

Evaluación del Desempeño en la Administración Pública del Principado de Asturias: Análisis de las Propiedades Psicométricas

Performance Appraisal in the Public Administration of the Principality of Asturias: An Analysis of Psychometric Properties

Jesús F. Salgado

Universidad de Santiago de Compostela

Ángel L. Cabal

Gobierno del Principado de Asturias

Resumen. Este artículo presenta los resultados de tres estudios sobre el desarrollo y las propiedades psicométricas del método de evaluación del desempeño de los empleados de la Administración Pública del Principado de Asturias. En el estudio 1 se describe el desarrollo y la selección de las competencias a evaluar. Para cada una de las tres grandes dimensiones del desempeño laboral se identificaron cinco competencias que debían ser evaluadas. En el estudio 2 se describe cómo se elaboraron las escalas de valoración con anclajes conductuales para evaluar las competencias de desempeño de tarea, contextual y organizacional. En el estudio 3 se examinó la fiabilidad, la validez de constructo y la validez concurrente de las valoraciones del desempeño y se analizó el modelo mediante análisis factorial confirmatorio. Los resultados indican una elevada fiabilidad, tanto para las dimensiones como la evaluación global del desempeño. Un análisis de componentes principales indica la presencia de dos factores en los que cargan las valoraciones de los superiores inmediatos y los criterios objetivos externos respectivamente. El análisis factorial confirmatorio indicó que el desempeño puede representarse mediante un modelo jerárquico. Finalmente, los criterios objetivos externos correlacionan con las valoraciones de los superiores jerárquicos. Las implicaciones de los resultados se comentan en la discusión.

Palabras clave: desempeño, EBEP, evaluación, fiabilidad, validez.

Abstract. This article presents the results of three independent studies on the development and psychometric properties of the method of evaluating the performance of employees of Public Administration of the Principality of Asturias. Study 1 describes the development and selection of competencies to be assessed. For each of the three major dimensions of job performance, five competencies were identified that should be evaluated. Study 2 describes how we developed behaviorally-anchored rating scales for measuring competencies of task performance, contextual performance, and organizational performance. Study 3 examined reliability, construct validity and concurrent validity of performance appraisals. The results showed high reliability coefficients for both the dimensions and the overall performance assessment. An exploratory principal component analysis indicated the presence of two factors in which loaded (a) the evaluations of the immediate superiors and (b) external objective criteria, respectively. A confirmatory factor analysis showed that performance can be represented by a hierarchical model. Finally, external criteria and the ratings of managers showed to be correlated. The implications of the findings are commented on in the discussion.

Keywords: appraisal, EBEP, performance, reliability, validity.

La historia de la Psicología del Trabajo y de las Organizaciones, más de cien años, está plagada de casos exitosos de predicción del desempeño, la productividad y la eficiencia de los empleados tanto en organizaciones privadas como en organizaciones

públicas (Salgado, Anderson y Hülsheger, 2010) y dicha historia contiene también multitud de ejemplos de definición y medida exitosa de diferentes partes del desempeño, pero no fue hasta hace unos veinticinco años cuando los investigadores tuvieron conciencia de que sabían mucho sobre cómo predecir, pero apenas sabían qué predecían. Por ejemplo, la Society for Industrial and Organizational Psychology (SIOP), en la tercera edición de sus *Principles for the Validation and Use of Personnel Selection Procedures* (1987) definía el desempeño como “la efectividad y el valor de la conducta laboral y sus resultados”. Pero definido de este modo el desempeño era igual que la produc-

La investigación incluida en este artículo ha sido parcialmente financiada con cargo al proyecto EDU2008-03592. Los autores agradecen la colaboración de los miembros de la Dirección General de Planificación y Evaluación de Recursos Humanos. La correspondencia debe enviarse a Jesús F. Salgado, E.U. Relaciones Laborales, Universidad de Santiago de Compostela, Campus Vida, 15782 Santiago de Compostela. E-mail: jesus.salgado@usc.es

tividad o la efectividad. En otras palabras, no existían definiciones precisas de términos tales como desempeño, productividad y eficiencia. Un primer intento de separar estos tres conceptos fue hecho por Murphy (1990, p. 159), para quien productividad sería la razón entre las variables resultantes (de salida o outputs) y el coste de las variables invertidas (variables de entrada o inputs). También hizo Murphy una observación importante al señalar que muchos empleados usan su tiempo en actividades diferentes a ejecutar tareas relativas al puesto (p. ej. formación, atender a los compañeros o a gestionar conflictos o dificultades no previstas). En otras palabras, el tiempo de trabajo no está dedicado sólo a realizar tareas identificables en un análisis del puesto. Por tanto, evaluar sólo la ejecución de las tareas dejaría muchas conductas, acciones y actividades del empleado sin evaluar, conductas que, además, resultan importantes para sus organizaciones. Así, Murphy definió el desempeño como “*el conjunto de conductas que son relevantes para las metas de la organización o la unidad organizativa en la que la persona trabaja.*” (Murphy, 1990, p. 162). De entre tales metas, algunas están relacionadas con tareas específicas y otras no están relacionadas con tareas (p. ej., mantener buenas relaciones interpersonales con los compañeros y los superiores). Murphy (1990) desarrolló un marco que sugiere que la mayoría de los puestos de trabajo contienen cuatro tipos generales de conductas: (a) conductas orientadas a la tarea, (b) conductas orientadas interpersonalmente, (c) conductas relacionadas con la pérdida de tiempo (p. ej. absentismo, demoras) y (d) conductas destructivas y azarosas.

Esta concepción de Murphy fue ampliada y desarrollada en diversas propuestas que en los años noventa y en la década pasada han hecho relevantes investigadores. Una de las primeras fue la llevada a cabo por John P. Campbell y sus colegas (1990; Campbell, Gasser y Oswald, 1996; McCloy, Campbell y Cudeck, 1994). Campbell definió el desempeño como *cualquier conducta cognitiva, psicomotora, motora o interpersonal, bajo el control del individuo, graduable en términos de habilidad y relevante para las metas organizacionales.* A partir de una investigación en el ejército de USA, desarrolló un modelo factorial jerárquico compuesto de ocho factores, de los cuales el primero, el segundo, el tercero y el séptimo corresponden a las conductas necesarias para realizar las tareas técnicas del puesto de trabajo y los factores cuatro, cinco, seis y ocho corresponden a conductas que dan apoyo al ambiente social, psicológico y organizacional.

En paralelo a las propuestas definitorias del desempeño de tarea, diversos autores hicieron contribuciones en un aspecto hasta entonces pasado por alto: las conductas prosociales, conductas extra-rol y conductas que iban más allá por lo prescrito en las descripciones de los puestos de trabajo. En esta línea, Organ y sus colegas introdujeron la expresión “*conducta organizacional cívica*” para referirse a aquellas conductas que

representan lealtad, cooperación o ayuda y que van más allá de las obligaciones técnicas del trabajo (Bateman y Organ, 1983; Smith, Organ y Near, 1983). Este concepto fue ampliado por Borman y sus colaboradores, quienes han desarrollado un modelo de facetas del desempeño contextual y lo han redenido desempeño cívico (Borman 2004; Borman y Busch, 1993; Borman, Penner, Allen y Motowidlo, 2001; Motowidlo, Borman y Schmit, 1997). Campbell et al. (1996) explícitamente indicaron que los factores cuatro, cinco, seis y ocho de su modelo podían identificarse sin ninguna dificultad con las facetas del modelo de Borman de desempeño contextual o cívico. Por su parte, Dorsey, Cortina y Luchman (2010) han sugerido que si los sistemas de evaluación del desempeño no incluyen medidas de desempeño contextual o cívico, entonces tales sistemas estarían infrarepresentando el dominio del desempeño.

Así pues, actualmente hay un elevado acuerdo en que el dominio del desempeño del puesto incluye dos grandes dimensiones: desempeño de tarea y desempeño contextual o cívico. Aguinis (2007, p. 76) ha resumido las contribuciones previas, señalando: (1) que el desempeño es lo que hacen los empleados y sus conductas y no lo que el empleado produce o los resultados de su trabajo, aunque en ocasiones como las conductas o actividades no son fácilmente observables, es necesario inferirlas a partir de sus resultados; (2) también señala dos características del desempeño: Es evaluable y es multidimensional; y (3) siguiendo a investigadores previos, Aguinis (2007, p. 79) sostiene que hay al menos dos grandes dimensiones: desempeño de tarea y desempeño contextual, que deben considerarse separadamente porque no necesariamente ocurren en tándem.

No obstante, el resumen de Aguinis no considera el tercer y cuarto tipo de conductas señaladas por Murphy (1990), que han sido muy investigadas en la última década y han sido colectivamente denominadas conductas contraproductivas. La definición más utilizada de las mismas es la propuesta por Sackett y DeVore (2001), para quienes una conducta contraproductiva sería “*toda conducta intencional por parte de un miembro de la organización vista por esta como contraria a sus legítimos intereses.*” Por su parte, Spector y sus colegas (Fox y Spector, 2005; Rotundo y Spector, 2010) definen las conductas contraproductivas como cualquier conducta intencional, realizada por un empleado que daña o pretende dañar a la organización o a los miembros de la organización, lo que incluye a los empleados y a los clientes o usuarios. Por tanto, se excluyen las conductas que no están bajo control del empleado o son casuales. La investigación llevada a cabo por Robinson y Bennet (1995; Bennet y Robinson, 2000) ha sido de gran ayuda para establecer a su vez una clasificación de las conductas contraproductivas. De acuerdo a esta investigación, tales conductas se clasificarían en un espacio bidimensional, siendo una de las dimensiones el objeto de la conduc-

ta (organización o personas) y la otra dimensión la gravedad de la conducta (grave frente a leve). De este modo, habría dos tipos generales de conductas contra la producción y contra la organización y otras dos contra las personas. Este esquema clasificatorio ha sido reproducido en diversas investigaciones y, aunque no agota todas las posibles conductas desviadas, sí permite clasificar a la mayoría de ellas. Recientes investigaciones como la de Gruys (2000; Gruys y Sackett, 2003; Gruys, Stewart y Bowling, 2010) o la de Rotundo y Spector (2010) muestran una buena convergencia con el sistema clasificatorio de Robinson y Bennett y la necesidad de considerar las conductas contraproductivas como una entidad diferenciable del desempeño de tarea y del desempeño contextual.

