

FACTORES RELEVANTES PARA AUMENTAR LA PRECISIÓN, LA VIABILIDAD Y EL ÉXITO DE LOS SISTEMAS DE EVALUACIÓN DEL DESEMPEÑO LABORAL

RELEVANT FACTORS TO INCREASE THE ACCURACY, FEASIBILITY AND SUCCESS OF JOB PERFORMANCE EVALUATION SYSTEMS

**Dolores Díaz Cabrera, Estefanía Hernández Fernaud, Rosa Isla Díaz, Naira Delgado Rodríguez,
Luis Díaz Vilela y Christian Rosales Sánchez**
Universidad de La Laguna

La situación de crisis económica en nuestro país ha provocado, por una parte, un incremento del interés de las organizaciones en la evaluación del desempeño laboral. Por otra parte, está produciendo una mayor resistencia a la evaluación por parte de los evaluados, que miran con desconfianza y preocupación su evaluación en momentos de dificultades económicas. Por ello, en situaciones económicas como la actual, es muy importante desarrollar escalas de evaluación que permitan lograr la efectividad, la credibilidad y el éxito de estos sistemas. El objetivo general de este trabajo es aportar resultados dirigidos a que los sistemas de evaluación sean valorados positivamente tanto por la organización como por los participantes. Para ello es necesario que: (a) las escalas de respuesta usadas ofrezcan datos precisos y fiables; y (b) los procedimientos aplicados para la evaluación sean percibidos como justos y fiables por los participantes.

Palabras clave: Evaluación del desempeño laboral, Escalas de desempeño de tareas y cívico, Reacciones a la evaluación del desempeño, Justicia percibida.

The situation of economic crisis in Spain has motivated an increasing interest in employee performance evaluation in organizations. On the other hand, it is also producing greater resistance to this evaluation on the part of appraisees, who regard these processes with suspicion and concern in these times of economic difficulties. Therefore, in economic situations like the current one, it is very important to develop assessment scales that facilitate the effectiveness, credibility and success of these systems. The general goal of this work is to provide results that can help evaluation systems to be positively valued by organizations as well as by their members. For this it is necessary that: (a) the response scales used offer precise and reliable data; and (b) the procedures applied to assess performance are perceived as fair and reliable by the participants.

Key words: Work performance assessment/appraisal, Task and citizenship performance scales, Reactions to performance assessment/appraisal, Perceived fairness/justice.

La situación de crisis económica en nuestro país ha aumentado el interés de las organizaciones en el estudio de la eficiencia y el desempeño laboral. Sin embargo, también se ha incrementado la resistencia a la evaluación por parte de los evaluados, e incluso de los evaluadores, que miran con desconfianza y preocupación su evaluación en estos momentos de dificultades económicas. Esta resistencia es producto, principalmente, de que la evaluación del desempeño suele asociarse con consecuencias negativas para los trabajadores. Y estas consecuencias son percibidas aún más negativamente en situaciones de crisis económica.

Correspondencia: Dolores Díaz Cabrera. Departamento de Psicología Cognitiva, Social y Organizacional. Facultad de Psicología. Campus de Guajara. 38205 La Laguna. Tenerife. España. E-mail: mddiaz@ull.es

.....
Esta investigación se ha desarrollado con la financiación conjunta del convenio Diseño de un sistema de evaluación del desempeño para los empleados públicos de las Administraciones Públicas Canarias del Gobierno de Canarias, y del proyecto del Plan Nacional de Investigación PSI2010-17327

La resistencia a las evaluaciones, y las reacciones asociadas a ellas por parte de los trabajadores, conduce con cierta frecuencia a que las organizaciones retrasen la implementación de estos sistemas de evaluación, al estar inmersas en procesos de reorganización, reestructuración o simplemente por la cautela que les merece la cuestión de evaluar el desempeño de su plantilla en estos momentos.

Sin embargo, otras organizaciones, ante la situación económica actual, se plantean como uno de sus objetivos centrales la mejora del desempeño laboral con el fin de aumentar su eficacia. Un riesgo importante que se deriva de este escenario es que la urgencia por disponer de indicadores de desempeño en poco tiempo dificulte la transparencia, la precisión de la evaluación, y, por tanto, reduzca la confianza de los agentes implicados en estos procesos.

En situaciones económicas como la actual es aún más importante tomar en consideración los tres criterios propuestos por Cardy y Dobbins (1994), que pretenden lograr la efectividad, credibilidad y el éxito de estos sistemas. Estos tres criterios son: (a) considerar los aspectos cualitativos del proceso de evalua-

ción; (b) reducir los sesgos propios de los evaluadores del desempeño; y (c) incrementar la precisión de la medida utilizada para realizar la evaluación.

