



Artículo

Investigaciones sobre la Entrevista Conductual Estructurada (ECE) en la Selección de Personal en la Administración General del País Vasco: Meta-análisis de la Fiabilidad¹

A study of Structured Behavioural Interview in Personnel Selection in the Basque Country General Administration: A meta-analysis of reliability

JESUS F. SALGADO

Universidad de Santiago de Compostela

SILVIA MOSCOSO

Universidad de Santiago de Compostela

MIKEL GORRITI

Oficina para la Modernización Administrativa, Gobierno Vasco

RESUMEN

Este artículo tiene por objetivos: (1) presentar las entrevistas conductuales estructuradas (ECE) como una técnica apropiada para la selección de personal en las organizaciones tanto privadas como públicas; (2) describir el proceso de implementación de las entrevistas conductuales estructuradas en la Administración General del País Vasco (AGPV) y mostrar

¹ Las opiniones expresadas en este artículo son de los autores del mismo y no necesariamente representan la opinión de la Administración General del País Vasco y sus gestores. Los autores quieren agradecer a las autoridades de la Consejería de Hacienda y Administración Pública del Gobierno Vasco y en especial a la Vicoconsejería de Función Pública su apoyo para la realización del presente artículo. También quieren mostrar su agradecimiento a todos los entrevistadores homologados del Gobierno Vasco, que con su esfuerzo, dedicación y rigor han hecho posible la puesta en práctica del procedimiento descrito en el artículo: Cecilia Anitua, Edurne Barañano, Carmelo Basoredo, Carlos Izaguirre, Rafael Jungitu, Santiago Martínez, Miren Ortega, Koldobike Uriarte, Imanol Urigüen, Arantza Irusta. Copias de este artículo pueden solicitarse a Jesús F. Salgado, Departamento de Psicología Social y Básica, Universidad de Santiago de Compostela, 15782 Santiago de Compostela (e-mail: psjesal@usc.es), a Silvia Moscoso en la misma dirección postal (e-mail: psilvia@usc.es) o a Mikel Gorriti, Oficina de Modernización Administrativa, Vicepresidencia del Gobierno, c/ Donostia San Sebastián 1, 01010 Vitoria-Gasteiz (e-mail: M-Gorriti@ej-gv.es). Los tres autores contribuyeron por igual a artículo.

cómo instrumentos desarrollados inicialmente en el ámbito de las organizaciones privadas pueden utilizarse sin merma de calidad ni de garantías procesales en las organizaciones públicas, tales como la administración general del estado; y (3) presentar los resultados de la evidencia empírica disponible en la actualidad, después de un año de utilización, sobre la fiabilidad de los entrevistadores y de los procesos de entrevista llevados a cabo en la Administración General del País Vasco. Dos meta-análisis fueron llevados a cabo con los resultados de la investigación realizada sobre 20 diseños de entrevista diferentes. La evidencia disponible indica que la fiabilidad de los entrevistadores usando la entrevista conductual estructurada en la AGPV es de .83 y que la fiabilidad de las decisiones del panel de entrevistadores (3 ó 4 entrevistadores) es de .93. Estos hallazgos superan ampliamente los resultados obtenidos en investigaciones previas y sus implicaciones para la práctica de la selección de personal en la Administración Pública son comentadas.

ABSTRACT

This paper has a number of goals: 1) introducing structured behavioral interview as a very appropriate tool for personnel selection in organizations whether public or private; 2) describing the implementation process of structured behavioral interviews in the General Administration of the Basque Country, showing how tools that have been developed for the private sector can also be used in public organizations without detriment to quality nor to procedural warranties ;and 3) reporting evidence one year after about reliability of interviewers and interview process carried out in the Basque Country General Administration. Two meta-analyses were carried out with results of research about 20 different interview designs. Evidence suggests that interviewer reliability using structured behavioral interview is .83 and reliability of interviewer panel (3 or 4) is .93. These findings go far beyond those of prior research. Implications for personnel selection practice in the public administration are discussed.

PALABRAS CLAVE

Entrevista Conductual Estructurada, Selección de Personal, Meta-análisis, Administración Pública.

KEY WORDS

Structured Behavioral Interview, Personnel Selection, Reliability, Validity, Public Administration.

Este artículo tiene tres objetivos relacionados. En primer lugar, introducir las entrevistas conductuales estructuradas (ECE) como una técnica apropiada para la selección de personal en las organizaciones tanto privadas como públicas. En segundo lugar, describir el proceso de implementación de las entrevistas conductuales estructuradas en la Administración General del País Vasco (AGPV) y mostrar cómo instrumentos desarrollados inicialmente en el ámbito de las organizaciones privadas pueden utilizarse sin merma de calidad ni de garantías procesales en las organizaciones públicas, tales como la Administración Pública del Estado. El tercer objetivo es presentar los resultados de la evidencia empírica disponible en la actualidad, después de un año de utilización, sobre la fiabilidad de los entrevistadores y de los procesos de entrevista llevados a cabo en la Administración General del País Vasco. Por conveniencia expositiva, hemos dividido el artículo en tres partes. La primera esta destinada describir lo que son las entrevistas conductuales estructuradas de selección de personal y los resultados de la investigación llevada a cabo sobre sus propiedades psicométricas. En la segunda parte mostramos cómo se ha llevado a cabo el proceso de implementación de este tipo de entrevistas en la Administración General del País Vasco, indicando como se procede desde que existe una vacante hasta que dicha vacante es cubierta mediante el oportuno concurso en el que se utiliza la entrevista conductual estructurada. También mencionamos en esta parte los preceptos legales en los que se basa la utilización de las ECE en la Administración General del País Vasco. En la tercera parte, que dedicamos a la investigación reciente, presentamos tres estudios empíricos llevados a cabo con los datos obteni-

dos a partir de la aplicación de la ECE en la Administración General del País Vasco. Por último, cerraremos el artículo con una valoración global de la entrevistas conductuales estructuradas y de la experiencia llevada a cabo en este último año, así como con sugerencias tanto de investigación como prácticas para un futuro inmediato.