Diversos meta-análisis han mostrado que los tres grandes espacios o dimensiones del desempeño del trabajo son entidades diferentes, aunque relacionadas entre sí. Agrupan conductas distintas, tienen implicaciones particulares para el comportamiento laboral y se predicen y explican por variables diferentes. Por ejemplo, el meta-análisis de Conway (1999) mostró que la correlación entre el desempeño de tarea y el desempeño contextual era de .49, y el meta-análisis de Dalal (2005) mostró que la correlación entre el desempeño contextual y las conductas contraproductivas oscilaban entre -.11 y -.33 según fueran dirigidas hacia las personas o hacia la organización. En cualquier caso, los resultados indicaron que son entidades diferentes. Además, la investigación también ha demostrado que el desempeño de tarea se explica mejor mediante capacidades cognitivas y conocimiento técnico, mientras que el desempeño cívico y las conductas contraproductivas se explican mejor mediante variables de personalidad y situacionales (Borman, Bryant y Dorio, 2010; Dalal, 2005; Dorsey, Cortina y Luchman, 2010; Rotundo y Spector, 2010; Salgado, Moscoso y Anderson, 2011). En definitiva, lo que ha demostrado la investigación es que el desempeño es multidimensional y tiene múltiples facetas, comúnmente agrupadas en torno a tres grandes dimensiones: (1) desempeño de tarea, (2) desempeño contextual (a veces llamada conductas prosociales o desempeño cívico) y (3) conductas desviadas o contraproductivas (conductas organizacionales en su polo positivo).

En paralelo con lo anterior, la evaluación del desempeño de los empleados es uno de los retos fundamentales con los que se enfrentan en esta década las Administraciones Públicas españolas, desde que la aprobación del *Estatuto Básico del Empleado Público* (EBEP) la ha convertido en una obligación legal. Por ello, las administraciones españolas, en mayor o menor plazo, tendrán que comenzar a realizarla. Además, las Administraciones se encuentran con un segundo problema que se deriva de la falta de tradición evaluativa del desempeño en la Administración española: la ausencia de herramientas y modelos de evaluación del desempeño específicos para las Administraciones Públicas.

El EBEP define la evaluación del desempeño como “*el procedimiento mediante el cual se mide y valora la conducta profesional y el rendimiento o el logro de resultados.*” (Art. 20.1) y señala que tendrá efectos “*en la carrera profesional horizontal, la formación, la provisión de puestos de trabajo y en la percepción de las retribuciones complementarias*” (Art. 20.3) y en la “*continuidad en un puesto de trabajo obtenido por concurso*” (Art. 20.4), y asimismo podría servir para la “*aplicación del régimen disciplinario*” (Art. 52). De todo lo anterior se desprende que la Administración entiende por desempeño la conducta profesional y el rendimiento o logro. En definitiva, en su definición de desempeño individual, la Administración está muy cerca de las definiciones más aceptadas por la Psicología del Trabajo y de las Organizaciones. Por ejemplo, es semejante a la definición de Campbell o a la propuesta de SIOP.

Ahora bien, la definición proporcionada en el Art. 20.1 no aclara a qué tipo específico de conducta profesional se está refiriendo. Para determinar esto último es preciso recurrir a otras partes del EBEP, entre las cuales el Capítulo VI, destinado a determinar los deberes de los empleados públicos y el código de conducta, es de especial relevancia para conceptualizar y delimitar las conductas y resultados que, además de las relativas a la tarea, conforman el desempeño del empleado público. Así, ya el comienzo del artículo 52 es relevante al decir que “*los empleados públicos deberán desempeñar con diligencia las tareas que tengan asignadas*” (Art. 52), lo que indica que el desempeño de tarea debe ser evaluado, pero también los artículos 53 y 54, sobre los principios éticos y el código de conducta, incluyen un número importante de clarificaciones. Una selección de los principios éticos relevantes a efectos de la evaluación del desempeño es la siguiente:

1. La actuación de los empleados públicos “*perseguirá la satisfacción de los intereses generales de los ciudadanos*” (Art. 53.2),
2. dicha actuación estará ajustada a “*los principios de lealtad y buena fe con la Administración en la que prestan servicios, y con sus superiores, compañeros, subordinados y con los ciudadanos*” (Art. 53.3),
3. y a los “*principios de eficacia, economía y eficiencia y vigilarán la consecución del interés general y el cumplimiento de los objetivos de la organización*” (Art. 53.8),
4. y “*cumplirán con diligencia las tareas que les corresponda o se les encomienden y, en su caso, resolverán dentro de plazo los procedimientos o expedientes de su competencia*” (Art. 53.10).

Por lo que se refiere a los Principios de Conducta, en nuestra opinión, los más relevantes en relación con la determinación del desempeño en el trabajo son los señalados seguidamente. Los empleados públicos:

1. “*Tratarán con atención y respeto a los ciudadanos, a sus superiores y a los restantes empleados públicos*” (Art. 54.1),

2. “El desempeño de las tareas correspondientes a su puesto de trabajo se realizará de forma diligente y cumpliendo la jornada y el horario establecidos” (Art.54.2),
2. “Administrarán los recursos y bienes públicos con austeridad. Tendrán, asimismo, el deber de velar por su conservación” (Art. 54.5),
4. “Mantendrán actualizada su formación y cualificación” (Art. 54.8),
5. “Observarán las normas sobre seguridad y salud laboral” (Art. 54.9),
6. “Pondrán en conocimiento de sus superiores o de los órganos competentes las propuestas que consideren adecuadas para mejorar el desarrollo de las funciones de la unidad en la que estén asignados” (Art. 54.10).

Además de las anteriores conductas, también son relevantes, a efectos de definir el espacio de desempeño laboral de los empleados públicos, las conductas tipificadas como faltas muy graves, entre las cuales hay tres que merecen especial consideración:

1. “El abandono del servicio, así como no hacerse cargo voluntariamente de las tareas o funciones que tienen encomendadas” (Art. 95.2c),
2. “El notorio incumplimiento de las funciones esenciales inherentes al puesto de trabajo o funciones encomendadas” (Art. 95.2g),
3. y “La desobediencia abierta a las órdenes o instrucciones de un superior” (Art. 95.2i).

En definitiva, el EBEP establece tres grandes ámbitos o tipologías de conductas relacionadas con las conductas laborales del empleado público, y aunque no las denomina así explícitamente, tales conductas se corresponden con las dimensiones de desempeño que la investigación ha etiquetado como desempeño de tarea, desempeño contextual y conductas contraproductivas (desempeño organizacional).

Con objeto de paliar la limitación antes mencionada, referida a la ausencia de modelos e instrumentos de evaluación del desempeño en la Administración Pública española, el gobierno del Principado de Asturias decidió elaborar un modelo de evaluación del desempeño y de seguimiento de la carrera profesional de sus empleados. En lo que sigue del artículo se dará cuenta del proceso de desarrollo de dicho modelo de evaluación de desempeño y de los resultados de los análisis efectuados para conocer sus propiedades psicométricas. A dicho efecto se presentan tres estudios independientes, en los que se describe cómo se identificaron las competencias a evaluar, cómo se desarrollaron las escalas evaluativas de dichas competencias y qué resultados mostraron los indicadores de las propiedades psicométricas (es decir, fiabilidad, validez de contenido, validez de constructo y validez de criterio).

Antes de finalizar esta introducción, debe recordarse que, en todo caso, el objetivo de una evaluación del desempeño con propósitos administrativos es, en últi-

ma instancia, producir una clasificación o ranking de los empleados en el conjunto de medidas de que se dota la organización para definir el desempeño. En este sentido, la resultante final será siempre una dimensión de ejecución en el trabajo, que indicará el grado de realización de las conductas y el logro de objetivos asignados al empleado. Este objetivo último no será logrado si el método de evaluación no posee las características psicométricas adecuadas.

ESTUDIO 1

Identificación y Estimación de la Fiabilidad de las Dimensiones de Evaluación del Desempeño de los Puestos Pertenecientes a la Administración Pública del Principado de Asturias (APPA)

Como indica el título de este primer estudio, su objetivo era identificar las dimensiones de desempeño de tarea, desempeño contextual y conductas organizacionales (conductas contraproductivas) que se van a medir en la APPA. Una vez identificadas las dimensiones a evaluar es preciso establecer el grado de fiabilidad del proceso de identificación, para comprobar si existe un buen acuerdo en la selección hecha.

Muestra

La muestra estuvo compuesta por 202 empleados de la administración del Principado de Asturias, desempeñando una variedad de puestos en diversos centros, instalaciones e instituciones. Por grupos se reparten del modo que figura en la Tabla 1.

Tabla 1. Distribución Porcentual de los Empleados según su Pertenencia Grupal

Grupo	Porcentaje
A1	77.22
A2	17.78
B	5.00

Por lo que respecta a su distribución según el nivel que poseen, en la Tabla 2 aparece como están repartidos.

Tabla 2. Distribución Porcentual de los Empleados según su Nivel

Nivel	Porcentaje
30	7.43
28	64.00
26	14.86
25 (o inferior)	13.71

Procedimiento

Los participantes respondieron a un cuestionario en el que estaban incluidas 50 competencias (facetas) correspondientes a las tres grandes dimensiones del desempeño en el trabajo, divididas en tres secciones. La primera incluía aquellas competencias del puesto (capacidades, conocimientos y habilidades) que se refieren al desempeño de tarea, totalizando 29. La segunda sección se refiera al desempeño contextual e incluye no sólo habilidades, sino también lo que podríamos llamar actitudes (p.e. orientación hacia la calidad, orientación al servicio del usuario/ciudadano, etc.), siendo un total de 12. La tercera sección incluía competencias correspondientes a la dimensión de comportamiento organizacional, con un total de 9 competencias. Con objeto de agotar al máximo las posibilidades de seleccionar las competencias apropiadas, en el cuestionario se incluyó también una cuarta sección de respuesta libre en la cual podían incluir posibles facetas relevantes no incluidas entre las anteriores.

El cuestionario iba precedido de unas instrucciones en las que, en síntesis, se le pedía a quien lo contestase que indicase, en una escala de 1 a 5 el grado de importancia o relevancia de cada faceta para los puestos de trabajo en cuestión. Los participantes contestaron dos veces el cuestionario. Una para los puestos de trabajo de los grupos A1 y A2 y otra para los grupos B, y C.