Este artículo presenta los resultados obtenidos por nuestro equipo de investigación en varios estudios que abordan aspectos relacionados con estos tres criterios. En el marco de un convenio de I+D+I financiado por la Administración Pública Canarias en 2009-2010 se inició el diseño de un sistema de evaluación del desempeño para los empleados públicos de la Administración Pública Autónoma. Además, en la actualidad nuestro equipo desarrolla un proyecto de investigación financiado por el Plan Nacional (2011-2014) sobre las relaciones entre el desempeño laboral, y algunas variables personales y organizacionales, y el papel de la formación de los evaluadores en la valoración del desempeño.

En este artículo abordamos diversos factores que consideramos relevantes para que los sistemas de evaluación sean valorados positivamente tanto por la organización como por los evaluadores y evaluados, reduciendo así la resistencia a la participación. En nuestra opinión, para ello es necesario, en primer lugar, que los procedimientos aplicados para la evaluación sean percibidos como justos por los participantes, a través de, entre otros aspectos, ofrecer información sobre el sistema de evaluación y fomentar la participación de los implicados. En segundo lugar, que las escalas de respuesta usadas ofrezcan datos precisos y fiables, y esto depende de que se tengan en cuenta los dos primeros criterios planteados anteriormente: reducción de sesgos y desarrollo de escalas precisas.

Específicamente, en este artículo se explora, en primer lugar, qué variables influyen en las percepciones de los trabajadores hacia la implementación de un sistema de evaluación del desempeño (p. ej., actitudes y creencias, justicia percibida y la aceptación de estos sistemas). En segundo lugar, se analizan las propiedades psicométricas de dos escalas de evaluación (escala tipo Likert -o escala gráfica de conductas- y escala tipo estandarizada mixta), y sus efectos sobre los sesgos de respuesta. Antes de abordar estos puntos, en el siguiente apartado se comentan algunos conceptos básicos sobre el desempeño laboral y su relevancia para la práctica profesional.

LA IMPORTANCIA DE LA EVALUACIÓN DEL DESEMPEÑO EN LA GESTIÓN DE LOS RECURSOS HUMANOS

La evaluación del desempeño laboral tiene importantes ventajas tanto para las organizaciones como para los trabajadores. Una de sus principales aportaciones es la de aumentar el conocimiento de la organización sobre cómo se está realizando el trabajo. La información obtenida orientará, además, en el diseño y desarrollo de los procesos de selección y en la detección de las necesidades de formación, entre otras estrategias de gestión de los Recursos Humanos. Asimismo, otras utilidades de las evaluaciones para las organizaciones están relaciona-

das con el reconocimiento de los trabajadores y el desarrollo de planes motivacionales. Por su parte, los trabajadores obtendrán información útil que les permite tener una mayor claridad sobre sus funciones y tareas, tomar conciencia sobre cómo realizan su trabajo y mejorar sus prácticas laborales.

La evaluación del desempeño ha sido objeto de interés de numerosos investigadores y profesionales. Una definición general del desempeño laboral comprende los comportamientos y resultados observables y cuantificables en los que los empleados se implican o participan de ellos, contribuyendo a la consecución de las metas organizacionales (Viswesvaran y Ones, 2000).

En la evaluación del desempeño se considera importante incluir todas aquellas conductas que repercuten, positiva o negativamente, en los resultados organizacionales, y no limitarse a las tareas asignadas formalmente al puesto y a la posición específica (Salgado y Cabal, 2011). Desde los años noventa se distinguen dos tipos de desempeño laboral, el desempeño en la tarea y el desempeño contextual (también denominado cívico). El primero se refiere al rol prescrito que debe cumplir el empleado, relacionado directamente con la eficacia con la que éste realiza actividades que contribuyen a las metas organizacionales (Borman y Motowidlo, 1993).

El desempeño cívico se define como el conjunto de comportamientos que no se encuentran directamente relacionados con el desarrollo de las tareas asignadas al puesto de trabajo, pero que tienen gran importancia, ya que promueven un adecuado contexto organizacional, social y psicológico para un desarrollo eficaz de las tareas encomendadas (Borman y Motowidlo, 1993). Lo importante en este tipo de desempeño es la iniciativa, el apoyo y la persistencia que muestra el empleado más que la competencia técnica demostrada (Poropat, 2002).