LA ENTREVISTA CONDUCTUAL ESTRUCTURADA DE SELECCIÓN DE PERSONAL

Schmidt y Hunter (1998) en su revisión de 85 años de investigación en selección de personal concluyen que las entrevistas de selección y, más concretamente, las entrevistas estructuradas, son uno de los mejores predictores del desempeño laboral y que su validez se generaliza a través de puestos, criterios y organizaciones. A las mismas conclusiones llega Salgado (1999; Salgado, Viswesvaran y Ones, 2001) en una revisión de la investigación reciente sobre los métodos de selección de personal. Este dato es de especial relevancia si tenemos en cuenta que la entrevista es el método de selección de personal más utilizado por los profesionales de recursos humanos en la mayoría de los países (Salgado, Viswesvaran y Ones, 2001). En todos, o casi todos, los programas de selección que se realizan en las organizaciones privadas y en muchas públicas, aunque se incluyan otros instrumentos, la entrevista constituye la parte central y la decisión de contratación se toma la mayoría de las veces únicamente a partir de ella. Incluso en muchas ocasiones es el único método que se utiliza para evaluar a los candidatos. Estudios realizados en Europa, América y Asia muestran que la entrevista se utiliza en más del 90% de los casos.

Sin embargo, aunque se suele hablar de la entrevista de selección en general, lo cierto es que existen distintos tipos de entrevistas, con importantes diferencias entre ellos. Así, podemos clasificar las entrevistas en dos grandes categorías: (a) en función de su contenido y (b) en función de su formato o grado de estructura. A su vez estas diferencias hacen que su frecuencia de uso y sus propiedades psicométricas varíen mucho. Janz (1982; Janz, Hellervik y Gilmore, 1986) desarrolló una taxonomía sobre las preguntas de la entrevista y las clasifica en cinco grandes grupos: (A) credenciales, logros y datos biográficos, (B) conocimiento técnico, (C) descripciones sobre la experiencia, (D) información autoevaluativa (p.e., puntos fuertes y débiles, gustos, actitudes, metas) y (E) descripciones de conducta (p.e., cuestiones relacionadas con la conducta pasada o futura). A partir de esta clasificación de las preguntas, podemos dividir las entrevistas en dos tipos: entrevistas convencionales y entrevistas conductuales estructuradas. Janz (1982) encontró que lo que nosotros llamamos entrevistas convencionales estaban compuestas básicamente de preguntas del tipo A (15%), C (32%) y D (48%), y son el tipo de entrevistas que se suelen utilizar en la gran mayoría de los procesos de selección, aunque su rigor en términos de estructura varía mucho de unos procesos y organizaciones a otros. En algunas organizaciones, la entrevista convencional dura muy pocos minutos y las preguntas no son las mismas para todos los candidatos, mientras que en otras la duración suele ser amplia y se utiliza el mismo repertorio de preguntas para todos los entrevistados. A las primeras se les puede denominar entrevistas convencionales no estructuradas o semiestructuradas y a las últimas entrevistas convencionales estruc-

turadas. Por su parte, las entrevistas conductuales estructuradas se centran en preguntas sobre conocimientos técnicos y, sobre todo, descripciones de conducta (40% y 33% respectivamente).

Cuando hablamos de las entrevistas en función de su formato o grado de estructuración nos podemos centrar en dos cuestiones: (a) el grado de estructuración de las preguntas de la entrevista y (b) el grado de estandarización de la valoración de las respuestas. Con respecto a la primera cuestión, Huffcutt y Arthur (1994) y Campion, Palmer y Campion (1997) plantean dos taxonomías muy similares basadas en cuatro niveles. El primer nivel se caracterizaría por una total ausencia de restricciones formales, el entrevistador puede plantear las preguntas que él considere oportunas durante el transcurso de cada una de las entrevistas que lleve a cabo. El segundo nivel estaría formado por aquellas entrevistas en las cuales se le proporciona al entrevistador una guía de las áreas a cubrir durante la misma (p.e. las utilizadas por Fear, 1984, Ghiselli, 1966 o Morgan y Cogger, 1974). El tercer nivel se caracterizaría por una especificación previa de las preguntas, aunque los entrevistadores tienen flexibilidad para adaptar la entrevista a los diferentes candidatos y, también, se les permite a los entrevistadores elegir entre cuestiones alternativas y sondear a los solicitantes sobre cuestiones específicas. Un ejemplo de este nivel sería la entrevista de descripción de conducta de Janz (1982). El cuarto nivel implica una estandarización completa, es decir, a los solicitantes se les plantean exactamente las mismas cuestiones y en el mismo orden. Aquí se incluiría la entrevista situacional de Latham, Saari, Pursell y Campion (1980) y la ECE (Salgado y Moscoso, 2001).

La valoración de las respuestas del candidato también es susceptible de dividirse en varios niveles progresivamente más estandarizados (Huffcutt y Arthur, 1994; Campion, Palmer y Campion, 1997). El primer nivel estaría formado por aquellas entrevistas en las cuales el entrevistador hace una evaluación global del candidato, basada en la información total obtenida. Este es el tipo de valoración que se hace con las entrevistas convencionales (p.e. la utilizada por Ghiselli, 1966). El segundo nivel lo formarían aquellas entrevistas en las cuales se realizan múltiples valoraciones referidas a cada una de las dimensiones en las que están basadas las preguntas de la entrevista o a diferentes rasgos (por ejemplo el modelo de entrevista de Janz, 1982). Y por último el tercer nivel se caracterizaría por la valoración independiente de cada una de las respuestas dadas por los solicitantes a las preguntas en diferentes escalas adaptadas a cada una de esas cuestiones (un ejemplo es la entrevista situacional y la ECE).

Las importantes diferencias entre los distintos tipos de entrevistas se han puesto también de manifiesto en la investigación realizada sobre sus propiedades psicométricas. Desde la década de 1940, periódicamente se han hecho revisiones de la investigación sobre la validez de la entrevista de selección de personal. Hasta el año 1987, se habían realizado 7 grandes revisiones de la literatura, incluyendo meta-análisis, y todas ellas concluían unánimemente afirmando que la entrevista se caracterizaba por tener una baja fiabilidad, es decir, que dos entrevistadores diferentes no coincidían apenas en sus apreciaciones de un candidato, y, en segundo lugar, que su validez era escasa o incluso nula y que no aportaba nada a la pre-

dicción del rendimiento obtenida mediante otros instrumentos como por ejemplo los tests de habilidades cognitivas (Arvey y Campion, 1982; Mayfield, 1964; Schmitt, 1976; Ulrich y Trumbo, 1965; Wagner, 1949; Wright, 1969). A resultados semejantes llegaron los meta-análisis realizados por Hunter y Hunter (1984) y Reilly y Chao (1982).