Análisis Estadístico y Psicométrico

Una vez obtenidas las respuestas a las dos versiones del cuestionario se procedió a realizar el análisis de los resultados en dos pasos. El primer paso consistió en identificar en cada una de los secciones cuáles son las dimensiones que resultaron con mayor relevancia, lo que se hizo usando dos criterios: (a) puntuación media más alta y (b) desviación típica más baja (lo que se interpretará como mayor acuerdo entre los encuestados). En este caso se seleccionaron 5 facetas por cada una de las grandes dimensiones del desempeño ocupacional. El segundo paso fue establecer la fiabilidad (consistencia interna) de cada una de las dimensiones del desempeño ocupacional, utilizando el coeficiente alfa de Cronbach. Estos dos pasos se hicieron doblemente, dado que se llevaron a cabo independientemente para los grupos A1 y A2 y B y C, respectivamente.

Resultados

Paso 1. Selección de Facetas de las Dimensiones del Desempeño Laboral.

Como se indicó más arriba, cada miembro de la muestra valoró cada una de las facetas del desempeño laboral utilizando una escala de 1 a 5, donde 1 signifi-

caba que la faceta no es nada relevante; 2 que es poco relevante; 3 que es algo relevante; 4 que es bastante relevante y 5 significa que es fundamental. Por tanto, cuanto más se aproxime la media a 5, mayor será el grado de relevancia que la faceta tiene para el desempeño. De acuerdo con los dos criterios señalados más arriba (puntuación media más alta y desviación típica más baja), las cinco facetas de cada una de las tres dimensiones del desempeño laboral que fueron seleccionadas para constituir el sistema de evaluación del desempeño son las que aparecen en la Tabla 3.

Tabla 3. Estadísticos Descriptivos de las Competencias Seleccionadas para la Evaluación del Desempeño

Grupos A1 y A2		
	Media	Desv.
Desempeño de Tarea		
Capacidad de toma de decisiones	4.10	0.85
Capacidad de organización y planificación	4.30	0.74
Capacidad de resolución de problemas	4.31	0.74
Conocimientos técnicos	4.44	0.72
Productividad (cantidad y calidad del trabajo)	4.23	0.77
Promedio	4.28	0.76
Desempeño Contextual		
Orientación a objetivos y resultados	4.28	0.83
Iniciativa	4.07	0.78
Compromiso con la organización	4.28	0.79
Colaboración y cooperación con compañeros	4.23	0.75
Compartir y transmitir conocimiento	4.02	0.77
Promedio	4.18	0.78
Desempeño Organizacional		
Mantenimiento voluntario del rendimiento laboral	4.51	0.82
Mantenimiento voluntario de la calidad del trabajo	4.52	0.80
Uso adecuado del tiempo y los recursos laborales	3.34	1.01
Asistencia al trabajo y uso eficiente del tiempo	3.96	0.95
Ritmo voluntario eficiente del trabajo	4.59	0.85
Promedio	4.18	0.89
Grupos B y C		
	Media	Desv.
Desempeño de Tarea		
Capacidad de aprendizaje	4.06	0.79
Mínuciosidad y responsabilidad	4.35	0.87
Esfuerzo y perseverancia	3.92	0.88
Conocimientos técnicos	4.09	0.79
Productividad (cantidad y calidad del trabajo)	4.35	0.73
Promedio	4.15	0.81
Desempeño Contextual		
Orientación a objetivos y resultados	3.68	1.04
Iniciativa	3.92	0.88
Compromiso con la organización	4.02	0.97
Colaboración y cooperación con compañeros	4.25	0.76
Compartir y transmitir conocimiento	3.60	1.05
Promedio	3.89	0.94
Desempeño Organizacional		
Mantenimiento voluntario del rendimiento laboral	4.47	0.85
Mantenimiento voluntario de la calidad del trabajo	4.46	0.85
Uso adecuado del tiempo y los recursos laborales	3.41	0.98
Asistencia al trabajo y uso eficiente del tiempo	4.03	0.90
Ritmo voluntario eficiente del trabajo	4.54	0.87
Promedio	4.18	0.89

Los resultados indican que las competencias seleccionadas para cada uno de los grupos tienen un promedio de relevancia en torno a 4 o superior, lo que claramente indica que son todas bastante importantes para el desempeño eficaz del puesto de trabajo. Al mismo tiempo, la desviación típica es en promedio inferior a 1 en todos los casos, lo que indica que los distintos encuestados están bastante de acuerdo entre sí. Queda patente también, como ya había sido demostrado previamente en la literatura científica, que las competencias de desempeño contextual y desempeño organizacional son comunes a todos los puestos de trabajo sin distinción, mientras que las facetas de desempeño de tarea, a excepción de conocimientos técnicos y productividad (cantidad y calidad del trabajo), son particulares para los distintos grupos laborales.

Complementariamente hemos examinado si las competencias seleccionadas aparecen reflejadas en el EBEP. Hemos constatado que además de las competencias desempeño de tarea, la competencia de orientación a objetivos y resultados figura en el artículo 53.8; la de iniciativa en los artículos 54.8 y 54.10; la de compromiso organizacional en el artículo 53.11; la de colaboración y cooperación con compañeros en el 53.3; la de compartir y transmitir conocimiento en el 54.8; la de mantenimiento voluntario del rendimiento laboral en el 95c; la del mantenimiento voluntario de calidad del trabajo en el artículo 95g; la dimensión de uso adecuado del tiempo y los recursos en los artículos 54.2 y 54.5; la asistencia al trabajo en los artículos 54.2 y 95c; y la dimensión de ritmo voluntario eficiente en los artículos 53.8 y 53.10.

Por tanto, las competencias indicadas por los encuestados como relevantes para la evaluación del desempeño están todas ellas mencionadas en el EBEP y acotan el dominio del desempeño en el trabajo en la Administración Pública del Principado de Asturias (APPA). Para comprobar más precisamente el grado de validez de contenido de las competencias seleccionadas, hemos calculado dos coeficientes de validez de contenido, el Índice de Validez de Contenido (IVC; Lawshe; 1975) y el coeficiente V (Aitken, 1980). Los resultados han sido estadísticamente significativos para ambos coeficientes. Así, el IVC resultó igual a .80 ($p < .05$) y el coeficiente V fue igual a .89 ($p < .01$). Debe recordarse que los valores de dichos coeficientes oscilan en -1 y +1, por lo que los resultados obtenidos son muy cercanos a los niveles óptimos de validez de contenido.

Paso 2. Análisis de la Fiabilidad de las Competencias del Desempeño Laboral

Una vez seleccionadas las competencias del desempeño laboral a evaluar, el siguiente paso fue establecer si la combinación de las mismas es fiable, para lo cuál hemos utilizado el coeficiente alfa de Cronbach. Los

resultados de los análisis de fiabilidad, para cada uno de los dos grupos en que se han clasificados los diversos puestos de trabajo existentes en la APPA, figuran en las Tablas 4 y 5 que aparecen a continuación.

Tabla 4. Competencias Seleccionadas para los Puestos de Trabajo de los Grupos A1 y A2 y Consistencia Interna (Coeficiente Alfa de Cronbach) de las Selecciones

Dimensiones de Tarea (Fiabilidad. *Coeficiente Alfa = .88*)

1. Capacidad de toma de decisiones
2. Capacidad de organización y planificación
3. Capacidad de resolución de problemas
4. Conocimientos técnicos
5. Productividad (cantidad y calidad del trabajo)

Dimensiones Contextuales (Fiabilidad. *Coeficiente Alfa = .75*)

6. Orientación a objetivos y resultados
7. Iniciativa
8. Compromiso con la organización
9. Colaboración y cooperación con compañeros
10. Compartir y transmitir conocimiento

Dimensiones Organizacionales (Fiabilidad. *Coeficiente Alfa = .74*)

11. Mantenimiento voluntario del rendimiento laboral
 12. Mantenimiento voluntario de la calidad del trabajo
 13. Uso adecuado del tiempo y los recursos laborales
 14. Asistencia al trabajo y uso eficiente del tiempo
 15. Ritmo voluntario eficiente del trabajo
-

Tabla 5. Competencias Seleccionadas para los Puestos de Trabajo de los Grupos B y C y Consistencia Interna (Coeficiente Alfa de Cronbach) de las Selecciones

Dimensiones de Tarea (Fiabilidad. *Coeficiente Alfa = .89*)

1. Capacidad de aprendizaje
2. Minuciosidad y responsabilidad
3. Esfuerzo y perseverancia
4. Conocimientos técnicos
5. Productividad (cantidad y calidad del trabajo)

Dimensiones Contextuales (Fiabilidad. *Coeficiente Alfa = .76*)

6. Orientación a objetivos y resultados
7. Iniciativa
8. Compromiso con la organización
9. Colaboración y cooperación con compañeros
10. Compartir y transmitir conocimiento

Dimensiones Organizacionales (Fiabilidad. *Coeficiente Alfa = .72*)

11. Mantenimiento voluntario del rendimiento laboral
 12. Mantenimiento voluntario de la calidad del trabajo
 13. Uso adecuado del tiempo y los recursos laborales
 14. Asistencia al trabajo y uso eficiente del tiempo
 15. Ritmo voluntario eficiente del trabajo
-

El resultado en las tres dimensiones para los grupos A1 y A2, como puede apreciarse en la Tabla 4, es altamente fiable y supera los límites recomendables para un trabajo aplicado y en el caso del desempeño de tarea el coeficiente alfa de Cronbach es notablemente alto (.88). Lo mismo ha ocurrido en el caso de las tres

dimensiones de desempeño ocupacional en los grupos B y C como puede verse en la Tabla 5. También en este caso, la fiabilidad para las dimensiones de desempeño de tarea resultó muy elevada (.89) y aceptablemente alta en el caso de las otras dos dimensiones, con valores de .76 y .72, respectivamente para desempeño contextual y desempeño organizacional.

Finalmente se ha examinado la fiabilidad global de las quince competencias del desempeño para los grupos A1 y A2 y para los grupos B y C. El resultado ha sido un coeficiente alfa de .78 en el primer caso y un coeficiente alfa de .79 en el segundo. Es decir, dos valores notablemente elevados.