En los siguientes apartados abordamos algunos de los resultados obtenidos por nuestro equipo de investigación en tres estudios que han tenido como objetivo general lograr una mayor efectividad y éxito de los sistemas y procedimientos de evaluación del desempeño. En concreto, en el siguiente punto nos centraremos en uno de los tres criterios planteados por Cardy y Dobbins (1994): los aspectos cualitativos del sistema de evaluación. Se describen brevemente los principales planteamientos que se han realizado acerca de estos aspectos, así como algunos de los resultados obtenidos en nuestra investigación.

FOMENTANDO LA ACEPTACIÓN DE LOS PARTICIPANTES HACIA EL SISTEMA DE EVALUACIÓN: JUSTICIA PERCIBIDA Y CONFIANZA

La mayoría de los estudios realizados sobre la evaluación del desempeño se han orientado principalmente al desarrollo de escalas de respuesta que reduzcan, en la mayor medida posible, los sesgos en la evaluación, principalmente el sesgo de be-

nevolencia o tendencia de los evaluadores a asignar puntuaciones altas. Aunque estos objetivos son muy relevantes para lograr evaluaciones precisas, se ha criticado la escasa atención prestada a otros criterios como son las reacciones de los participantes a las evaluaciones de desempeño y los factores que contribuyen a esas reacciones (Cardy y Dobbins, 1994; Murphy y Cleveland, 1995, cfr. Cawley, Keeping y Levy, 1998). Esta carencia de atención ha conducido a Murphy y Cleveland (1995) a denominar este tercer criterio, las percepciones y reacciones de los participantes, como “el criterio descuidado” (p. 310).

Las actitudes de los trabajadores hacia los sistemas de evaluación influyen en su participación en las evaluaciones de desempeño. En una encuesta realizada en el Reino Unido (Industrial Society, 1997, cfr. Fletcher, 2002) sobre la participación en los procesos de evaluación del desempeño se obtuvo que menos de las dos terceras partes de las organizaciones valoradas lograron un 67% del total de las evaluaciones posibles. Efectivamente, aunque se cuente con un sistema de evaluación altamente preciso, si no es aceptado y apoyado por los miembros de la organización, su efectividad y viabilidad estará limitada (Farr y Jacobs, 2006; Keeping y Levy, 2000).

La importancia del contexto social y organizacional en el que se desarrollan las evaluaciones es acentuada en lo que Folger, Konovsky y Cropanzano (1992) denominan como la metáfora de “*due-process*” (del proceso debido). Esta metáfora defiende la importancia de que los procesos a desarrollar tengan en cuenta aspectos como: (a) que los participantes tengan información previa sobre los objetivos y criterios que se van a usar y que puedan participar en su desarrollo, o al menos opinar sobre ellos; (b) que los evaluados tengan la oportunidad de explicar los resultados que han obtenido en la evaluación; (c) que los evaluadores estén familiarizados con los puestos de trabajo a evaluar y con el sistema de evaluación; y (d) que los evaluadores tengan la oportunidad de realizar juicios basados en la evidencia, sin presiones externas o prejuicios personales, y partiendo de valores como la honestidad y la justicia.

Uno de los constructos con mayor impacto sobre la aceptación y la efectividad de un sistema de evaluación es el que Farr y Jacobs (2006) proponen etiquetar como “*la confianza colectiva de los participantes en el proceso de evaluación*” (p. 324), que recoge la idea de que los participantes deben confiar en el proceso antes de que la evaluación pueda ser efectivamente implementada en la organización.

Por su parte, la Teoría de la justicia organizacional incide en la importancia de constructos similares, al poner el acento en que el éxito de los procesos de evaluación dependerá de las reacciones de los participantes. Estas reacciones dependen, entre otros aspectos, de: (a) las actitudes y creencias de los participantes sobre cuestiones como su comodidad con los

cuestionarios y las escalas de puntuación utilizadas, sus creencias acerca de la finalidad de esas evaluaciones, la satisfacción con el evaluador, las percepciones de los evaluadores sobre su autoeficacia para la evaluación, etc.; (b) el contexto de la evaluación: por ejemplo, el grado de participación de los miembros organizacionales en el desarrollo del proceso de evaluación y su satisfacción con el sistema de evaluación; y (c) la satisfacción con la retroalimentación de los resultados obtenidos en la evaluación (Giles y Mossholder, 2000; Levy y Williams, 2004; Tziner, Murphy y Cleveland, 2001).

Otros estudios han mostrado además la relación entre variables como el compromiso afectivo con la organización (sentimientos positivos de identificación, implicación y apego hacia la organización) y la percepción del sistema de evaluación como justo (Colquitt, Conlon, Wesson, Porter y Yee, 2001; Colquitt et al., 2013; Dirks y Ferrin, 2001; Salles, Amin, Muda y Halim, 2013). Por tanto, la aceptación de los sistemas de evaluación por parte de los participantes y su grado de implicación en éstos está determinada por su confianza en el proceso concreto desarrollado y por la justicia percibida en relación a su desarrollo, a su aplicación y a la retroalimentación que se aportará.

