fueron convocados a una entrevista semiestructurada cuya duración media fue de veinticinco minutos. Después, basándonos en estos datos, se tomó la decisión de aptitud para el acceso a la formación. Los cursos de albañilería tienen una duración de mil trescientas cincuenta horas, repartidas en cuarenta horas semanales durante nueve meses. En el caso que nos ocupa no hay vínculo laboral alguno entre el alumno y la organización, la actividad que desarrollan es similar a la que se realiza en el contexto laboral, y la única compensación económica que reciben son las ayudas para el desplazamiento hasta el centro y una subvención para el comedor.

Un mes antes del fin de la acción formativa se aplicaron los predictores (hoja de datos biográficos, las pruebas de aptitudes y la prueba de personalidad) debido a que al comienzo del noveno y último mes los alumnos disfrutaban de permisos para la búsqueda de empleo, por lo que, implícitamente, estamos incorporando el criterio de permanencia en el curso como indicador de éxito. El tiempo medio total empleado en cumplimentar los cuestionarios y en realizar las pruebas fue, aproximadamente, de dos horas. Se les explicó la importancia de realizar las pruebas, pues es información que maneja el Departamento de Recursos Humanos para la Bolsa de Trabajo del Centro de Formación, con el objetivo de que los sujetos colaborasen y estuvieran motivados para cumplimentarlas. Como criterio se aplicó la escala Evaluación del Rendimiento Profesional-Formativo (ERPF), cuya consistencia interna, según el coeficiente alfa de Cronbach, fue de 0,96. Una vez recogida toda esta información se procedió al análisis de los datos.

**Resultados**

Utilizamos la regresión logística aplicando el método paso a paso con el test de razón de máxima verosimilitud para la estima-

ción de parámetros. Como medida de ajuste se examina la significación estadística de la R<sup>2</sup> de Nagelkerke. También utilizamos la regresión múltiple mediante el método paso a paso, y como medida de ajuste el coeficiente de determinación (R<sup>2</sup>). Todos los análisis de datos se realizaron siguiendo el criterio *listwise* para el tratamiento de los datos perdidos.

Para elaborar los criterios realizamos un análisis de componentes principales de la escala ERPF y obtuvimos dos factores que explican una varianza total del 77,5%. Al primero lo denominamos «Rendimiento en las tareas»(RT), que explica el 68,5% de la varianza, y contiene los ítems «Calidad», «Rapidez», «Habilidad», «Autonomía», «Comprensión», «Interés», «Constancia», «Atención» y «Orden»; el segundo lo denominamos «Cumplimiento de pautas sociales» (CS), explica el 9% de la varianza, y está formado por los ítems «Compañeros», «Superiores» y «Puntualidad». Para asignar las variables a un componente u otro hemos seguido la recomendación que proponen Hair, Anderson, Tatham y Black (1999), según la cual para una potencia estadística del ochenta por ciento y un nivel de significación del 0,05 la carga factorial admisible debe estar comprendida entre 0,65 y 0,70 como mínimo, dado nuestro tamaño muestral. El criterio compuesto elaborado, denominado «Total», surge de la suma de todos los ítems de la escala ERPF.

Las ecuaciones de regresión múltiple (RM) y de regresión logística (RL) resultantes presentan ajustes del modelo adecuados que, ordenándolas de mayor a menor ajuste, nos pronostican el criterio múltiple «RT» y «CS» y el criterio compuesto «Total». Los predictores que aparecen en los dos modelos de regresión elaborados (RM y RL) sobre el criterio múltiple «RT» son: «Experiencia en el sector», «Inteligencia» y la variable de personalidad «Metódico». Respecto al criterio «CS», en la RM aparecen como predictores la variable «Autoafirmativo» y «Experiencia en otros sectores», ambos con signo negativo. Cuando se utiliza la RL se mantiene la variable «Autoafirmativo», aunque en este caso le acompaña en segundo lugar otra variable de personalidad, «Gregario social», también las dos con signo negativo. Es destacable que ninguno de los predictores que miden capacidad intelectual forma parte del modelo de predicción con este criterio.