La mayoría de las entrevistas que se habían utilizado hasta este momento en la investigación y en la práctica de la selección de personal eran entrevistas no estructuradas o semiestructuradas de tipo convencional y en las que, en muchas ocasiones, se trataba de diagnosticar la personalidad del entrevistado. En estas entrevistas, el entrevistador posee una gran libertad para decidir qué y cómo preguntar. Esto lleva a que las cuestiones planteadas muchas veces poco tengan que ver con los requisitos del puesto en cuestión. Para un mismo puesto no hay acuerdo entre los entrevistadores con respecto a cual sería la mejor estrategia a seguir o cual sería el contenido adecuado de las cuestiones que se le van a plantear al candidato. Incluso un mismo entrevistador puede modificar tanto la estrategia como el contenido ante diferentes candidatos para un mismo puesto.

A partir de 1988 se llevaron a cabo una serie de meta-análisis sobre la fiabilidad y la validez de la entrevista y cuyos resultados varían sustancialmente de los encontrados hasta este momento. Así, Weisner y Cronshaw (1988) llevaron a cabo un meta-análisis sobre la validez predictiva de la entrevistas y dividieron éstas en dos categorías: estructuradas y no estructuradas. Sus resultados muestran una validez corregida de .31 para las entrevistas no estructu-

radas y de .62 para las entrevistas estructuradas. Por su parte, McDaniel, Whetzel, Schmidt y Maurer (1994) llevaron a cabo el meta-análisis más amplio realizado hasta la fecha sobre la validez predictiva de las entrevistas de selección. Estos investigadores consideraron dos criterios para clasificar las entrevistas: el contenido y la estructura. Con respecto al contenido, dividieron las entrevistas en tres categorías: situacionales, relacionadas con el puesto y entrevistas psicológicas. Encontraron una validez de .50, .39 y .29 respectivamente para cada una de las entrevistas. Cuando clasificaron las entrevistas según su estructura, los autores encontraron un coeficiente de validez de .44 para las entrevistas estructuradas y de .33 para las no estructuradas. En el mismo año, Huffcutt y Arthur (1994) llevaron a cabo otro meta-análisis en el que dividían las entrevistas en cuatro niveles según su grado de estructura. Sus resultados muestran unos coeficientes de validez que oscilan entre .20 para las entrevistas con el grado de estructura más bajo, hasta .56 para las entrevistas con mayor estructuración. Por esas mismas fechas, Salgado y Moscoso (1995) realizaron un meta-análisis sobre la validez predictiva de las entrevistas conductuales estructuradas (entrevistas de descripción de conducta, situacionales, relacionadas con el puesto y conductuales estructuradas) y encontraron una validez observada de .28 que, cuando se corrige por atenuación resulta una validez operativa de .57, la misma obtenida por Huffcutt y Arthur. Más recientemente, Taylor y Small (2002) llevaron a cabo una investigación en la que hacían una comparación meta-analítica de la validez predictiva de las entrevistas situacionales (Latham, Saari, Pursell y M. Campion, 1980; Latham, 1989) y las entrevistas de descripción de conducta (Janz,

1982, 1989). Sus resultados muestran unos coeficientes de validez de .45 y .56 para las entrevistas situacionales y de descripción de conducta respectivamente. Cuando dividen las entrevistas de descripción de conducta en función de la utilización en las mismas de escalas de valoración con anclajes conductuales (EVAS) o no, los coeficientes de validez encontrados son de .63 cuando utilizan EVAS y .47 cuando no las utilizan. Por lo tanto, parece que la utilización de escalas de valoración con anclajes conductuales es un moderador de la validez de estos tipos de entrevistas.

Con respecto a la fiabilidad, en el caso de la entrevista el método más adecuado es la fiabilidad interevaluador. En este sentido, Conway, Jako y Goodman (1995) llevaron a cabo el más completo meta-análisis realizado hasta ahora sobre la fiabilidad de las entrevistas de selección. Estos autores dividieron las entrevistas en cinco niveles en función de su grado de estructuración y encontraron un coeficiente de fiabilidad interevaluador promedio de .70. Los autores diferenciaron entre entrevistas de panel (varios entrevistadores en la misma entrevista) y entrevistas separadas o seriales (más de un entrevistador entrevista a un candidato en momentos diferentes). Para las entrevistas de panel los coeficientes de fiabilidad oscilaban entre .69 para el nivel más bajo de estructuración y .90 para el nivel más alto. Por su parte, las entrevistas seriales mostraban un rango que iba desde .37 hasta .66 (éste último se corresponde con el nivel cuarto de estructuración).

En esta investigación, Conway, Jako y Goodman realizan también un meta-análisis de los coeficientes de consistencia interna. Este tipo de coeficiente es apropiado esti-

marlo sólo si los entrevistadores evalúan a los solicitantes en múltiples preguntas o ítems que están midiendo la misma dimensión (si se utiliza un número considerable de preguntas o ítems para medir, por ejemplo, habilidades interpersonales). El cálculo de la consistencia interna puede ser especialmente útil en el caso de las entrevistas altamente estructuradas, donde se utiliza una valoración independiente para cada pregunta, para saber el número de ítems o preguntas necesarios para incluir en la entrevista, es decir, un coeficiente alfa muy alto podría indicar que se están incluyendo preguntas redundantes. Los resultados del meta-análisis de Conway, Jako y Goodman (1995), muestran que las entrevistas con apenas grado de estructuración, con únicamente dos valoraciones ya alcanzan un nivel alfa de .85, sin embargo, las entrevistas con un alto grado de estructuración presentan unos coeficientes alpha bajos aún utilizando un mayor número de valoraciones. Esto se traduce en que cuando la estandarización de las preguntas es muy baja, el número de valoraciones que podrían proporcionar información única es muy pequeño. Alternativamente, cuando la estructuración es muy alta, el número de valoraciones que podría continuar proporcionando información única es bastante alta.