El estudio realizado por nuestro equipo de investigación tenía como objetivo explorar el impacto de algunas de las variables comentadas anteriormente sobre las percepciones de los trabajadores hacia la implementación de un sistema de evaluación del desempeño que se estaba diseñando en esos momentos y que iba a ser implementado en un corto plazo, y por primera vez, en una unidad organizacional (Díaz-Cabrera, Hernández-Fernaud, Rosales-Sánchez, Isla y Díaz-Vilela, en revisión).

En concreto, se analizó en qué medida algunas actitudes y creencias (Satisfacción laboral, Actitud hacia el cambio, Autoeficacia, Iniciativa y Compromiso afectivo) explican la Justicia percibida de la evaluación, y la Aceptación del futuro sistema de evaluación. Participaron 142 empleados públicos, que constituían el total de la plantilla de una unidad de una Administración Autónoma.

Los resultados obtenidos indican que el contexto social y organizacional en el que se realiza la evaluación así como las reacciones de los participantes explican, al menos parcialmente, la percepción de los trabajadores sobre el sistema de evaluación del desempeño a ser implementado. Específicamente, y en primer lugar, los resultados muestran que los trabajadores con un compromiso afectivo con la organización más alto tienden a valorar como más justa la puesta en marcha de un sistema de evaluación de su desempeño. Y, en segundo lugar, la aceptación del sistema de evaluación del desempeño por parte de los participantes se explica en un 13.4% por la satisfacción laboral, el compromiso afectivo y la actitud hacia los cambios que se pueden producir en el contexto laboral.

Estos resultados ponen de manifiesto la importancia de considerar las reacciones de los participantes hacia los sistemas de evaluación con el fin de facilitar su implementación. En el siguiente apartado se aborda otro aspecto relevante para el logro de la efectividad y aceptación de estos sistemas.

FOMENTANDO LA REDUCCIÓN DE LOS SESGOS DE EVALUACIÓN MEDIANTE LA PRECISIÓN DE LAS ESCALAS DE MEDIDA

Folger et al. (1992) utilizan la “metáfora del test” para describir las asunciones implícitas que subyacen al desarrollo de las escalas de evaluación. Estas asunciones se refieren a: (a) que la naturaleza de los puestos de trabajo permite mediciones válidas y fiables de las conductas y de los resultados laborales; (b) que los evaluadores pueden y desean hacer evaluaciones precisas; y (c) que existe un criterio de evaluación racional y verdadero para cada puesto.

Los criterios propuestos por Cardy y Dobbins (1994) pretenden superar estas asunciones. Concretamente, de estos tres criterios para mejorar el éxito de los sistemas de evaluación, dos de ellos inciden directamente en la elaboración de escalas, favoreciendo la reducción de los sesgos propios de los evaluadores del desempeño y el incremento de la precisión de los instrumentos de medida.

Aunque existen diferentes tipos de sesgos en la evaluación (p. ej., halo, severidad, primacía), el de benevolencia es uno de los más frecuentes. El sesgo de benevolencia es la tendencia por parte de los evaluadores a asignar puntuaciones altas en las evaluaciones (Aguinis, 2013; Guilford, 1954). Este sesgo tiene un efecto importante en el proceso de evaluación del desempeño, ya que reduce la posibilidad de reconocer y recompensar adecuadamente el desempeño de los trabajadores, y, por tanto, la validez de las evaluaciones disminuye considerablemente (Bretz, Milkovich y Read, 1992). La ocurrencia de este sesgo se atribuyó inicialmente a factores situacionales y organizacionales, considerándose posteriormente variables del evaluador como una posible causa. Así, la evaluación positiva del desempeño de forma recurrente se ha relacionado con el propósito de la evaluación del desempeño, la evitación de confrontación con los evaluados, la existencia de relaciones personales entre evaluadores y evaluados, o con algunas características de personalidad del evaluador, entre otras variables (Bernardin, Cooke y Villanova, 2000).