En lo que al criterio compuesto se refiere, en la RM los predictores seleccionados son «Experiencia en el sector» y «Metódico», y mediante la RL aparece de nuevo la variable «Metódico» y «Ex-

*Tabla 1*  
Resumen de resultados

| Criterios       | Regresión múltiple        |                   |                | Regresión logística   |                           |                   |                 |  |
|-----------------|---------------------------|-------------------|----------------|---|---------------------------|-------------------|-----------------|--|
|                 | Clasificaciones correctas | Razón de eficacia | R <sup>2</sup> | Predictores   | Clasificaciones correctas | Razón de eficacia | R̄ <sup>2</sup> | Predictores  |
| Múltiple RT     | 76,38%                    | 81,35%            | 0,349          | «Experiencia sector»<br>«Inteligencia (g)»<br>«Metódico»    | 73,61%                    | 76,27%            | 0,392           | «Experiencia sector»<br>«Inteligencia (g)»<br>«Metódico» |
| Múltiple CS     | 79,16%                    | 93,22%            | 0,217          | «Autoafirmativo (-)»<br>«Experiencia en otros sectores (-)» | 76,38%                    | 77,42%            | 0,282           | «Autoafirmativo (-)»<br>«Gregario social (-)»            |
| Compuesto total | 88,13%                    | 90,27%            | 0,158          | «Experiencia sector»<br><br>«Metódico»                      | 77,27%                    | 77,96%            | 0,227           | «Experiencia en otros sectores (-)»<br>«Metódico»        |

n= 66

perencia en otros sectores», esta última con signo negativo. Cuando se utiliza la regresión múltiple encontramos un mejor ajuste del modelo para los criterios múltiples que para el criterio compuesto.

La mayor razón de eficacia la proporciona el criterio múltiple «RT», seguido del criterio compuesto y del criterio múltiple «Cumplimiento de pautas». En cuanto a las clasificaciones correctas, el porcentaje que proporciona el criterio compuesto supera a los dos criterios múltiples.

A modo de resumen, y a la vista de los resultados encontrados y que mostramos en la tabla 1, podemos decir que el criterio compuesto nos ayuda a tomar mejores decisiones que el criterio múltiple, ya que permite realizar más clasificaciones correctas (88,13% con el criterio «RT» en RM y 77,27% con el criterio «CS» en RL, frente a 76,38% con el criterio «RT» en RM, y 79,16% con el criterio «CS» en RL). Asimismo, resulta ser más eficaz (90,27% RM y 77,96% RL) que el criterio múltiple de rendimiento en las tareas utilizando tanto la regresión múltiple como la regresión logística (81,35% en RM y 76,27% en RL, respectivamente). Pero esto no ocurre con el criterio «CS» (93,22% RM y 77,42% RL) frente al criterio compuesto (90,27% RM y 77,96% RL), donde vemos que es más eficaz el criterio múltiple en RM (93,22% respecto a 90,27%), aunque algo menos eficaz en RL (77,42% respecto a 77,96%).

#### Discusión y conclusiones

En los modelos de predicción existe una gran similitud en cuanto a la proporción de varianza explicada tanto por la RM y la RL; y en lo que a los predictores se refiere, coinciden en casi todos los casos, excepto con el criterio «CS», donde en el modelo de RM aparece la variable «Experiencia en otros sectores» y en la RL la variable «Gregario social». Si bien era de esperar la capacidad predictiva de las aptitudes (Carreta y Ree, 2000; Salgado, 2003a; Schmidt y Hunter, 1998) y de la experiencia laboral (Quiñones, Ford y Teachout, 1995; Sonnentag, 2000), no lo es tanto la capacidad de las pruebas de personalidad, aunque cada vez pueda ser mayor (Blasco y Prieto, 1993). Destacamos la capacidad predictiva del rendimiento de «Metódico» (*Conscientiousness*), hallazgo que parece suficientemente consolidado (Tett, Jackson y Rothstein, 1991; Salgado, 1998, 2003b).

Respecto al criterio «CS», la asociación de la variable predictora «Gregario social» (faceta de extraversión), con signo negativo, es congruente con el modelo de Eysenck (introversión y mayor respuesta a la sanción) y también la asertividad medida con la variable «Autoafirmativo» estaría asociada a una mayor percepción de conflictividad, y, por lo tanto, reciben peores evaluaciones.

Respecto a «Experiencia en otros sectores» y su relación con el cumplimiento de pautas sociales, puede provenir de las dificultades de socialización laboral en un nuevo sector. Aquí se abre una línea de investigación interesante, ya que ésta es una de las áreas más deficientes en lo que a investigación se refiere.