Además de los anteriores estudios, también es de interés mencionar que en los últimos años, algunos autores han llevado a cabo estudios primarios sobre la fiabilidad de diferentes tipos de entrevistas conductuales estructuradas. Por ejemplo, Latham y sus colegas (1980) utilizando tres entrevistadores encontraron unos coeficientes de fiabilidad interevaluador de .76 ($n=49$) y .79 ($n=63$). Por su parte, Campion, Pursell y Brown (1988) informan de

un coeficiente de fiabilidad de .88 ($n=149$) utilizando tres entrevistadores. Janz (1982), comparando las entrevistas convencionales con las entrevistas de descripción de conducta, encontró un coeficiente de .71 para las entrevistas convencionales y de .46 para las denominadas entrevistas de descripción de conducta. Una explicación de estos resultados puede ser que las condiciones asignadas a la entrevistas no fueron aleatorias. Salgado y Moscoso (1995) en una investigación meta-analítica en la que tomaban conjuntamente los datos existentes sobre la fiabilidad interevaluador de las entrevistas conductuales estructuradas, encontraron un coeficiente promedio de .75. En su reciente investigación meta-analítica, Taylor y Small (2002) encontraron un coeficiente de fiabilidad de .79 para las entrevistas situacionales y de .76 para las entrevistas de descripción de conducta. Cuando dividían éstas últimas en función de si utilizaban escalas de valoración con anclajes conductuales o no, los coeficientes de fiabilidad fueron de .76 y .73 respectivamente. De todos modos estos valores están calculados a partir de número no muy elevado de coeficientes, 8 en el caso de las entrevistas que utilizaban EVAS y 5 que no las utilizaban.

Resumiendo, hoy día no existen dudas sobre la validez predictiva de la entrevista de selección y sobre su fiabilidad. Concretamente, la investigación meta-analítica ha mostrado que las entrevistas estructuradas presentan un coeficiente de validez similar al de los mejores instrumentos de selección (tests de habilidad cognitiva, tests de muestras de trabajo, etc.) y su fiabilidad interevaluador es más que aceptable. Pero dentro de lo que denominamos entrevistas conductuales estructuradas, si categorizamos las entrevistas en función de su con-

tenido, o entrevistas estructuradas (cuando hablamos del formato), existen varias categorías. Así, hemos mencionado que existen meta-análisis que analizan la validez y la fiabilidad de cada uno de estos tipos de entrevistas. Sin embargo, aún no existen demasiados estudios primarios como para ser concluyentes por ejemplo, con respecto a la fiabilidad interevaluador de las entrevistas situacionales y las entrevistas de descripción de conducta. La validez de cualquier instrumento depende de su fiabilidad y aquí podría estar la clave o parte de la explicación de por qué un tipo de ECE muestran mayor validez predictiva que otras. Por ejemplo, en el meta-análisis de Taylor y Small (2002) las entrevistas de descripción de conducta presentan un coeficiente de validez sensiblemente mayor que las entrevistas situacionales, sin embargo, el número de estudios incluidos en el meta-análisis es bastante reducido.

Características formales de las Entrevistas Conductuales Estructuradas

Las entrevistas conductuales estructuradas (ECE) fueron desarrolladas en la década de los ochenta por varios psicólogos. Estas entrevistas presentan una serie de características que controlan o eliminan muchos de los sesgos que se producen con las entrevistas convencionales y hacen que se consigan elevados coeficientes de validez predictiva. Existen varias alternativas que se podrían englobar en esta categoría general denominada "entrevista conductual estructurada" (Motowidlo, Carter, Dunnette, Tippins, Werner, Burnett, Vanghan, 1992). Se incluiría bajo esta etiqueta la entrevista "estructurada desarrollada a partir de un análisis del puesto (EEAP)"

(Campion, Pursell y Brown, 1988; Wright, Lichtenfels y Pursell, 1989), la entrevista "situacional (ES)" (Latham, Saari, Pursell y M. Campion, 1980; Latham, 1989), la entrevista "estructurada de descripción de conducta (EEDC)" (Janz, 1982, 1989) o la entrevista multimodal (Schuler, 1989), la entrevista conductual estructurada (ECE) (Salgado y Moscoso, 2001). Cada una de las anteriores entrevistas presenta una serie de características particulares, sin embargo, también es cierto que el desarrollo, aplicación y evaluación en las mismas sigue un modelo común. Campion, Palmer y Campion (1997) hacen una revisión sobre la estructura de las entrevistas de selección e identifican 15 componentes diferenciables que se podrían resumir en las siguientes seis características:

(1) La característica más destacable y general es que las preguntas que se utilizan mediante este tipo de entrevistas se centran en las conductas eficaces para el desempeño en el puesto. Es decir, se evita toda información que sea irrelevante para poder predecir el futuro rendimiento del candidato en el puesto. Así el primer paso en el desarrollo de este tipo de entrevistas es realizar un análisis del puesto. El objeto del mismo es conocer las habilidades, las destrezas y los requisitos necesarios e importantes para un rendimiento adecuado en el puesto. Esta información va a servir para desarrollar las dimensiones relevantes del puesto y las preguntas que se derivarían de esas dimensiones. En la mayoría de las entrevistas se utiliza la técnica de los incidentes críticos (Flanagan, 1954), este método consiste en recoger ejemplos y anécdotas de conductas adecuadas e inadecuadas para el rendimiento en el puesto y representativas de las tareas y situaciones requeridas en el mismo.

(2) Una segunda característica de las entrevistas conductuales estructuradas es que se hacen el mismo tipo de preguntas a todos los candidatos. Para poder elegir al mejor candidato es necesario tener la misma información de todos ellos para poder compararlos y decidir cuál es el mejor. Dependiendo del tipo de entrevista conductual estructurada que estemos utilizando estas preguntas pueden ser exactamente las mismas o pueden variar ligeramente en función de la información que vaya proporcionando el candidato. En cualquier caso el contenido de la información obtenida debe ser el mismo para todos los candidatos.

(3) Otra característica es que para la evaluación de las entrevistas se utilizan escalas de valoración con anclajes conductuales (EVAS) (Smith y Kendall, 1963; Quijano, 1992). Los anclajes conductuales son ejemplos específicos de conductas observables en un puesto que nos indican diferentes niveles de desempeño en una dimensión o competencia determinada. Estos anclajes ayudan al entrevistador a valorar las respuestas de los candidatos, haciendo más precisas sus valoraciones. Cada anclaje representaría un nivel determinado de desempeño.

(4) Se debe realizar el mismo proceso con todos los candidatos. El mismo entrevistador debe preguntar todas las cuestiones a cada uno de los candidatos. Si hay más de un entrevistador es aconsejable no comentar ninguna de las entrevistas hasta que todas sean completadas. Después es cuando se discuten las posibles discrepancias y las valoraciones se promedian. El entrevistador va tomando notas o grabando cada entrevista.