La reducción del sesgo de benevolencia se ha enfocado tradicionalmente hacia la elaboración de programas de formación dirigidos a los evaluadores, así como al desarrollo de formatos mejorados de las escalas de respuesta (Woehr y Huffcutt, 1994). Investigaciones previas han mostrado que el diseño de la escala de medida influye en la distribución de las respuestas a lo largo de la misma. Concretamente, Viswesvaran y Ones (2005) exponen que las escalas de elección forzada y las esca-

las estandarizadas mixtas (MSS) permiten reducir el sesgo de benevolencia cuando los evaluadores intentan sesgar sus respuestas intencionadamente. Nuestro equipo de investigación ha trabajado en los últimos años en el desarrollo de un formato de escala de respuesta diferente a la escala gráfica tradicional, y en la comparación del efecto que cada una de ellas tiene sobre el sesgo de benevolencia al evaluar el desempeño de tarea. Con estos objetivos en mente, hemos desarrollado dos investigaciones, una con trabajadores administrativos, y otra con estudiantes universitarios.

Para el desarrollo de la primera investigación se realizó, en primer lugar, un análisis de puestos a través de cuestionarios y entrevistas con el fin de identificar las tareas administrativas más relevantes en los puestos y que permitieran la evaluación del desempeño. A partir de esta información se elaboró el inventario de tareas que cumplimentaban los trabajadores, valorando la frecuencia, importancia y complejidad de cada tarea de los puestos de trabajo. Un comité de expertos, formado por investigadores, supervisores y ocupantes de los puestos, decidió la pertinencia de evaluar cada una de las tareas administrativas en función de su relevancia, incluyéndose 60 tareas (Díaz-Vilela, Delgado, Isla-Díaz, Díaz-Cabrera y Hernández-Fernaud, en revisión).

Con estas 60 tareas administrativas se elaboró una Escala de conductas observadas (BOS – Behavior Observation Scale) para la evaluación del desempeño, que fue cumplimentada, en formato lápiz y papel, por 53 trabajadores de una administración pública que desempeñaban tareas administrativas. Los trabajadores debían evaluar la frecuencia con la que realizaban cada una de las tareas de forma adecuada, mediante una escala gráfica de 10 puntos. Se comprobó que 14 de las tareas incluidas en el cuestionario eran desempeñadas por la mayoría de administrativos, mientras que las restantes presentaban mayor variabilidad entre puestos (Díaz-Vilela et al., 2013).

En segundo lugar, se elaboró un nuevo instrumento de evaluación del desempeño de trabajadores que realizan funciones administrativas, utilizando un formato de escala estandarizada mixta (MSS). Este instrumento incluía las mismas 14 tareas anteriores, aunque con una redacción y presentación de los enunciados diferente. Esta nueva escala se presentaba mediante una aplicación informática, de forma que para cada tarea aparecían en pantalla tres alternativas de respuesta que representaban niveles de desempeño distintos (Excelente, Bueno, Mejorable). Estas alternativas de respuesta se presentaban en posiciones y orden aleatorio en cada una de las tareas a evaluar. Es decir, las respuestas no siempre estaban ordenadas desde un desempeño deficiente a uno excelente, o al contrario. Una vez seleccionado el nivel de desempeño, se abrían otras alternativas de respuesta en las que debía indicarse con qué frecuencia se realiza ese nivel de desempeño.

Con esta nueva escala de respuesta se pretende: 1) evitar la deshabilitación social de las respuestas, al no estar etiquetado cada nivel de desempeño como Excelente, Bueno o Mejorable, ni presentarse necesariamente ordenados desde el peor al mejor desempeño, o al revés; 2) fomentar que los evaluadores presten mayor atención al contenido de las alternativas de respuesta, ya que al aparecer en distinto orden y posición cada vez requieren una lectura consciente de cada alternativa para seleccionar el nivel de desempeño deseado; y 3) reducir el sesgo de benevolencia, al reducir la posibilidad de responder a cada ítem de forma automática y solicitar la valoración de la frecuencia con que se realiza la tarea al nivel de desempeño seleccionado (Algunas veces, Habitualmente o Siempre).

En un primer estudio de comparación de estos dos formatos de escala de respuesta, escala gráfica tipo Likert y MSS, se contrastaron los datos obtenidos en dos muestras independientes de administrativos (Díaz-Cabrera et al., 2013). De los 61 administrativos que componían la muestra, el 37,7% autoevaluó su desempeño y fue evaluado por sus supervisores con la Escala BOS con escala gráfica de respuesta tipo Likert en formato lápiz y papel, y el 62,3% restante valoró su desempeño y fue valorado a través de la aplicación informática con el cuestionario con una escala de respuesta MSS. De forma que para cada participante se contaba con dos medidas: su auto-evaluación del desempeño, y la evaluación realizada por su supervisor.