ESTUDIO 2

Elaboración de escalas de valoración con anclajes conductuales (BARS) para las dimensiones seleccionadas

Una vez que mediante el Estudio 1 identificamos las competencias de desempeño ocupacional en las que deberían ser evaluados los diversos empleados de la APPA, este segundo estudio consistió en el desarrollo de los diferentes anclajes a utilizar en cada una de las escalas de valoración de las competencias y en el examen psicométrico de sus propiedades.

Es sabido que las BARS son un método de evaluación de conductas relacionadas con el trabajo, que se caracterizan, por un lado, por el uso de una graduación de niveles de eficiencia y, por otro, por una tipificación conductual de dichos niveles (Aamodt, 1999). La investigación ha demostrado que por su fiabilidad se convierten en uno de los tipos más utilizados en los ámbitos laborales con propósitos de evaluación del desempeño y para constatar la eficacia de las intervenciones organizacionales (p. ej., selección, formación, compensación, etc) (Aamodt, 1999; Bernardin, La Shells, et al., 1976; Bernardin y Smith, 1980; Smith y Kendall; 1963). Entre sus ventajas se encuentran la mejora del sistema evaluativo, la reducción del efecto halo, un feedback más efectivo de los valoradores a los evaluados y una mayor validez de contenido. Presentan el inconveniente de que la forma tradicional de elaboración requiere más tiempo y medios que otras escalas más convencionales. No obstante, a lo largo de los años se han sugerido sistemas alternativos de desarrollo que minimizan o reducen las dificultades de producción de las BARS. Por su ductilidad hace que sean utilizadas, además de en procesos de evaluación del desempeño, en otros procesos tales como las entrevistas de selección de personal o la evaluación académica de los estudiantes. El procedimiento típico para el desarrollo de las BARS es un análisis de puesto y entre las diversas técnicas, la de los incidentes críticos goza de preferencia porque permite identificar simultáneamente las dimensiones a evaluar y aporta ejemplos para la creación de los anclajes. No obstante, otras téc-

nicas, como la conferencia de expertos o los diarios, son utilizadas también para el desarrollo de las BARS, en especial cuando las dimensiones ya han sido identificadas.

En el presente caso, hemos utilizado un procedimiento abreviado para la confección de las BARS que se describe seguidamente. Para la elaboración de los anclajes se partió de un conjunto de incidentes críticos que comúnmente se producen en los puestos de trabajo de las Administraciones Públicas, y que habían sido reunidos con antelación. Dichos incidentes fueron examinados por un equipo de seis sujetos expertos en la materia y entrenados en análisis de puestos de trabajo, con objeto de que pudiera identificarse el tipo de conductas (criterio) que incluían y que permitiese señalar el grado de eficiencia de la misma. Por lo general, para cada competencia, los incidentes retenidos reflejaban de dos a cuatro tipos de conductas (criterios) que el empleado debería demostrar para que fuese valorado en un grado (nivel) determinado de desempeño en la competencia. Para unificar las distintas competencias, se decidió fijar en tres el número de conductas criterio que cada competencia permitía evaluar. A su vez, y dado que la APPA había decidido que las escalas de evaluación del desempeño tendrían cuatro grados, y después de trasladar conductualmente los criterios de cada dimensión, se adoptó la siguiente graduación. La puntuación más baja (1) sería para aquel caso que nunca o casi nunca era capaz de reunir los tres criterios. La siguiente puntuación (2) sería para aquel caso en el que normalmente se cumplen los criterios, pero no siempre todos y no efectivamente. La puntuación de 3 se asignaría a aquel caso en el que sobrepasase las conductas requeridas para el grado 2 pero no alcanzase el grado 4. En el caso concreto de la puntuación de 3 no se indicaban anclajes concretos por considerar suficiente la descripción conductual del grado 2 y del grado 4. Por último, el grado más elevado se asignaría a aquel caso en el que los tres criterios conductuales que definen la eficacia en la competencia se demuestran en la mayoría de las ocasiones. Siguiendo las recomendaciones de Smith y Kendall (1963), Bernardin, La Shells, et al. (1976), Bernardin y Smith (1980) y Aamodt (1999), las declaraciones incluidas en los anclajes son de tipo genérico en vez de ser específicos de un incidente determinado, por que tal incidente podría no haber ocurrido y, sin embargo, la conducta a evaluar se produzca con frecuencia.

El equipo de seis técnicos en recursos humanos, identificó independientemente cada uno de los criterios a incluir en la BARS, a partir de la lectura de los correspondientes incidentes críticos. Seguidamente cada técnico los clasificó en tres grupos: (a) fundamentales para definir la competencia conductual; (b) importantes pero no fundamentales; (c) prescindibles. Una vez realizada la anterior clasificación se procedió a calcular el grado de acuerdo entre los técnicos en cada uno de los incidentes y criterios. Aquellos que alcanza-

ron un acuerdo de más del 80% y que pertenecían al grupo de los fundamentales fueron los retenidos para escribir los anclajes.

El siguiente paso fue retrasladar los incidentes en términos generales y con los criterios mencionados a anclajes de 1, 2 y 4 y escribir las descripciones correspondientes. Una vez establecida la definición de la competencia, los criterios que la componen y los anclajes conductuales, las escalas resultantes fueron enviadas a la Dirección general de Evaluación de la APPA para su aprobación final. El Director General y la Jefe de Servicio de Evaluación, con varios técnicos de la Dirección examinaron el contenido de todas las BARS y realizaron la inspección final. Una vez aceptadas, cada una de las BARS se escribió en una ficha específica para ser incluida posteriormente en el manual del evaluador.

ESTUDIO 3

Análisis de las características psicométricas de la evaluación del desempeño de los empleados de la APPA

El tercer estudio llevado a cabo tuvo como objetivo establecer las propiedades psicométricas del procedimiento de evaluación del desempeño realizado, teniendo como metas la estimación de la fiabilidad, el examen de la estructura interna, la correlación con criterios externos de desempeño y la validación de constructo del modelo mediante un análisis factorial confirmatorio. En las páginas que siguen se describen el desarrollo y los resultados de dicho estudio.

Muestra

La muestra estuvo compuesta por 4422 empleados de los distintos grupos y escalas del personal funcionario que prestan sus servicios en la APPA. Estaba repartida del siguiente modo: Grupo A1 = 701; A2 = 735; C1 = 652; C2 = 1795; E = 539.

Procedimiento

Evaluación del Desempeño del Empleado

El desempeño de cada empleado durante el periodo de doce meses anterior a la fecha de la evaluación fue realizado por su jefe inmediato (Responsable de UGE). Para hacer la valoración del desempeño se utilizaron las 15 escalas de valoración con anclajes conductuales descritas anteriormente, correspondiendo cinco de ellas a competencias del desempeño de tarea, cinco a competencias de desempeño contextual y cinco a competencias de desempeño organizacional (ausencia de conductas contraproductivas). Las escalas utilizadas fueron las siguientes:

Desempeño de Tarea: En el caso de los puestos de los grupos A y B se utilizaron las siguientes escalas: capacidad de toma de decisiones, capacidad de organización y planificación, capacidad de resolución de problemas, conocimientos técnicos y productividad (cantidad y calidad del trabajo). En el caso de los grupos C1, C2 y agrupaciones profesionales se utilizaron las siguientes escalas: capacidad de aprendizaje, minuciosidad, esfuerzo y perseverancia, conocimientos técnicos y productividad (cantidad y calidad del trabajo).

Desempeño Contextual: Fue medido en todos los grupos con las mismas cinco escalas, que fueron las siguientes: Orientación a objetivos y resultados, iniciativa, compromiso organizacional, colaboración y cooperación, compartir y transmitir conocimiento.

Desempeño Organizacional: Fue medido en todos los grupos de la misma manera, mediante las siguientes escalas: Mantenimiento voluntario del rendimiento laboral, mantenimiento voluntario de la calidad del trabajo, uso adecuado de los recursos laborales, asistencia al trabajo y uso eficiente del tiempo, ritmo voluntario eficiente del trabajo.

Cada dimensión podía ser evaluada con una puntuación de 1 a 4, donde el 1 indicaba que el desempeño era inferior al promedio de dicha dimensión, el 2 era el promedio, 3 superior al promedio pero sin llegar a excelente y 4 un desempeño excelente. De este modo, voluntariamente la Administración Pública de Asturias ha introducido un marco valorativo que tiende a sesgar las puntuaciones hacia arriba (positivamente), al no estar balanceadas las posibilidades de desempeño superior al promedio e inferior al promedio.

Criterios Externos

Además de las variables indicadas más arriba, en el presente estudio se incluyeron, como elementos complementarios, la puntuación de varios factores de diversa naturaleza. Tales factores incluían el logro de los objetivos individuales asignados, la permanencia en el puesto durante el periodo objeto de evaluación (desempeño de puestos) y la asistencia a cursos formativos. También se incluían dos factores más, que no siempre están bajo el control de todos los empleados, denominados logro de objetivos colectivos y colaboración en la docencia. Una última variable que se ha registrado de cada empleado es su antigüedad en la Administración Pública. A este conjunto de factores los hemos denominado criterios externos. Seguidamente se ofrece una descripción de cada uno de ellos y de cómo fue medido.

Antigüedad: Permanencia en la Administración Pública, contabilizada como el número de años como empleado público, según los registros informáticos del Principado.

Objetivos Individuales: A todos los empleados de la Administración del Principado de Asturias se le asignó

el mismo objetivo personal en esta primera evaluación: proporcionar una descripción detallada de su puesto de trabajo, usando las herramientas informáticas disponibles y la asesoría de la Dirección Personal, si lo precisaban. A las personas que realizaron su objetivo personal se les asignó una puntuación de 5 puntos y a los que no lo hicieron 0.