Cuando analizamos las clasificaciones correctas y las razones de eficacia, con la RM se obtienen resultados más precisos, lo cual es de esperar dado que en el proceso de dicotomización del criterio en la RL se pierde información. De lo anterior se puede concluir que siempre que los datos lo permitan sería aconsejable elegir el modelo de RM antes que el de RL, aunque la RL produce coeficientes de determinación mayores, y, por lo tanto, mejores ajustes del modelo. Sin embargo, la RL puede ser muy adecuada en los casos en que además de pronosticar es conveniente elaborar

una clasificación basada en las características de los datos, es recomendable frente al análisis discriminante (Wilson y Hardgrave, 1995), y además parece ser más adecuada para pronosticar el rendimiento académico (García y Alvarado, 2000), aspecto que no hemos encontrado en nuestro estudio, hecho que se puede deber a las características de la muestra (homogeneidad y escaso tamaño), lo que sugiere la necesidad de ahondar en este punto.

De lo anterior parece desprenderse que, en principio, la opción que más acierto nos proporciona es el modelo compensatorio con una estrategia de regresión múltiple que, a modo de resumen, permite extraer las siguientes variantes relevantes para pronosticar el rendimiento en este contexto: inteligencia, experiencia en el sector, ausencia de experiencia en otros sectores y un estilo de personalidad poco asertivo y metódico. A lo que podríamos sumar la extraversión social blanda (Gregario social) que aporta la regresión logística, aunque quizá no sería necesario, ya que puede existir multicolinealidad con la baja asertividad. No obstante, ésta es una línea que parece oportuno seguir en un modelo posterior, para realizar entrevistas e incluso para elaborar un perfil del trabajador tipo del sector.

Hemos comparado los dos criterios múltiples obtenidos del análisis de componentes principales de la escala ERPF sobre rendimiento, «RT» y «CS», factores que han recibido un amplio contraste empírico en la bibliografía, en consonancia con los aportados por Katz y Kahn (1966), Hersey y Blanchard (1982) y Borman y Motowidlo (1993). Esto supone partir de una consideración multidimensional del comportamiento en el trabajo. En este sentido, una de las propuestas con mayor aceptación es la de Campbell, McCloy, Oppler y Sager (1993). Para estos autores el rendimiento es el desempeño, el comportamiento, porque está bajo el control del trabajador; mientras que la medida del rendimiento a través de los resultados que alcanza la persona se considera una medida muy contaminada porque depende de muchos factores que escapan a la esfera de control del evaluado. Sin embargo, los criterios compuestos son más adecuados que los criterios únicos en aquellos casos donde haya que tomar decisiones, aunque cuentan con el inconveniente de que es difícil comprender qué es lo que está midiendo (Schmidt y Kaplan, 1971).

Los criterios múltiples presentan numerosas ventajas una vez que se ha realizado el esfuerzo de su elaboración, ya que permiten posteriores tratamientos que los puedan convertir en criterios compuestos y tomar decisiones de manera más simple. Por lo tanto, a la vez que proporcionan información pormenorizada sobre aquellos aspectos relevantes, se pueden simplificar para un uso más ágil.

La mayoría de las propuestas respecto al criterio tienen en común la perspectiva de los intereses de la organización, ya que son los dirigentes quienes deciden cuándo un comportamiento en el trabajo es adecuado o no. Sin embargo, Schuler, Farr y Smith (1993) incorporan un nuevo punto de vista donde la satisfacción del empleado es también una medida del criterio, lo que abre un nuevo enfoque centrado en el interés para el trabajador que se puede complementar con el anterior. Ésta es una perspectiva que requiere más investigación, entre otros motivos por los cambios sociales y laborales. Desde esta perspectiva se puede ver un escenario de relaciones laborales más completo, donde se consideran los intereses de ambas partes y se pueden tomar decisiones más eficaces, imparciales y equánimes, al tiempo que se evitan prácticas discriminatorias y se respetan los derechos de los candidatos a un puesto de trabajo (García-Izquierdo, 2002).