(5) Se debe entrenar a los entrevistadores. La aplicación de este tipo de entrevistas requiere un pequeño curso de formación en el que el entrevistador además de familiarizarse con las características del puesto, aprende a utilizar correctamente la entrevista, cómo realizar las preguntas y cómo evaluarlas.

(6) Por último, la decisión final de selección se toma después de haber realizado todas las entrevistas. Una de las críticas que se le hacía a las entrevistas convencionales era que los entrevistadores tomaban la decisión de contratación en los primeros minutos de la entrevista y esta decisión normalmente está sesgada (el entrevistador se precipita y hace sus juicios muy pronto). La solución a esto es extraer de la entrevista el momento de la decisión, y esto es lo que se hace en las entrevistas conductuales estructuradas, la decisión de selección se toma después de haber valorado todas las entrevistas. Las escalas de valoración que se utilizan proporcionan una puntuación total para cada candidato, los que tienen las puntuaciones más altas son los elegidos. Dependiendo de las características del puesto y del proceso se establece un determinado punto de corte que se considera la puntuación mínima aceptable para considerar a los candidatos aptos o no para contratarlos.

Los dos tipos de entrevistas conductuales estructuradas más utilizadas son la entrevista de descripción de conducta de Janz y la entrevista situacional de Latham. La entrevista de descripción de conducta se basa en el principio de que el mejor predictor de la conducta futura es la conducta pasada en similares circunstancias. Es decir, se trata de conocer el comportamiento del candidato en situaciones pasadas que

se relacione o sea lo más parecido posible a las conductas que el candidato deberá realizar en el puesto actual. Este principio se llama también principio de consistencia interna y presenta dos corolarios. El primer corolario dice que cuanto más reciente es la conducta pasada, mayor es su poder predictivo. El segundo afirma que cuanto mayor es el tiempo de existencia de la conducta, mayor es su poder predictivo. Como puede observarse, cada pregunta consta de dos partes. En primer lugar se plantea la situación sobre la que se quiere preguntar para localizar un ejemplo concreto en el pasado del candidato y que este se centre en él. Después, para conocer cómo se comportó el solicitante en esa situación y cuáles fueron las consecuencias de su conducta, se hacen las preguntas pertinentes. El entrevistador debe indagar en la conducta del candidato hasta tener un conocimiento amplio de la situación planteada y así poder valorar correctamente al candidato en la dimensión evaluada. Una variante de este tipo de entrevistas son las entrevistas conductuales estructuradas (ECE, Salgado y Moscoso, 2001) que además de utilizar preguntas de descripción de conducta, usan escalas de valoración con anclajes conductuales (EVAS) para valorar las respuestas.

Con respecto a las entrevistas situacionales desarrolladas por Latham y sus colegas (1980), están basadas en el principio de que las intenciones de una persona están relacionadas con su conducta, es decir, lo que una persona dice que va a hacer correlaciona altamente con lo que esa persona hará, por lo tanto el futuro rendimiento del solicitante puede predecirse si se conocen sus metas e intenciones. Esta premisa surge de la teoría del establecimiento de metas propuesta por Locke (1968; Locke y Latham, 1984). Estas pre-

guntas representan una situación hipotética que se puede presentar en el puesto y los entrevistados deben contar que es lo que ellos harían en esas situaciones. Estas entrevistas utilizan escalas de valoración con anclajes conductuales (EVAS) para la valoración de las respuestas de los candidatos.

Las entrevistas conductuales estructuradas presentan las siguientes ventajas: en primer lugar, poseen una validez predictiva elevada, similar a la de los mejores instrumentos de selección; en segundo lugar, al utilizar un sistema de valoración objetivo es mucho más imparcial y se pueden evitar muchos de los sesgos propios de la entrevista convencional. En tercer lugar, presenta una utilidad económica importante en términos de ganancia de productividad, es decir, coloca al personal más adecuado en los puestos, lo que reduce considerablemente el número de malas colocaciones y el coste que esto supone a medio o largo plazo.

A su vez, las entrevistas conductuales estructuradas presentan ciertas restricciones a su empleo sistemático como instrumento de selección. Algunas de las razones esgrimidas para justificar el elevado uso de la entrevista (ver Harris, 1998; Schneider y Schmitt, 1986), eran su facilidad de uso (cualquiera con apenas preparación podía realizarla), su generabilidad (podía emplearse prácticamente en cualquier puesto) y su escaso coste de producción (limitado al tiempo empleado en la realización práctica de la entrevista). Ninguno de estos factores es característico de la entrevista conductual estructurada. Aunque es posible que con un entrenamiento adecuado cualquier profesional pueda realizarla, se requiere una formación y experiencia

previa profunda tanto en su conducción como en su valoración, lo cual reduce considerablemente el número de posibles usuarios. En segundo lugar, dado que la entrevista conductual estructurada está condicionada a un puesto concreto, su generabilidad es mucho menor que los otros tipos de entrevistas. En tercer lugar, los costes de producción de una entrevista conductual estructurada son elevados, ya que requiere la realización de análisis de puestos sistemáticos (generalmente mediante la técnica de los incidentes críticos), precisa de un número elevado de expertos en el puesto y obliga a importantes análisis psicométricos previos a su validación y utilización.

LA ENTREVISTA CONDUCTUAL ESTRUCTURADA EN LA ADMINISTRACIÓN GENERAL DEL PAÍS VASCO

El segundo objetivo de este artículo es describir pormenorizadamente el proceso llevado a cabo en la Administración General del País Vasco para aplicar las entrevistas conductuales estructuradas en los procesos de cobertura de vacantes. Esta descripción trata de mostrar cómo el desarrollo y aplicación rigurosa de estas técnicas puede ser beneficioso para la administración pública, tanto en términos de eficiencia de los procesos de selección de personal como en términos de satisfacción de los participantes en el proceso, básicamente, los miembros del tribunal y los candidatos, pero también los expertos consultados y todos aquellos que directa o indirectamente se ven implicados en los procesos de selección. En cierto sentido la utilización de la ECE ha supuesto una "revolución" en el sistema español de

cobertura de vacantes y no existe en el momento presente ninguna otra administración pública, ni autonómica ni estatal, que tenga un sistema tan elaborado, público, riguroso y orientado al puesto de trabajo como el que actualmente se emplea en la Administración General del País Vasco.