Desempeño del Puesto: Permanencia en el puesto durante el periodo anual correspondiente a la evaluación del desempeño, corregido a la baja en función de los días de absentismo por IT (enfermedad común). Se puntúa con 0.5 puntos por mes, con un máximo de 6 puntos por año. En el caso de los responsables de UGE se pondera por un factor corrector que va de 1.2 a 1.5, por lo que en estos casos, el rango posible de valores máximos podría ser de 7.2 a, 8.1 ó 9 puntos, dependiendo del número de trabajadores que formen parte de la UGE (hasta 24, hasta 49 ó más de 50). La puntuación se minora dividiendo el número de días de baja entre 30 y multiplicado por un coeficiente de 0.75. Así, si una persona ha faltado 10 días en ese año, se descontará $(10/30) \times 0.75$, o lo que es lo mismo se restarán 0.25, lo que daría lugar a una puntuación de 5.75 para el factor en un puesto que no tenga dirección de UGE. En definitiva, el factor está evaluando permanencia y asistencia simultáneamente. Se valora a través del sistema informático. Esta variable sería realmente una medida del absentismo registrado.

Formación: Asistencia a actividades formativas. Se valora con 0.20 puntos por hora, cuando tenga un certificado de aprovechamiento y 0.15 sin certificado, con un máximo de 40 horas año. Se valora a través del sistema informático.

Resultados

Análisis de la Fiabilidad (Alfa de Cronbach) de las Evaluaciones del Desempeño de los Responsables de UGE

La evaluación del desempeño no tiene valor ni eficacia alguna si no es fiable, es decir, si el error de medición que se comete en toda evaluación no está dentro de los límites de lo aceptable en los estándares científicos. En el caso de las valoraciones hechas por los responsables de las UGE hay tres estimadores de la fiabilidad que son susceptibles, a priori, de ser calculados: la consistencia interna, la estabilidad temporal y el acuerdo entre valoradores. La estimación de la estabilidad temporal de las medidas (test-retest) requiere que se efectúen dos evaluaciones por los mismos valoradores separadas en el tiempo (normalmente varios meses). Esto no es posible en este momento en la Administración Pública y tampoco serviría correlacionar las medidas de un año con las del año siguiente, porque no se estarían refiriendo exactamente a las mis-

mas conductas, por lo que no podrían considerarse medidas paralelas en sentido estricto. Por lo que se refiere a la estimación del acuerdo entre valoradores, siempre deseable, no suele ser posible porque muchos empleados sólo tienen un supervisor directo y, por tanto, faltaría la segunda fuente de datos. Por estas razones, la forma más habitual de estimar la fiabilidad de la evaluación del desempeño es a través de la consistencia interna, utilizando para ello el coeficiente alfa de Cronbach. Este coeficiente nos informa sobre el grado de consistencia y homogeneidad de los distintos elementos (items) que componen una medida. Como ya se ha señalado en páginas precedentes, el coeficiente alfa es el estimador más utilizado de la fiabilidad en la investigación científica y hay bastante acuerdo en que, con propósitos de investigación, dicho coeficiente debería poseer al menos un valor de .70 o superior (Nunnally, 1978). Con propósitos de toma de decisiones administrativas, es recomendable que el valor alfa sea superior a .80.

En el presente caso hemos calculado el coeficiente alfa para cada una de las dimensiones del desempeño (tarea, contextual y organizacional) y también para la medida agregada del desempeño (desempeño global) y lo hemos hecho tanto para cada grupo de empleados como para el total de la muestra. Así pues, tenemos 4 estimadores por grupo, por cinco grupos, más cuatro estimadores para la muestra total. En definitiva, hemos calculado 24 coeficientes alfa de Cronbach, cuyos valores figuran en la Tabla 6. Dado que los tamaños de las muestras son grandes en todos los casos, podemos tener bastante confianza en que el error de muestreo apenas tiene efectos sobre dichos coeficientes.

Tabla 6. Fiabilidad (Alfa de Cronbach) de los Componentes de la Evaluación del Desempeño y del Desempeño Global en cada Grupo Ocupacional

GRUPO	DT	DC	DO	DG
A1 (n = 703)	.823	.835	.863	.939
A2 (n = 735)	.852	.843	.883	.947
C1 (n = 652)	.858	.891	.897	.957
C2 (n = 1796)	.879	.879	.878	.956
E (n = 539)	.840	.853	.872	.946
Total (n = 4424)	.864	.872	.883	.953

Nota: DT = Desempeño de Tarea; DC = Desempeño Contextual; DO = Desempeño Organizacional; DG = Desempeño Global.

Como puede apreciarse en la Tabla 6, los coeficientes alfa para las tres dimensiones del desempeño y en todos los grupos laborales son muy superiores a .80, oscilando entre .823 y .897, por lo que podemos concluir, sin ningún género de dudas, que estamos ante medidas muy fiables de las tres grandes dimen-

siones del desempeño laboral. Por lo que se refiere a la medida de desempeño global, la fiabilidad obtenida en los cinco grupos es extraordinariamente elevada, oscilando entre .939 en el grupo A1 y .957 en el grupo C1. Puesto que todos los coeficientes han sido muy elevados y superiores a los requisitos más exigentes, el mejor indicador de la fiabilidad de la medida de la evaluación del desempeño es el que se obtiene con la muestra total. En este caso, los coeficientes son todos ellos muy elevados respaldando la conclusión de que nos encontramos ante una medición del desempeño extraordinariamente fiable. Así los coeficientes alfa de Cronbach son .864, .872 y .883 para la medida de desempeño de tarea, la medida de desempeño contextual y la medida de desempeño organizacional, respectivamente. Dado que, en última instancia, todas las medidas del desempeño se agregan para producir una evaluación del desempeño en el trabajo en los empleados de la APPA, la mejor estimación es la que se obtiene para el desempeño global con la totalidad de la muestra. El resultado es un coeficiente alfa de .953, lo que indica una total fiabilidad en las valoraciones de los responsables de las diferentes UGE.

Para finalizar este apartado, queremos introducir una nota de cautela. Si bien los estimadores de la fiabilidad son excelentes y la conclusión referida a la fiabilidad de la estimación del desempeño no cambia en absoluto, es preciso también tener en cuenta que no se puede descartar que haya existido también un cierto grado de error halo en las valoraciones, que a su vez haya contribuido a que los coeficientes sean más altos de lo que en realidad serían sin dicho error halo. A pesar de todo, como señala Murphy (1982, Murphy, Jako y Anhalt, 1993), el error halo no siempre limita la precisión de las evaluaciones, sino que la puede aumentar, por ejemplo, en los procesos de evaluación para decisiones de promoción, ascenso o retribución, ya que el error halo no tiene efectos sobre la clasificación de los individuos, que es lo que se persigue en estos casos.

Análisis de Componentes Principales de las seis variables de desempeño individual bajo control del empleado

Seguidamente para examinar la estructura empírica de las variables, hemos realizado un análisis de componentes principales en el que incluimos las tres dimensiones de desempeño laboral y los tres criterios externos bajo control del empleado (desempeño de puestos, formación y objetivos individuales). El análisis que se presenta se refiere a la totalidad de la muestra (N = 4424), si bien los análisis realizados con los grupos específicos producen resultados semejantes. En la Tabla 7 se muestran los autovalores de los componentes extraídos.

Tabla 7. Autovalores de los Componentes Extraídos

Componente	Autovalores Iniciales		
	Total	% Varianza	% Acumulado
1	2.831	47.18	47.18
2	1.139	18.98	66.16
3	0.917	15.28	
4	0.847	14.12	
5	0.142	2.36	
6	0.124	2.08	

Una vez extraídos los componentes, se debe tomar la decisión de cuántos factores se retienen para su rotación posterior. Par tomar dicha decisión, en el presente caso se tuvieron en cuenta tres criterios: Kaiser, scree test y análisis paralelo. Tanto el criterio de Kaiser (autovalor > 1), como el Scree Test, como el análisis paralelo sugieren que sólo se deben retener y rotar los dos primeros componentes. Los dos componentes retenidos explican el 66.16% de la varianza total de las variables. El método de rotación seleccionado ha sido el Varimax de Kaiser por considerar que los dos componentes son variables ortogonales (independientes). Una vez rotados los componentes y como puede apreciarse en la Tabla 8, en el primer componente cargan con pesos significativos las tres dimensiones de desempeño evaluadas por los responsables de UGE, mientras que las otras tres medidas tienen pesos totalmente despreciables. En el segundo componente ocurre exactamente lo opuesto, son las tres medidas externas las que tienen pesos relevantes, mientras que las valoraciones de los responsables muestran pesos cercanos a cero. Así pues, estamos ante dos componentes claramente distintos del desempeño del puesto y totalmente independientes entre sí. El primero hace referencia a las tres dimensiones del desempeño del puesto, por lo que podemos denominarlo evaluación del desempeño, mientras que el segundo componente se refiere a los criterios externos de éxito en el empleo utilizados en el procedimiento de seguimiento de carrera.

Tabla 8. Matriz de Componentes Principales Rotados (Rotación Varimax)

Variable	Componente 1	Componente 2
Desempeño de Puestos	.142	.573
Formación	.017	.722
Objetivos Individuales	.048	.606
Desempeño de Tarea	.952	.098
Desempeño Contextual	.947	.116
Desempeño Organizacional	.946	.094

Relaciones entre las Dimensiones de Evaluación del Desempeño y los Criterios Externos de Ejecución Laboral

Otra evidencia complementaria para establecer la validez de constructo de la evaluación del desempeño efectuada en la APPA ha consistido en analizar la correlación existente entre las tres dimensiones del desempeño laboral, la medida global del desempeño y

Tabla 9. Correlaciones entre Evaluaciones del Desempeño y Criterios Objetivos Externos