## Referencias

- Blasco, R. (1996). Selección de personal. *Dossier de evaluación y selección de personal*. Barcelona: Universitat de Barcelona.
- Blasco, R. y Prieto, J.M. (1993). Validité et utilité de l'évaluation de la personnalité en gestion du personnel. *Revue Européenne de Psychologie Appliquée*, 43 (3), 231-239.
- Blum, M. y Naylor, J.C. (1968/1976). *Psicología industrial. Sus fundamentos teóricos y sociales*. México: Trillas.
- Borman, W.C. y Motowidlo, S.J. (1993). Expanding the criterion domain to include elements of contextual performance. En N. Schmitt y W.C. Borman (eds.): *Personnel selection in organizations* (pp. 71-98). San Francisco: Jossey-Bass.
- Brogden, H.E. y Taylor, E.K. (1950). The dollar-criterion applying the cost accounting concept to criterion construction. *Personnel Psychology*, 3, 133-153.
- Calles, A. (1975). B-53 Test de inteligencia no verbal. En J. García Yagüe, A. Cordero y A. Murga: *Tests empleados en España*. Madrid: Instituto Nacional de Psicología Aplicada y Orientación Profesional.
- Campbell, J.P., Dunnette, M.D., Lawler, E.E. y Weick, K.E. (1970). *Managerial behavior, performance and effectiveness*. Nueva York: McGraw-Hill.
- Campbell, J.P. (1990). Modelling the performance prediction problem in industrial and organizational psychology. En M.D. Dunnette y L.M. Hough (eds.): *Handbook of Industrial and Organizational Psychology* (pp. 687-732). Palo Alto: Jossey Bass.
- Campbell, J.P., Blake, M. y Oswald, F. (1996). The substantive nature of job performance variability. En K.R. Murphy (ed.): *Individual differences and behavior in organizations* (pp. 258- 299). San Francisco: Jossey Bass.
- Carless, S.A. y Robertson-Thompson, G.P. (2001). Self-ratings in training programs: an examination of level of performance and the effect of feedback. *International Journal of Selection and Assessment*, 9 (3), 217-225.
- Carretta, T.R. y Ree, M.J. (2000). General and specific cognitive and psychomotor abilities in personnel selection: the prediction of training and job performance. *IJSA*, 8 (4), 227-236.
- Drenth, P. y Algera, J. (1988). Personnel selection. En P. Warr (ed.): *Psychology at work* (pp. 113-400). Penguin. Londres.
- Dunnette, M.D. (1963). A note on the criterion. *Journal of Applied Psychology*, 47, 251-254.
- Escudero, J.T., Balagué, G. y García-Mas, A. (2002). Comportamientos que influyen en el rendimiento deportivo de jugadores de baloncesto desde el punto de vista de los entrenadores. *Psicothema*, 14 (1), 34-38.
- García, M.V., Alvarado, J. y Jiménez, A. (2000). La predicción del rendimiento académico: regresión lineal versus regresión logística. *Psicothema*, 12 (2), 248-252.
- García, M.A., Barbero, M.I., Ávila, I. y García, M.C. (2003). La motivación laboral de los jóvenes en su primer empleo. *Psicothema*, 15 (1), 109-113.
- García-Izquierdo, A.L. (2001). *Validación orientada al criterio de procedimientos de selección de personal de oficio para el pronóstico del rendimiento laboral y formativo en el sector de la construcción*. Universidad de Murcia. Tesis doctoral no publicada.
- García-Izquierdo, A.L. (2002). *Selección, discriminación e igualdad de oportunidades en el acceso al empleo*. Universidad de Oviedo. Tesina de licenciatura no publicada.
- García-Izquierdo, A.L. y García-Izquierdo, M. (2002). Pronóstico de la inserción laboral en el sector de la construcción. *Psicothema*, 14 (4), 726-732.
- Ghiselli, E.E. (1956). Dimensional problems of criteria. *Journal of Applied Psychology*, 40, 1-4.
- Guion, R.M. (1961). Criterion measurement and personnel judgments. *Personnel Psychology*, 14, 141-149.
- Guion, R.M. (1979). *Principles of work sample testing: III. Construction and evaluation of work sample tests*. TR- 79- A10. Alexandria: United States Army Research Institute for the Behavioral and Social Sciences.
- Guion, R.M. (1997). Criterion measures and the criterion dilemma. En N. Anderson y P. Herriot (eds.): *International Handbook of Selection and Assessment* (pp. 267-286). Chichester: John Wiley and Sons.
- Hair, J.F., Anderson, R.E., Tatham, R.L. y Black, W.C. (1999). *Análisis multivariante*. Madrid: Prentice-Hall.
- Hersey, P. y Blanchard, K. (1982). *Management of organizational behavior: utilizing human resources*. Nueva Jersey, Englewood Cliff: Prentice Hall.
- Hunter, J.E. y Schmidt, F.L. (1990). *Methods of meta-analysis*. Newbury Park, California: Sage.