Los resultados de esta investigación mostraron una tendencia menor a asignar puntuaciones altas en la evaluación del desempeño cuando se utiliza la escala de respuesta MSS frente a la escala gráfica tradicional. Es decir, se reduce el sesgo de benevolencia. Los supervisores, en comparación con las autoevaluaciones, asignan puntuaciones más bajas, tanto cuando se utiliza la escala gráfica tipo Likert como cuando se evalúa con la escala MSS.

Además, se observa una mayor desviación típica utilizando la escala MSS en comparación con la tipo Likert. Esto es, quienes utilizan la escala MSS asignan en mayor medida puntuaciones diferentes dependiendo de la tarea concreta que estén evaluando, reduciéndose la tendencia de los evaluadores a evaluar de forma similar el desempeño del trabajador en cada una de las tareas. Las evaluaciones de los supervisores también presentan mayor variabilidad de respuesta en sus evaluaciones del desempeño cuando usan la escala MSS frente a la escala gráfica.

Estos datos apoyan nuestra hipótesis de que utilizar la escala MSS reduce el sesgo de benevolencia y favorece una mayor variabilidad de respuestas intra-sujeto que cuando se usa la escala tipo Likert. No obstante, este primer estudio presenta dos limitaciones importantes desde nuestro punto de vista, el reducido tamaño de la muestra, y que las muestras que evaluaron el desempeño con los dos tipos de escala eran diferentes, es

decir, se utilizó un diseño inter-sujeto. Con el objeto de subsanar estas deficiencias y confirmar los resultados obtenidos se llevó a cabo la segunda investigación.

El objetivo de este segundo estudio fue analizar si existen diferencias significativas en la evaluación del desempeño de una misma persona cuando se utiliza una escala de conductas observadas con respuesta en escala gráfica tipo Likert, frente a una escala con anclajes conductuales tipo estandarizada mixta (Hernández-Fernaud et al., 2013). Al igual que en la investigación anterior se elaboraron dos instrumentos que contenían las mismas conductas a evaluar, ocho tareas necesarias para realizar el trabajo en grupo, y que variaban solo en el tipo de escala de respuesta.

La primera fase de esta investigación se realizó con 68 estudiantes universitarios que habían realizado una actividad de prácticas en pequeño grupo, y que era el objeto de la evaluación del desempeño. Se utilizó un diseño de medidas repetidas, de forma que cada participante autoevaluaba su desempeño en el trabajo en grupo dos veces, una utilizando la escala gráfica, y otra usando la escala MSS. Además, se consideró el orden de cumplimentación de las escalas como variable, de manera que la mitad de la muestra contestó en primer lugar la escala gráfica y un mes después la MSS, y la otra mitad de la muestra en el orden contrario. Ambas escalas se presentaban mediante una aplicación informática. Los resultados mostraron un efecto estadísticamente significativo del tipo de escala, de forma que cuando se utiliza la escala MSS la autoevaluación del desempeño es menor que cuando se utiliza una escala gráfica tipo Likert. También se observó que el orden de cumplimentación de las escalas no modificaba los resultados.

En la segunda fase de esta investigación se amplió la muestra a 188 participantes, sin tener en cuenta el orden de cumplimentación de los dos instrumentos. Los resultados fueron en la misma línea que los anteriores, la media en autoevaluación del desempeño es significativamente mayor utilizando la escala gráfica tipo Likert que usando la escala MSS, por lo que ésta disminuye el sesgo de benevolencia. Asimismo, al igual que en la primera investigación comentada, se observó que la desviación típica media que se obtiene con la escala gráfica es menor que la obtenida para la escala MSS, de forma que los participantes puntuaron en un rango mayor sus respuestas con ésta última que utilizando la escala gráfica.

Los resultados que hemos comentado indican que utilizar instrumentos con escalas de respuesta que presenten anclajes conductuales sin una identificación explícita del nivel de desempeño que representan, disminuye la tendencia a la evaluación positiva del desempeño, y favorece que el evaluador utilice más puntos de la escala de respuesta que cuando se utiliza una escala gráfica tradicional, disminuyendo así el sesgo de benevolencia en las evaluaciones del desempeño de tarea.

CONCLUSIONES E IMPLICACIONES PRÁCTICAS

La evaluación del desempeño laboral es una estrategia indispensable en la gestión de las personas en las organizaciones, siendo especialmente relevante en situaciones de crisis económica como la actual. La evaluación del desempeño cumple un doble papel en las organizaciones. Por un lado, contribuye a la mejora de las organizaciones al facilitar el logro de las metas organizacionales de eficacia, eficiencia y transparencia, ya que, entre otros aspectos, permite rediseñar los puestos de trabajo, planificar los planes de formación y los procesos de selección, o tomar decisiones sobre la política de promociones e incentivos. Por otro lado, incide en la motivación de los trabajadores a través de la participación de los mismos en el proceso, a través de la retroalimentación recibida respecto a su desempeño, que les permite mejorarlo, ser reconocidos y recompensados.