Variables	DT	DC	DO	DG
<i>Grupo Profesional A1 (n = 696)</i>				
Objetivos Grupales	-.004	.040	.049	.029
Objetivos Individuales	-.026	.006	-.014	-.012
Desempeño de Puestos	.199**	.182**	.177**	.195**
Formación	-.027	.017	.027	.007
Docencia	-.056	.062	.041	.056
Antigüedad (n = 701)	-.082*	-.125**	-.098**	-.108**
<i>Grupo Profesional A2 (n = 733)</i>				
Objetivos Grupales	.081*	.057	.050	.066
Objetivos Individuales	.133*	.154**	.154**	.155**
Desempeño de Puestos	.073*	.128**	.095*	.105**
Formación	.070	.100**	.115**	.101**
Docencia	.067	.078*	.069	.076*
Antigüedad (n = 641)	.023	-.019	-.013	-.004
<i>Grupo Profesional C1 (n = 611)</i>				
Objetivos Grupales	-.119**	-.151**	-.124**	-.138**
Objetivos Individuales	.372**	.406**	.361**	.398**
Desempeño de Puestos	.107**	.110**	.089**	.106**
Formación	.197**	.223**	.186**	.212**
Docencia	.117**	.130**	.132**	.132**
Antigüedad (n = 652)	.023	.042	.054	.042
<i>Grupo Profesional C2 (n=1756)</i>				
Objetivos Grupales	.027	.043	.072**	.050*
Objetivos Individuales	.127**	.129**	.119**	.422**
Desempeño de Puestos	.139**	.141**	.135**	.144**
Formación	.077**	.071**	.054*	.070**
Docencia	.071**	.076**	.010	.054*
Antigüedad (n = 1795)	-.009	.004	-.012	-.006
<i>Grupo Profesional E (n = 533)</i>				
Objetivos Grupales	.021	.070	.027	.042
Objetivos Individuales	.160**	.177**	.126*	.162**
Desempeño de Puestos	.194**	.197**	.190**	.204**
Formación	.085*	.064	.064	.075
Docencia	-.008	-.037	-.012	-.020
Antigüedad (n = 539)	-.070	-.025	-.061	-.055
<i>Total (n = 4329)</i>				
Objetivos Grupales	.001	.009	.021	.010
Objetivos Individuales	.160**	.178**	.156**	.170**
Desempeño de Puestos	.167**	.175**	.162**	.175**
Formación	.103**	.113**	.102**	.112**
Docencia	.083**	.093**	.074**	.088**
Antigüedad	-.011	-.006	-.016	-.011

Nota: * $p < .05$; ** $p < 0.01$ (bilateral). Con excepción de Antigüedad, todas las variables han sido transformadas a puntuaciones estándar. DT = Desempeño de Tarea; DC = Desempeño Contextual; DO = Desempeño Organizacional; DG = Desempeño Global

los criterios externos, tanto los que están directamente bajo control del empleado como los que están menos controlados. El análisis se ha hecho para cada grupo laboral separadamente, porque las distribuciones en los criterios externos no necesariamente tienen que ser semejantes. Los resultados aparecen en la Tabla 9.

El examen de las correlaciones muestra que son todas ellas, en general, bajas lo que indica que la evaluación del desempeño es relativamente independiente de cada uno de los criterios externos de evaluación de los empleados públicos y que dichos criterios son una aproximación escasa del conjunto de conductas laborales realizadas por los empleados públicos en sus puestos de trabajo. Por las importantes consecuencias de orden práctico que tiene, debe destacarse la nula correlación entre la evaluación del desempeño y la antigüedad en la Administración Pública, que es prácticamente cero en todos los grupos. Este resultado coincide exactamente con el de estudios meta-analíticos en los que se ha examinado la relación entre desempeño y edad. Por ejemplo, Ng y Feldman (2010) encontraron que la correlación entre antigüedad y desempeño es de .10. En tanto que la antigüedad es directamente dependiente de la edad, son también de interés los meta-análisis de Waldman y Avolio (1986), McEvoy y Cascio (1989) y, más recientemente, Hardigree (2006), quienes encontraron correlaciones de -.14, .03 y .09, respectivamente entre edad y desempeño.

Una última evidencia para establecer la validez de constructo de la evaluación del desempeño efectuada en la APPA ha consistido en analizar la correlación existente entre las tres dimensiones del desempeño, la medida global del desempeño y el criterio externo compuesto, es decir, la suma estandarizada de logro de objetivos individuales, más la asistencia a cursos formativos, más el desempeño de puestos de trabajo (permanencia+asistencia). La razón de estandarizar las puntuaciones del criterio externo se debe a que están medidas con escalas diferentes y, por tanto, a efectos de eliminar contaminación en la puntuación compuesta resultante, la transformación a puntuaciones z hace que se puedan agregar usando la misma escala. Debe recordarse que son dos fuentes informativas totalmente independientes, una procedente de las valoraciones del responsable inmediato y la segunda procedente de los archivos de la Administración Pública del Principado. No obstante, ambas mediciones están relacionadas, como exige el EBEP, con las conductas, el rendimiento y logro de resultados de los empleados públicos en el ejercicio de su trabajo. En este sentido, una correlación significativa entre ambos tipos de medidas es un indicador de que en efecto están midiendo aspectos complementarios del desempeño pero relativamente independientes. En la Tabla 10 aparecen las correlaciones. Como puede apreciarse, los criterios externos globalmente considerados, correlacionan algo más de .30 con las evaluaciones del desempeño hechas por el superior jerárquico. Esta correlación es un indi-

Tabla 10. Correlaciones entre el Criterio Compuesto Externo y las Dimensiones del Desempeño Ocupacional

Variable	r_{obs}	r_{oper}	r_{verd}
Desempeño de Tarea	.20	.30	.32
Desempeño Contextual	.21	.32	.34
Desempeño Organizacional	.19	.29	.31
Desempeño Global	.21	.32	.33

Nota: Todas las correlaciones son significativas ($p < .001$). r_{obs} = correlación observada; r_{oper} = correlación observada corregida por atenuación en el criterio externo; r_{verd} = correlación observada corregida por atenuación en X e Y.

cadador de la validez de las evaluaciones del desempeño, dado que son criterios externos objetivos independientemente registrados.

Análisis Factorial Confirmatorio

El resultado del análisis factorial exploratorio de primer orden mostró que las medidas del desempeño estaban asociadas con dos componentes ortogonales. Queda pendiente comprobar si el espacio total del desempeño en el trabajo consiste en un único factor (componente) o por el contrario es más de uno. El análisis factorial confirmatorio nos permitirá responder a la anterior cuestión. Nuestra hipótesis es que existirán dos factores de primer orden y un factor general del desempeño. Asimismo, el primer factor, que llamaremos evaluación del desempeño en el trabajo, cargarán significativamente las evaluaciones de desempeño de tarea, contextual y organizacional, y en el segundo factor, que llamaremos criterio compuesto externo, cargarán el logro de objetivos individuales, la formación y el desempeño de puestos (permanencia y asistencia). Por último, los dos factores de primer orden cargarán significativamente en el factor general de desempeño (factor de segundo orden). El análisis factorial confirmatorio en el presente caso se hizo con el programa Lisrel. Los resultados se indican en la Tabla 11.

Así pues, como puede apreciarse, cada una de las variables pesa significativamente en el correspondiente factor y los dos factores de primer orden cargan significativamente en el factor general (segundo orden). Por tanto, desde este punto de vista la hipótesis se cumple. En cuanto a los estadísticos de ajuste del modelo, todos ellos han resultado totalmente satisfactorios, indicando que el modelo se ajusta totalmente a los datos. En consecuencia, el modelo de evaluación del desempeño y seguimiento de la actividad profesional utilizado por la APPA quedaría compuesto del modo representado en la Figura 1.

Discusión y Conclusiones

La investigación empírica llevada a cabo en los últimos veinticinco años ha demostrado que el dominio

Tabla 11. Cargas Factoriales del Análisis Factorial Confirmatorio (AFC) y Estadísticos de Ajuste del Modelo de un Factor General del Desempeño y dos Factores de Primer Orden

Variable	Factor 1	Factor 2	Factor General
DT	.94		
DC	.93		
DO	.92		
OI		.45	
DP		.49	
FO		.33	
Factor 1			.44
Factor 2			.75

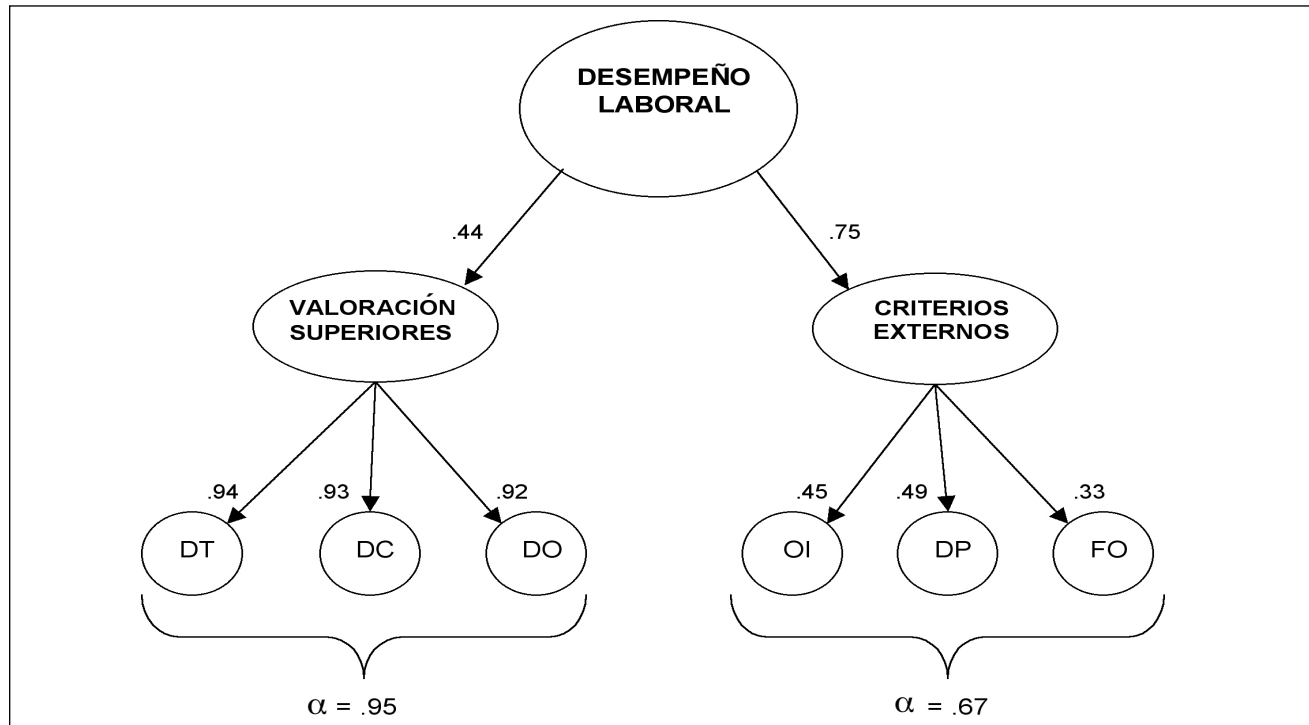
Estadísticos de Bondad de Ajuste

$\chi^2 =$	21.10 ($p = .0036$)
GFI=	0.99
RMR=	0.014
NFI=	0.98
CFI=	0.99
SRMR	0.014