- Ilgel, D.R. y Pulakos, E.D. (1999). *The changing nature of performance*. San Francisco: Jossey Bass.
- Katz, D. y Kahn, R.L. (1966/1978). *Psicología social de las organizaciones*. México: Trillas.
- Latham, G.P. (1984). The appraisal system as a strategic control. En C.J. Fombrun y N.M. Tichy (eds.): *Strategic human resource management* (pp. 87-100). Nueva York: Wiley.
- Latham, G.P. (1986). Job performance and appraisal. En C.L. Cooper e I.T. Robertson (eds.): *International Review of Industrial and Organizational Psychology* (pp. 117-155). Chichester: John Wiley and Sons.
- Muchinsky, M.P. (1994). *Psicología aplicada al trabajo*. Bilbao: Desclée de Brouwer.
- Murphy, K.R. (1997). Meta- analysis and validity generalization. En N. Anderson y P. Herriot (eds.): *International Handbook of Selection and Assessment* (pp. 323- 342). Chichester: John Wiley and Sons.
- Murphy, K.R. y Cleveland, J.N. (1995). *Understanding performance appraisal: social, organizational and goal-based perspectives*. Thousand Oaks: Sage.
- Nagle, B.F. (1953). Criterion development. *Personnel Psychology*, 6, 271-288.
- Nijenhuis, J.T. y Van Der Flier, H. (2000). Differential prediction of immigrant versus majority group training performance using cognitive ability and personality measures. *IJSA*, 8 (2), 54-60.
- Quijano, S.D. (1993). *La Psicología social en las organizaciones: fundamentos*. Barcelona: PPU.
- Quiñones, M., Ford, J.K. y Teachout, M.S. (1995). The relationship between work experience and job performance: a conceptual and meta-analytic review. *Personnel Psychology*, 48, 887- 910.
- Sackett, P. (1998). Performance assessment in educational and professional certifications: lessons for personnel selection? En M.D. Hakel (ed.): *Beyond Multiple Choice* (pp. 113-130). Mahwah: LEA.
- Salgado, J.F. (1998). Big five personality dimensions and job performance in army and civil occupations: a european perspective. *Human Performance*, 11 (2/3), 271-288.
- Salgado, J.F. (2003a). International validity generalization of GMA and cognitive abilities: a european community meta-analysis. *Personnel Psychology*, 56, 573-605.
- Salgado, J.F. (2003b). Predicting job performance using FFM and non-FFM personality measures. *Journal of Occupational and Organizational Psychology*, 76, 323-346.
- Saville, P., Holdsworth, R., Nyfield, G., Cramp, L. y Mabey, W. (1985). *Occupational Personality Questionnaire: Manual*. Esher- Surrey: SHL. Ltd.
- Schmidt, F.L. y Kaplan, L.B. (1971). Composite vs multiple criteria: a review and resolution of the controversy. *Personnel Psychology*, 24, 419-434.
- Schmidt, F.L. y Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: practical and theoretical implications of 85 of research findings. *Psychological Bulletin*, 124 (2), 262-274.
- Schmitt, N. y Chan, D. (1998). *Personnel selection. A theoretical approach*. Thousand Oaks: Sage.
- Schmitt, N., Gooding, R.Z., Noe, R.D. y Kirsch, M. (1984). Meta-analysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Schuler, H., Farr, J.L. y Smith, M. (1993). The individual and organizational sides of personnel selection and assessment. En H. Schuler, J.L. Farr y M. Smith. (eds.): *Personnel Selection and Placement* (pp. 1-9). New Jersey: Lawrence Erlbaum.
- Sonnentag, S. (2000). Expertise at work: experience and excellent performance. *International Review of Industrial and Organizational Psychology*, 15, 223- 264.
- Steele-Johnson, D., Osburn, H.G. y Pieper, K.F. (2000). A review and extension of current models of dynamic criteria. *IJSA*, 8 (3), 110-136.

- Tett, R., Jackson, D. y Rothstein, M. (1991). Personality measures as predictors of job performance: a meta-analytic review. *Personnel Psychology*, 44, 703-742.
- Tziner, A., Murphy, K.R. y Cleveland, J.N. (2001). Relationships between attitudes toward organizations and performance appraisal systems and rating behavior. *IJSA*, 9 (3), 226-239.
- Toops, H.A. (1944). The criterion. *Educational and Psychology Measurement*, 4, 271-297.
- Viswesvaran, C. y Ones, D.S. (2000). Perspectives on models of job performance. *IJSA*, 8 (4), 216-226.
- Wilson, R.L. y Hardgrave, B.C. (1995). Predicting graduate student success in an MBA program: regression versus classification. *Educational and Psychological Measurement*, 55, 186-195.
- Wiggins, J.S. (1973). *Personality and prediction*. Reading, Massachusetts: Addison-Wesley.