La Orden de 26 de febrero de 2002, de la Consejera de Hacienda y Administración Pública, regula el procedimiento para la provisión transitoria, en comisión de servicios o como personal funcionario interino, de puestos de trabajo de la Administración General de la Comunidad Autónoma del País Vasco y de sus Organismos Autónomos (Boletín Oficial del País Vasco, nº 57 de 22 de marzo de 2002, pp 5130-5142) cita el Decreto 74/2001, de 2 de mayo, por el que se aprueba el Acuerdo regulador de las condiciones de trabajo para el año 2001 del personal funcionario al servicio de la Administración de la Comunidad Autónoma de Euskadi y de sus Organismos Autónomos que en su artículo 53, modifica la provisión transitoria de los puestos de trabajo sin ocupantes en régimen de comisión de servicios. La Orden afirma que entre los cambios fundamentales de este Decreto está "*la incorporación de criterios objetivos para la selección de las personas candidatos*" buscando "*la selección de las personas más adecuadas*". Criterios objetivos (igualdad) y persona más adecuada (eficacia) se plasman como los cambios fundamentales a incorporar a partir de entonces a los procesos selectivos de la AGPV. Pero no en todos los casos. La citada Orden en su introducción dice: "*Con la finalidad de seleccionar a los y las aspirantes más adecuadas, se introduce un sistema de selección basado en la valoración de méritos... Además, para los puestos del grupo A o A/B de*

nivel 25 o superior se valorarán específicamente las competencias requeridas para el desempeño del puesto. A través de la presente Orden se determinan los méritos e instrumentos de evaluación". Y la Resolución de 26 de febrero de 2002, del Viceconsejero de Función Pública, por la que se establecen los baremos de aplicación en el procedimiento para la provisión transitoria, en comisión de servicios o como personal funcionario interino, de puestos de trabajo de la Administración General de la Comunidad Autónoma del País Vasco y de sus Organismos Autónomos (Boletín Oficial del País Vasco, nº 57 de 22 de marzo de 2002), en su Anexo 1, circunscribe el uso de la ECE *"Para los puestos de nivel 28, 27 y aquellos de nivel 26 que dependan directamente de una o un Alto Cargo o de un puesto de cobertura mediante libre designación"*.

La decisión por la que se ha incorporado la ECE a los niveles 26 con mando (jefaturas en estructura organizativa) y niveles 27 y 28 (los niveles 29 y 30 son de libre designación) obedece a dos razones fundamentales. La primera a que ha sido conciencia de los gestores de la Función General del País Vasco que un instrumento tan potente sólo puede introducirse de manera paulatina, dejando que su práctica demuestre sus bondades y extienda su uso al resto de niveles para la provisión interna entre el personal funcionario (Concursos de Traslados y Comisiones de Servicios), como en el acceso externo mediante Ofertas Públicas de Empleo (OPE). La segunda, a que se ha constatado la ineficacia de la forma de proveer los puestos de nivel superior, exclusivamente, mediante un concurso de méritos. Los gestores de la Función General del País Vasco necesitaban un instrumento que asegurase la ade-

cuación de los seleccionados a puestos de alta complejidad técnica en donde las consecuencias de errar son mayores y donde la variabilidad del desempeño tiene consecuencias graves bien sean económicas (hasta un 70% según Schmidt y Hunter, 1998), legales o de imagen.

El Anexo 1 de la citada Resolución tiene otras dos referencias importantes para entender la trascendencia de la incorporación de la ECE: (a) *"Las competencias necesarias para el desempeño del puesto se valorarán en una entrevista conductual estructurada cuya puntuación máxima será de 10 puntos. Además, se podrá solicitar, en su caso, la realización de un proyecto relacionado con las funciones del puesto que será defendido ante la Comisión de Evaluación."*, y (b) *"No podrá resultar seleccionado ni seleccionada ningún candidato ni candidata que no obtenga como mínimo 5 puntos en la entrevista. Si la diferencia entre la puntuación máxima y mínima, otorgada por cada miembro de la Comisión de Evaluación, es superior a 2,5 puntos, ésta deberá revisar la valoración efectuada, hasta que dicha diferencia sea menor a 2,5 puntos."*

En el caso de que la provisión de los puestos de estos niveles se realice primero mediante la superación de un proyecto, generalmente una prueba práctica, y después mediante la superación de una entrevista, ambos excluyentes, los 10 puntos se prorratean entre las dos pruebas. En el caso, la mayoría, de que estos 10 puntos se otorguen exclusivamente mediante la superación de la ECE, el porcentaje de puntuación que aporta este instrumento a la puntuación total varía entre 16.39% para el caso de obtener un 5 y 32.78% para el caso de obtener un 10.

El Proceso de Desarrollo de la ECE en la Administración General del País Vasco

Tal y como se ha comentado más arriba, la ECE solamente se utiliza para cubrir puestos de determinados niveles en Comisiones de Servicio del personal funcionario de la AGPV. La Comisión de Servicios es un procedimiento para la provisión transitoria de puestos de trabajo regulado por el artículo 54 de la Ley 6/89, de la Función General del País Vasco regulado por la Orden y la Resolución arriba citadas.

El inicio de una Comisión de Servicios tiene lugar a partir de la solicitud de un Departamento cualquiera del Gobierno Vasco a la Dirección de Función Pública para la cobertura transitoria de una o varias dotaciones vacantes de un puesto de trabajo. La vacante suele producirse, normalmente, por la creación de nuevos puestos o dotaciones o porque el titular se ha trasladado, también en Comisión de Servicios, a otro puesto o ha pasado a otra condición administrativa (servicios especiales, excepciones, etc.). A partir de dicha solicitud y tras recabar los representantes del órgano convocante, la Dirección de Función Pública forma una Comisión de Valoración compuesta por los siguientes miembros:

(1) Presidencia: El titular del órgano al que está adscrito el puesto convocado. Normalmente el director o directora del órgano al que está adscrito el puesto convocado.

(2) Vocalía: Un representante designado por la Dirección de Función Pública. Este suele ser el entrevistador homologado encargado de diseñar y realizar las ECE.

(3) Vocalía: Un representante de la Dirección a la que esté adscrito el puesto convocado. Cuando hay ECE este representante debe ser un puesto de jefatura del mismo órgano y si no fuera posible, que ejerza funciones similares. Este realiza las funciones de secretario.