del desempeño laboral es multidimensional y que cada dimensión tiene una serie de competencias o facetas (Aguinis, 2007). Además, investigaciones y propuestas teóricas independientes han convergido en considerar que en el dominio del desempeño deben considerarse tres grandes clases de conductas que determinan el éxito o el fracaso de los resultados del desempeño (Borman et al., 2001; Campbell et al., 2003; Gruys y Sackett, 2003; Murphy, 1990). Igualmente, la investigación ha indicado que el desempeño en el trabajo tiene como característica central que las conductas tienen que estar bajo control del individuo y, de este modo, ha separado los conceptos de desempeño laboral y productividad, siendo ésta última una resultante del primero, a la que además afectan otras variables fuera del control del empleado (p.e. circunstancias económicas, estilos de dirección, situaciones laborales) (Campbell et al., 1996; Murphy, 1990). El consenso entre las propuestas teóricas y la investigación sugiere que las grandes dimensiones (factores) del desempeño ocupacional serían tres: Desempeño de tarea (task performance), desempeño contextual (*contextual o civic performance*) y conductas contraproductivas (*counter-productive work behaviors*). Las dos primeras dimensiones incluyen conductas y competencias positivas o a favor de la organización, mientras que la última incluiría comportamientos contrarios a los intereses de las organizaciones. No obstante, debido a la naturaleza dimensional de los factores, la última dimensión puede verse también desde el polo positivo y, en este sentido, nosotros la redenominaamos desempeño organizacional. Los tres factores o dimensiones son entidades diferenciadas y guardan una cierta relación entre sí, aunque cada una de ellas acota un espacio propio del dominio del desempeño.

Los avances mencionados en la conceptualización y

Figura 1. Modelo de Evaluación del Desempeño y Seguimiento de la Actividad Profesional de la APPA



Nota: DT = Desempeño de Tareas; DC = Desempeño Contextual, DO = Desempeño Organizacional; OI = Objetivos Individuales; DP = Desempeño de Puesto (Permanencia+Absentismo); FO = Formación. Los números que aparecen en las líneas son los coeficientes Lisrel.

medida del desempeño en el trabajo son relevante en relación con las recientes disposiciones legales que establecen que todos los empleados públicos tienen el derecho a ser evaluados en su desempeño y la Administración Pública la obligación de llevarlo a cabo. En este sentido, El Estatuto Básico del Empleado Público (EBEP) define lo que la Administración Pública entiende por desempeño. Sin embargo, la falta de tradición en la evaluación del desempeño de los empleados públicos, unida a todavía escaso desarrollo de las disposiciones legales (p. ej. leyes autonómicas reguladoras), hacen que muchas Administraciones Públicas se encuentren sin modelos evaluativos y sin instrumentos de medida del desempeño. Además, los potenciales evaluadores tienen que ser formados y entrenados en los procesos evaluativos y de feedback y también es preciso desarrollar procesos de cambio organizacional que logran vencer las resistencias a la implantación de la evaluación del desempeño.

Un aspecto importante a tener en cuenta es que el EBEP ha definido el desempeño laboral de manera semejante a como lo ha hecho la investigación empírica (p. ej., Borman, 2004, Borman et al., 2010; Campbell et al., 1996; Murphy, 1996) y, a lo largo de los artículos 20, 52, 53, 54 y 95, ha señalado el tipo de conductas laborales de interés para el desempeño ocupacional en la Administración Pública española. Tales conductas se corresponden con el marco de las tres grandes dimensiones del desempeño ocupacional señaladas en la investigación. En este sentido, el desarrollo

de sistemas evaluativos en la Administración Pública española puede guiarse por la investigación empírica.

En cumplimiento de las exigencias legales que establece el EBEP, la Administración General del Principado de Asturias ha desarrollado un modelo de evaluación del desempeño y de seguimiento de la carrera profesional, específico para sus empleados y que se fundamenta en el EBEP y en las disposiciones autonómicas que le son propias. El modelo de evaluación del desempeño de la APPA se compone de indicadores de las tres grandes dimensiones y de criterios externos objetivos, independientes de los anteriores.

Dicho modelo y el sistema evaluativo correspondiente han sido aplicados a una muestra representativa de empleados públicos del Principado de Asturias, que incluía miembros de todos los grupos ocupacionales. La aplicación del sistema evaluativo ha permitido examinar empíricamente sus propiedades psicométricas y, a tal efecto, se han examinado la fiabilidad, la validez de constructo y la validez concurrente y discriminante con criterios externos de resultados laborales.

El examen empírico del modelo de la APPA ha demostrado que la evaluación de las dimensiones de desempeño es altamente fiable, con coeficientes alfa de Cronbach por encima de .85 para cada una de las dimensiones y de .95 para la puntuación global. Igualmente se ha demostrado que la evaluación del desempeño hecha por los superiores jerárquicos correlaciona significativamente con las tres medidas externas de resultados del desempeño: formación, logro de

objetivos individuales y desempeño de puestos (permanencia y absentismo).

Los resultados de la evaluación hecha por los superiores jerárquicos indican que existen en la misma un ligero sesgo de benevolencia (puntuaciones hacia el polo superior de la escala), explicable por la propia naturaleza administrativa de la evaluación y por el número de niveles que la APPA determinó para dicha evaluación. No obstante, un 3% de los empleados públicos de la muestra no habrían superado la evaluación si la puntuación de corte se sitúa en 2 puntos y un 15% de los mismos no la habría superado si la puntuación mínima se hubiese situado 2.7 puntos.

La validez de constructo de la evaluación ha sido demostrada tanto mediante análisis exploratorio de componentes principales como mediante análisis factorial confirmatorio. En este último caso, los índices de ajuste del modelo a los datos son excelentes. De acuer-

do a los resultados, el modelo de evaluación del desempeño y del seguimiento de la carrera en la APPA tendría dos grandes factores, uno correspondiente a las valoraciones hechas por los superiores jerárquicos y otra correspondiente a los tres criterios externos de resultados.

En resumen, el análisis de las propiedades psicométricas del instrumento de evaluación del desempeño de los empleados de la Administración del Principado de Asturias ha demostrado ser consistente con los artículos del EBEP y, por tanto, poseer validez de contenido, ha resultado ser altamente fiable (coeficientes alfa de Cronbach superiores a .80), tiene una estructura factorial clara y confirmada mediante análisis factorial confirmatorio y las relaciones con los criterios externos son las esperadas e indican convergencia entre las evaluaciones de los superiores jerárquicos y las medidas objetivas incluidas en el proceso evaluativo.

Extended Summary

Empirical research conducted in the last twenty years has shown that the domain of work performance is multidimensional and that each dimension has a set of competencies or facets (Aguinis, 2007). Additionally, independent research and theoretical approaches have converged to consider that job performance should be considered as consisting of three broad classes of behavior that determine the success or failure of performance outcomes (Borman et al., 2001; Campbell et al., 2003; Gruys and Sackett, 2003; Murphy, 1990). Similarly, research has indicated that the key characteristic of job performance is that behaviors must be under control of individuals. Consequently, the concepts of job performance and productivity must be separated, the latter being a result of the first, which is also affected by other variables beyond the control of the employee (e.g. economic conditions, leadership styles, work situations) (Campbell et al. 1996; Murphy, 1990). The agreement between the theoretical models and research results suggests that job performance would consist of three big dimensions (factors): task performance, contextual performance, and counterproductive behaviors. The first two dimensions include positive behaviors and competencies or towards organizations, while the latter would include behavior contrary or against to organizational interests. However, due to bipolar characteristics of job dimensions, the last dimension can also be seen from the positive pole and, in this sense, we rename it as organizational behavior. The three factors or dimensions are separate entities but with a certain relationship to each other, although each of them their own space narrows the domain of performance.

The aforementioned advances in the conceptualization and measurement of job performance are relevant in relation to recent laws passed in Spain which state

that all civil servants have a right to be assessed on their performance and Public Administration must to bring him out

In this sense, the Basic Statute for Civil Servants (EBEP) defines what performance means in Public Administration. However, the lack of tradition in performance evaluation of civil servants, coupled with the still legal provisions under developing (e.g. regulatory laws of Autonomous Communities), make that many public administrations do not even have an appraisal model and instruments for measuring job performance. Additionally, potential evaluators need to be trained and trained in appraisal and feedback processes. Organizational change processes have also to be developed in order to be successful in overcoming resistance to the introduction of performance appraisal.

An important aspect to consider is that EBEP has defined job performance similarly to as it has done empirical research (e.g., Borman, 2004; Borman et al., 2010; Campbell et al., 1996; Murphy, 1996) and, in articles 20, 52, 53, 54, and 95, clarifies the type of work behaviors that are and are not performance for the Spanish Public Administration. Such behaviors are consistent with the framework of the three major dimensions of occupational performance identified in the investigation. In this sense, the development of appraisal systems in the Spanish Public Administration can be guided by empirical research.

In compliance with the legal requirements established by EBEP, the General Administration of the Principality of Asturias has developed a model of performance assessment and follow-up of professional career, specific to their employees. The model is based on the EBEP and regional arrangements particular to them. The APPA appraisal model consists of indicators of the three major dimensions and objectives of external criteria, independent of the above.

Research reported in this article consisted of three studies. The main aim of study-1 was to identify what competencies would be included in the appraisal model. Due to the fact that we wished to include competencies (facets) of the three Big occupational dimensions, i.e. task performance, contextual performance, and organizational performance, a list of fifty independent competencies was prepared. The definition and examples of behaviors corresponding to every competency were included in a form. The questionnaire included 29 competencies of task performance, 12 competencies of contextual performance, and 9 competencies of organizational performance. A representative sample of 202 managers rated each competency in a 5-point scale, where 5 means absolutely necessary. Tables 1 and 2 show how the sample was divided in terms of group and level. Table 3 shows the descriptive statistics of the five competencies selected for the big three performance dimensions. As can be seen, on average, the selected competencies showed a rating above 4. Tables 4 and 5 show the consistency of the ratings in terms of Cronbach's alpha coefficients. This analysis was independently conducted for A and B groups and C and D groups. The results showed that the ratings were reliable and that they were more reliable for task performance in both analyses. A second goal of study-1 was to obtain evidence of content validity. Six expert matter subjects examined independently the content of the selected competencies and checked whether the content (definition and examples) was according to the corresponding articles of the EBEP. The agreement was near perfect for all competencies (Lawshe's ICV = .80, $p < .05$; Aitken's V = .89, $p < .01$).