Para lograr estos efectos positivos de la evaluación del desempeño es importante tener en mente dos objetivos. Como plantean Farr y Jacobs (2006) es necesario que la organización y los participantes en la misma valoren positivamente las evaluaciones de desempeño, lo que dependerá de que existan una serie de condiciones. En concreto, la organización valorará positivamente contar con datos precisos y fiables del desempeño, hecho que dependerá del uso de escalas precisas de evaluación y de que se reduzcan los sesgos de los evaluadores (objetivos de la metáfora del test). Asimismo, los participantes valorarán más positivamente las evaluaciones si éstas son justas y creíbles, y ello dependerá, a su vez, de la precisión de las evaluaciones (metáfora del test y del due-process). Además, la justicia percibida y, por tanto, el éxito y la efectividad del sistema de evaluación, estarán supeditados a que: (a) el procedimiento utilizado reconozca las diferentes metas y motivaciones individuales en juego; (b) el sistema a emplear ofrezca información clara a los participantes; y (c) se posibilite la participación de evaluadores y evaluados en el proceso de diseño del sistema.

REFERENCIAS

- Aguinis, H. (2013). *Performance management*. Upper Saddle River, New Jersey: Pearson
- Bernardin, H.J., Cooke, D.K., & Villanova, P. (2000). Conscientiousness and Agreeableness as predictors of rating leniency. *Journal of Applied Psychology, 85*(2), 232-234. <http://dx.doi.org/10.1037//0021-9010.85.2.232>
- Borman, W. C., & Motowidlo, S. (1993). Expanding the criterion domain to include elements of contextual performance. En N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations*, (pp. 71-98). San Francisco: Jossey - Bass.
- Bretz, R. D., Milkovich, G. T., & Read, W. (1992). The current state of performance appraisal research and practice: Concerns, directions, and implications. *Journal of Management, 18*, 321-352. <http://dx.doi.org/10.1177/014920639201800206>
- Cardy, R. L., & Dobbins G. H. (1994). *Performance appraisal: alternative perspectives*. South-Western, Cincinnati: OH.
- Cawley, B. D., Keeping, L. M., & Levy, P. E. (1998). Participation in the Performance Appraisal Process and Employee Reactions: A meta-analytic review of field investigations. *Journal of Applied Psychology, 83*(4), 615-633. <http://dx.doi.org/10.1037//0021-9010.83.4.615>
- Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O. L. H., & Yee Ng, K. (2001). Justice at the Millennium: A Meta-Analytic Review of 25 years of Organizational Justice Research. *Journal of Applied Psychology, 86*(3), 425-445. <http://dx.doi.org/10.1037/0021-9010.86.3.425>
- Colquitt, J. A., Scott, B. A., Rodell, J. B., Long, D. M., Zapata, C. P., Colon, D. E., & Wesson, M. J. (2013). Justice at the Millennium, a Decade Later: A Meta-Analytic Test of Social Exchange and Affect-Based Perspectives. *Journal of Applied Psychology, 98*(2), 199-236. <http://dx.doi.org/10.1037/a0031757>
- Díaz-Cabrera, D., Díaz-Vilela, L., Delgado, N., Isla-Díaz, R., Rosales-Sánchez, Ch. y Hernández-Fernaud, E. (2013). *Comportamiento de las distribuciones de medidas del desempeño en función del tipo de escala utilizada*. Trabajo presentado en el XIII Congreso de Metodología de las Ciencias Sociales y de la Salud (AEMCCO), La Laguna, España.
- Díaz-Cabrera, D., Hernández-Fernaud, E., Rosales-Sánchez, Ch., Isla, R. y Díaz-Vilela, L. (en revisión). Aceptación y justicia percibida de la evaluación del desempeño: análisis de variables predictoras.
- Díaz-Vilela, L., Delgado, N., Isla-Díaz, R., Díaz-Cabrera, D. & Hernández-Fernaud, E. (under review). Relationships between contextual and task performance and interrater agreement: Why there are none?.
- Díaz-Vilela, L., Díaz-Cabrera, D., Isla-Díaz, R., Hernández-Fernaud, E., Rosales-Sánchez, Ch. & Delgado, N. (2013). *Development of a computer based task performance scale for clerical workers in a Public Administration*. 16th Congress of the European Association of Work and Organizational Psychology (EAWOP), Münster, Germany.
- Dirks, K. T. & Ferrin, D. L. (2001). The role of trust in organizational settings. *Organization Science, 12*, 450-467. <http://dx.doi.org/10.1287/orsc.12.4.450.10640>
- Farr, J. L & Jacobs, R. (2006). Trust us: New perspectives on performance appraisal. En W. Bennett, C. Lance & D. Woehr (Eds.), *Performance Measurement: Current perspectives and future challenges* (pp. 321-337). Mahwah, New Jersey: LEA, Publishers.
- Fletcher, C. (2002). Appraisal: An individual psychological perspective. En S. Sonnentag (Ed.), *Psychological management of individual performance* (pp. 115-136). New York, NY: John Wiley & Sons, Ltd.
- Folger, R., Konovsky, M., & Cropanzano, R. (1992). A due process metaphor for performance appraisal. En B. Staw &