Las funciones de la Comisión son las siguientes:

(a) Elaborar la propuesta de convocatoria. El anuncio de la Comisión de Servicios.

(b) Declarar los admitidos y excluidos del proceso por cumplir o no los requisitos (pertenencia al Cuerpo o Grupo correspondiente, la titulación u otros requisitos específicos)

(c) Diseñar y corregir la prueba en caso de que la hubiere.

(d) Valorar los méritos alegados por los candidatos

(e) Valorar la entrevista. Es decir, la comisión es el panel de evaluadores formado normalmente por tres miembros, aunque pueden llegar a cuatro en caso de que la comisión se haya dotado, como asesor, de algún experto en el puesto convocado.

(f) Proponer el candidato o candidata seleccionados.

Como ya se ha dicho, el entrevistador es un vocal de la Comisión, pero un vocal que tiene la condición de entrevistador homologado, tal condición es reconocida por la Dirección de Función Pública tras haber acreditado suficiencia en el proceso formativo específico que

dicha Dirección convoca. La Dirección de Función Pública escoge al entrevistador en función de dos criterios: (a) La cercanía en los ámbitos laborales entre la plaza convocada y el puesto de trabajo del entrevistador. No hay que olvidar que los entrevistadores también son funcionarios de la AGPV y con funciones referidas a todos los ámbitos de los que se responsabiliza la AGPV, y (b) La disponibilidad de los entrevistadores. Desde el 22 de marzo de 2002, fecha en la que entró en vigor la utilización de la ECE, la demanda de Comisiones de Servicio con entrevista conductual estructurada ha sido tal que en ocasiones el primer criterio para seleccionar a los entrevistadores no se ha podido cumplir, ya que no estaban disponibles entrevistadores homologados. Teniendo en cuenta los criterios mencionados, la Dirección de Función Pública contacta con el entrevistador informándole de quiénes son los otros dos miembros de la Comisión. Asimismo le remite una información básica para que pueda comenzar a analizar el puesto objeto de la convocatoria. Esta información básica consiste en: (1) Monografía del puesto de trabajo, (2) Las Funciones Tipo, (3) Persona de la que depende estructuralmente el puesto convocado, (4) Personal técnico subordinado, (5) Antiguos ocupantes, y (6) Ocupantes de puestos similares o cercanos.

El proceso de elaboración de la ECE en la AGPV, siguiendo el modelo desarrollado por Salgado y Moscoso (2001), consta de tres fases: (1) análisis del puesto e identificación de las dimensiones relevantes del mismo, (2) redacción de las preguntas conductuales y (3) creación de las Escalas de Valoración con Anclajes Conductuales (EVAS).

El Análisis del Puesto de Identificación de las Dimensiones Relevantes

La técnica más utilizada para analizar los puestos de trabajo con el objetivo de elaborar una ECE es la de los Incidentes Críticos (IC) desarrollada por Flanagan (1954; véase Salgado y Moscoso, 2001 para un descripción amplia). En tanto que el proceso de diseño y elaboración de ECE tiene como uno de sus puntos esenciales la elaboración de las escalas con anclajes conductuales, la técnica de los incidentes críticos es casi insustituible porque aporta ejemplos conductuales desde los que elaborar dichas escalas. Además, los entrevistadores de la AGPV disponen en la actualidad de otros materiales referidos al puesto como son los Análisis de Tareas Críticas, resultantes de la implantación de la racionalización de estructuras llevada a cabo en esta Administración (Gorriti, 2002). Dichos análisis complementarán los incidentes críticos porque identifican para cada tarea representativa: estándares de ejecución, destrezas técnicas y de relación, factores de personalidad y condiciones del puesto.

La técnica de los IC es parte del curso para la homologación de los entrevistadores y no es nuestra intención detallarla aquí, valga nombrar cuáles son los criterios que la Dirección de Función Pública exige a sus entrevistadores para la identificación de estos incidentes: (a) ser específico de las tareas o cometidos del puesto convocado, (b) ser claramente distintivo de lo que es un rendimiento muy bueno o muy malo del puesto convocado, (c) hacer referencia a conductas observables por alguien no entrenado en dicho puesto de trabajo, (d) que describa el contexto del puesto del trabajo en el que se da o suele

dar dicho incidente, (e) que identifique las consecuencias que para el trabajador y para la AGPV tiene dicho incidente. La Dirección de Función Pública ha utilizado una ficha diseñada por Salgado (2002), en la que se especifican los contenidos ineludibles en la recogida de cualquier incidente crítico: (1) ¿Cómo fue la situación?, (2) ¿En qué consistió el incidente?, (3) ¿Qué hizo la persona?, (4) ¿Cuáles fueron las consecuencias o resultados de estas actuaciones?, (5) ¿Cuál es la importancia del incidente para el trabajo?, (6) ¿Con qué frecuencia se produce este incidente?

El procedimiento general para recoger incidentes críticos es acudir a otros funcionarios o a otras personas, incluso de fuera de la AGPV en caso de puestos de nueva creación o en aquellos casos para los que no se identifican expertos, y solicitarles que mediante una entrevista relaten anécdotas o sucesos que cumplan con los criterios arriba expuestos. Con objeto de normalizar lo más posible la recogida de datos, la Dirección de Función Pública señala algunos puntos a considerar: (a) intentar redactar unos 10 incidentes por experto, tantos positivos como negativos al objeto de conseguir, por lo menos, 30 incidentes, (b) limitar la extensión de los textos entre 10 y 30 líneas para que se puedan recoger todas las preguntas referidas al incidente, y (c) redactar incidentes que sean relevantes y discriminantes de ejecuciones muy buenas o muy malas del puesto convocado.