The objective of study 2 was to develop Behaviorally Anchored Rating Scales for the competencies selected in study. Based on a large number of critical incidents collected in occupations of Public Administration, six individuals with experience in job analysis and the critical incident technique examined and rated the incidents. For each competency, the incidents more close to 1, 3 and 5 were retained for additional inspection. Next, the incidents were re-translated to form the basis of the anchors included in the BARS. Finally, the General Manager of Personnel, together with the manager of Personnel Office and five additional managers checked the BARS and approved the final version.

The third study was aimed at establishing the psychometric properties of the performance evaluation procedure. The specific goals were to estimate the measure reliability, to analyze the internal structure of the model, and to correlate the ratings with external criteria of performance. A confirmatory factor analysis (CFA) was carried out to examine the fit of model to data. The sample consisted of 4422 employees of different groups and scales of civil servants who serve in the APPA. It was distributed as

follows: Group A1 = 701, A2 = 735, C1 = 652, C2 = 1795, E = 539.

The performance of each employee during the twelve months preceding the date of the evaluation was appraised by the immediate manager (Head of CGU). The performance appraisal was done using the 15 BARS developed in Study 2. Thus five BARS appraised task performance, five appraised contextual performance, and five appraised organizational performance.

Table 6 shows the Coefficient Alpha for the three performance dimensions and the global performance in each occupational group. For task performance, reliability coefficients ranged from .82 to .879. They ranged from .835 to .891 for contextual performance, and, finally, for organizational performance ranged from .863 to .897. Therefore, they are highly reliable. The coefficients ranged from .939 to .957 for the composite of global performance. Consequently, the global measure of performance is the more reliable estimator of civil servant performance.

Tables 7 and 8 show the results of an exploratory component analysis carried out with the ratings of the managers and with the external criteria. As can be seen, two components explain 66.16% of variance. The ratings of the three performance dimensions loaded in the first component and the three external criteria loaded in the second component. Moreover, there was any relevant secondary loading. Consequently, both components are totally orthogonal.

Table 9 and 10 show the correlations among job performance ratings and external criteria in each occupational group and in the total sample. In general, the correlations are small and similar to the coefficients found in previous research. A composite of external criteria was correlated with the three dimensions of job performance and with global performance. When the correlation is corrected for measurement error in both measures, the values ranged from .31 for organizational performance to .34 for contextual performance.

Finally, as can be seen in Table 11, a confirmatory factor analysis showed that the model fits very well to data, as the indices of goodness-of-fit are excellent. The model of job performance of the Public Administration of the Principality of Asturias is represented in Figure 1.

In summary, the analysis of the psychometric properties of the performance evaluation of employees of the Administration of the Principality of Asturias has been shown (1) be consistent with the Articles of EBEP and, hence, may have content validity, (2) has proven being highly reliable (Cronbach alpha coefficients greater than .80), (3) to have a clear factor structure and (4) be confirmed by confirmatory factor analysis and (5) that the relationships with external criteria indicate convergence between superiors assessments and objective measures included in the evaluation process.

Referencias

- Aamond, (1999). *Applied Industrial/Organizational Psychology* (3ª Edición). Belmont, CA: Wadsworth Publishing Company.
- Aguinis, H. (2007). *Performance management*. Upper Saddle River, NJ: Pearson/Prentice-Hall.
- Aitken, L. R. (1980). Content validity and reliability of single items questionnaires. *Educational and Psychological Measurement*, 40, 955-959.
- Bateman, D. T. y Organ, D. W. (1983). Job satisfaction and the good soldier: the relationship between affect and employee "citizenship". *Academy of Management Journal*, 26, 587-595.
- Bennett, R. J., y Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85, 349-360.
- Bernardin, H. J., LaShells, M. B., Smith, P. C. y Alvares, K. M. (1979). Behavioral expectation scales: Effects of developmental procedures and formats. *Journal of Applied Psychology*, 61, 75-79.
- Bernardin, H. J. y Smith, P. C. (1981). A clarification of some issues regarding the development and use of Behaviourally Anchored Rating Scales (BARS). *Journal of Applied Psychology*, 66, 458-463.
- Borman, W. C. (2004). The concept of organizational citizenship. *Current Directions in Psychological Science*, 13, 238-241.
- Borman, W. C. y Brusck, D. H. (1993). More progress toward a taxonomy of managerial performance requirements. *Human Performance*, 6, 1-22.
- Borman, W. C., Bryant, R. H. y Dorio, J. (2010). The measurement of task performance as criteria in selection research. En J. L. Farr y N. T. Tippins (Eds.). *Handbook of personnel selection* (pp. 439-461). New York: Routledge.
- Borman, W. C., Penner, L. A., Allen, T. D., y Motowidlo, S. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9, 52-69.
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. En M. D. Dunnette y L. M. Hough (Eds.). *Handbook of industrial and organizational psychology: Vol. 1*. (pp. 687-732). 2ª edición. Palo Alto, CA: Consulting Psychologist Press.
- Campbell, J. P., Gasser, M. B. y Oswald, F. L. (1996). The substantive nature of job performance variability. En K.R. Murphy (Ed). *Individual differences and behavior in organizations* (pp. 258-299). San Francisco, CA: Jossey-Bass
- Conway, J. (1999). Distinguishing contextual performance from task performance for managerial jobs. *Journal of Applied Psychology*, 84, 3-13.
- Dalal, R. S. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology*, 90, 1241-1255.
- Dorsey, D. W., Cortina, J. M. y Luchman, J. (2010). Adaptive and citizenship-related behaviors at work. En J. L. Farr y N. T. Tippins (Eds.). *Handbook of personnel selection* (pp. 463-487). New York: Routledge.
- Fox, S. y Spector, P. E. (2005). *Counterproductive work behavior: Investigations of actors and targets*. Washington, DC: American Psychological Association.
- Gruys, M. L. (2000). *The dimensionality of deviant employee behavior in the workplace*. Ph.D. dissertation, University of Minnesota. Retrieved August 7, 2008, from Dissertations y Theses: AyI database. (Publication No. AAT 9972966).
- Gruys, M. L., Stewart, S. M. y Bowling, N. A. (2010). Choosing to report: Characteristics of employees who report the counterproductive work behavior of others. *International Journal of Selection and Assessment*, 18, 439-446.
- Gruys, M. L. y Sackett, P. R. (2003). Investigating the dimensionality of counterproductive work behavior. *International Journal of Selection and Assessment*, 11, 30-42.
- Hardigree, A. L. (2006). *Meta-analysis of age and job performance relation: is job complexity a moderator?* Master Arts Dissertation no publicada, Rice University.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28, 563-575.
- Ley 7/2007, de 12 de abril, del *Estatuto Básico del Empleado Público*, BOE de 13 de Abril de 2007, pp. 16270-16299.
- McCloy, R. A., Campbell, J. P. y Cudeck, R. (1994). A confirmatory test of a model of performance determinants. *Journal of Applied Psychology*, 79, 493-505.
- McEvoy, G. M. y Cascio, W. F. (1989). Cumulative evidence of the relationship between employee age and job performance. *Journal of Applied Psychology*, 74, 11-17.
- Motowidlo, S., Borman, W. C. y Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, 10, 71-83.
- Murphy, K. R. (1990). Job performance and productivity. En K. R. Murphy y F. E. Saal (Eds.). *Psychology in Organizations: Integrating science and practice* (pp. 157-176). Hillsdale, N.J.: Erlbaum.
- Murphy, K. R. (1982). Difficulties in the statistical control of halo. *Journal of Applied Psychology*, 67, 161-164.
- Murphy, K. R., Jako, R. A. y Anhalt, R. L. (1993). Nature and consequences of Halo error: a critical analysis. *Journal of Applied Psychology*, 78, 218-225.
- Ng, T. y Feldman, D. C. (2010). Organizational tenure and job performance. *Journal of Management*, 36, 1220-1250.
- Nunnally, J. C. (1978). *Psychometric theory*. (2ª Edición). New York: McGraw-Hill.
- Robinson, S. L. y Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal*, 38, 555-572.
- Rotundo, M. y Spector, P. E. (2010). Counterproductive work behavior and withdrawal. En J. L. Farr y N. T. Tippins (Eds.). *Handbook of personnel selection* (pp. 489-511). New York: Routledge.

- Sackett, P. R., Berry, C. M., Wiemann, S. A. y Laczo, R. M. (2006). Citizenship and counterproductive work behavior: Clarifying relationships between the two domains. *Human Performance, 19*, 441-464.
- Sackett, P. R. y DeVore, C. J. (2001). Counterproductive behaviors at work. In N. Anderson, D. S. Ones, H. Sinangil y C. Viswesvaran (Eds.) *Handbook of Industrial, Work, and Organizational Psychology: Vol. 1.* (pp 144-164). London, United Kingdom: Sage.
- Salgado, J. F., Anderson, N. y Hülshager, U. R. (2010). Employee selection in Europe: Psychotechnics and the forgotten history of modern scientific employee selection. En J. L. Farr y N. T. Tippins (Eds.). *Handbook of personnel selection* (pp. 921-941). New York: Routledge.
- Salgado, J. F., Moscoso, S. y Anderson, N. (2011). Personality and counterproductive behaviors at work. En R. Tett y N. Christiansen (Eds.). *Handbook of personality at work*. En prensa.
- Smith, P. C. y Kendall, L. M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for ratings scales. *Journal of Applied Psychology, 47*, 149-155.
- Smith, C. A., Organ, D. W. y Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology, 68*, 653-663.
- Society for Industrial and Organizational Psychology, Inc. (1987). *Principles for the validation and use of the personnel selection procedures*. College Park, MD: Author.
- Waldman, D. A. y Avolio, B. J. (1986). Meta-analysis of age differences in job performance. *Journal of Applied Psychology, 71*, 33-38.

Manuscrito Recibido: 25/01/2011

Revisión Recibida: 15/05/2011

Aceptado: 20/05/2011