DOLORES DÍAZ CABRERA, ESTEFANÍA HERNÁNDEZ FERNAUD,
ROSA ISLA DÍAZ, NAIRA DELGADO RODRÍGUEZ, LUIS DÍAZ VILELA Y
CHRISTIAN ROSALES SÁNCHEZ

- L. Cummings (Eds.), *Research in organizational behavior*, Vol. 14 (pp. 127-148). Greenwich: JAI Press.
- Giles, W. F., & Mossholder, K. W. (1990). Employee reactions to contextual and session components of performance appraisal. *Journal of Applied Psychology*, 75, 371-377. <http://dx.doi.org/10.1037//0021-9010.75.4.371>
- Guilford, J. P. (1954). *Psychometric method*. Nueva York: McGraw-Hill.
- Hernández Fernaud, E., Díaz-Vilela, L., Isla-Díaz, R., Delgado, N., Rosales-Sánchez, Ch. y Díaz-Cabrera, D. (2013). *Efecto del tipo de escala de medida en la evaluación del desempeño en estudiantes universitarios*. Trabajo presentado en el XIII Congreso de Metodología de las Ciencias Sociales y de la Salud (AEMCCO), La Laguna, España.
- Keeping, L. M., & Levy, P. E. (2000). Performance appraisal reactions: measurement, modeling, and method bias. *Journal of Applied Psychology*, 85, 708-723. <http://dx.doi.org/10.1037//0021-9010.85.5.708>
- Levy, P. E., & Williams, J. R. (2004). The social context of performance appraisal: A review and framework for the future. *Journal of Management*, 30(6), 881-905. doi:10.1016/j.jm.2004.06.005
- Murphy, K. R., & Cleveland, J. N. (1995). *Understanding performance appraisal: Social, organizational and good-based perspectives*. Thousand Oaks: CA: Sage.
- Poropat, A. E. (2002). New Models of Work Performance and Their Implications for Employment Relations. En P. Ross, M. Lyons, C. Allan & K. Townsend (Eds.), *Proceedings of the Tenth Annual Conference of the International Employment Relations Association*. Sydney: International Employment Relations Association (IERA).
- Salgado, J. F. y Cabal, A. L. (2011). Evaluación del desempeño en la Administración Pública del Principado de Asturias: Análisis de las propiedades psicométricas. *Revista de Psicología del Trabajo y de las Organizaciones*, 27(2), 75-91.
- Salleh, M., Amin, A., Muda, S., & Halim, M. A. S. A. (2013). Fairness of performance appraisal and organizational commitment. *Asian Social Science*, 9(2), 121-128. <http://dx.doi.org/10.5539/ass.v9n2p121>
- Smith, P. (1976). Behavior, results, and organizational effectiveness: The problem of criteria. En M. D. Dunnette (Ed.), *Handbook of Industrial and organizational psychology* (pp. 745-775). Chicago: Rand McNally College Pub. Co.
- Tziner, A., Murphy, K. R., & Cleveland, J. N. (2001). Relationships between attitudes toward organizations and performance appraisal systems and rating behavior. *International Journal of Selection and Assessment*, 9(3), 226-239. <http://dx.doi.org/10.1111/1468-2389.00176>
- Viswesvaran, C., & Ones, D. S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment*, 8(4), 216-226. doi:<http://dx.doi.org/10.1111/1468-2389.00151>
- Viswesvaran, C., & Ones, D. S. (2005). Job performance: Assessment issues in personnel selection. En A. Evers, N. Anderson & O. Voskuijl (Eds.), *The Blackwell Handbook of Personnel Selection*, (pp. 354-375). Malden: Blackwell Publishing.
- Woehr, D. J., & Huffcutt, A. I. (1994). Rater training for performance appraisal: A quantitative review. *Journal of Occupational and Organizational Psychology*, 67, 189-205.