Una vez recogidos los incidentes y, puesto que el diseño de las ECE exige como elemento fundamental la previa identificación de las dimensiones básicas a medir, para lo cual los incidentes críticos recogidos son esenciales, el entrevistador y

los colaboradores identifican en cada incidente crítico una posible dimensión a la que se da un nombre. A continuación califican el incidente con un "1" si es un buen ejemplo de conducta negativa de dicha dimensión, como un "3" si es un ejemplo de conducta regular de dicha dimensión o con un "5" si es un buen ejemplo de conducta muy buena. Los números dos y cuatro también se usan aunque con menor frecuencia. Una vez que se ha hecho esto con todas las dimensiones y todas las conductas, se elabora una tabla con dicha información. De los datos aportados por esta tabla se pueden eliminar aquellas dimensiones para las que no existen suficientes anclajes conductuales y por tanto no van a poder ser medidas con objetividad. Si bien esta tabla aporta datos, la decisión sobre qué dimensiones son las que se van a medir en la entrevista se hace a partir de las valoraciones de todos los miembros de la comisión, a los que se pueden añadir los expertos que a juicio del entrevistador pueden aportar información relevante. Se les administra un cuestionario en el que se define cada dimensión y se les pide que valoren en una escala de 1 (Innecesaria) a 7 (Imprescindible) la importancia de cada dimensión para el puesto convocado. De estos resultados se determina primero la fiabilidad promedio interjueces y, si ésta es suficientemente alta (>.65), se desechan aquellas dimensiones con valores medios inferiores a 5. A las dimensiones restantes (normalmente entre 4 y 7) se les calcula el peso (importancia) de cada una con relación a las seleccionadas).

Redacción de las preguntas

Las preguntas deben ser capaces de provocar respuestas referentes a las conductas

relativas a las dimensiones conductuales básicas. En el caso de la ECE de la AGPV, los entrevistadores homologados elaboran las preguntas con los siguientes criterios: (a) Cada dimensión debe estar representada por al menos una pregunta, con sus correspondientes repreguntas. Algunas dimensiones pueden ser evaluadas a partir de más de una pregunta siendo factible hacerlo de forma bipolar en algunas dimensiones (una pregunta para un ejemplo bueno y otra para un ejemplo de malo), (b) Todos los candidatos deben ser evaluados en las mismas preguntas y en el mismo orden; y (c) Todas las preguntas deben ser conductuales, referidas a conductas laborales previas y a descripciones de dichas conductas.

Escalas de Valoración con Anclajes Conductuales (EVAS)

La utilización de estas escalas es una de las características distintivas de la ECE en el formato seleccionado por la AGPV. Los anclajes conductuales son ejemplos específicos de conductas observables en un puesto, que nos indican diferentes niveles de desempeño en una dimensión o competencia determinada. Estos anclajes ayudan al entrevistador y, en el presente caso, a la Comisión de Valoración (panel) a efectuar la evaluación de las respuestas de los candidatos, haciendo más precisas sus valoraciones. Para elaborar los anclajes es preciso redactar ejemplos de conducta para cada valor determinado según la escala decidida para la dimensión (por ejemplo de 1 a 5 ó de 1 a 7). Para ello se analizan los ejemplos identificados en los incidentes críticos y, partiendo de cada uno de ellos y teniendo en cuenta la dimensión a medir y las condiciones y especificidades

del puesto convocado, se redacta un anclaje para cada opción del gradiente en cada dimensión. La ventaja de los anclajes reside en la fidelidad que hay entre su enunciado y las conductas solicitadas por las preguntas de la ECE (véase Quijano, 1992 para una descripción general de estas escalas y Salgado y Moscoso, 2001, para su aplicación en el caso de la entrevista conductual estructurada).

Formación de los miembros del panel

La comisión de valoración es el panel que evalúa la entrevista y, excepto el vocal-entrevistador que es un entrevistador homologado, ninguno ha sido formado en la técnica de la ECE, por ello es imprescindible que a los miembros de la Comisión se les dé formación básica sobre cómo evaluar y en el caso de que quieran preguntar, cómo hacerlo. Nuestra experiencia es que la mayoría aprende rápido y son respetuosos con la mecánica de valoración aunque cuando se aventuran a preguntar, les resulta difícil respetar la naturaleza conductual de las preguntas, yéndose fácilmente a preguntas situacionales o de opinión. Los contenidos principales de dicha formación son los siguientes: (a) La importancia de la entrevista en el proceso selectivo de la Comisión de Servicios, su naturaleza no compensatoria y el peso de su puntuación en el total; (b) El principio de Consistencia Conductual que guía la teoría de la entrevista conductual estructurada: “El mejor predictor de la conducta futura es la conducta pasada en similares circunstancias”; y sus dos corolarios: “cuanto más reciente es la conducta pasada, mayor es su poder predictivo” y “cuanto mayor es el tiempo de existencia de la conducta, mayor es su poder predic-

tivo”, (c) Breve descripción de la mecánica de construcción de la ECE y justificación de las dimensiones, los anclajes y las preguntas, (d) Descripción sobre cómo se han obtenido los puntajes de importancia de cada dimensión y la mecánica de la valoración. Cómo obtener el puntaje de 0 a 10, (e) La naturaleza de la pregunta conductual, la importancia de diferenciarla de otros tipos de pregunta como son las de experiencia, situacionales, de credenciales, especulativas y las de opinión. No sólo para que comprendan su naturaleza sino también porque en la ECE que se hace en la AGPV, se permite que los miembros del panel formulen preguntas, aunque el peso de la entrevista lo lleva el entrevistador, y (f) Errores típicos de la medida: Tendencia central, halo, valores extremos y fuentes de contaminación, haciendo especial hincapié en la importancia del anclaje y en calificar la puntuación de la dimensión en el gradiente de la EVA y sólo según lo expuesto en el anclaje, no en función de un juicio global o sintético, más allá de lo dicho por el entrevistado.

Además, a todos los miembros del panel se les entregan dos documentos: (1) El diseño completo de la entrevista que incluye tanto los incidentes críticos como la definición de las dimensiones, los anclajes conductuales (EVAS) y las preguntas y repreguntas, y (2) un cuadernillo que lleva a la entrevista y en el que se incluyen las preguntas y repreguntas en una separata para consulta durante la entrevista, los anclajes por pregunta y espacio para la toma de notas, referencias y para la valoración del candidato en cada dimensión y una hoja final de resumen en las que vuelcan la valoración de cada dimensión y que se multiplica por el peso de la misma, y donde se obtiene la puntuación final.

Feedback recibido

A pesar de que como se ha dicho era conciencia general la necesidad de un instrumento objetivo que adecuara las personas a los puestos, la implantación de la ECE como paso excluyente para el acceso a puestos de alto nivel, ha sido visto como algo necesario pero duro. Dureza que, salvo en el caso de los entrevistados, se refiere más al contraste con la laxitud anterior, o por lo menos con la pasividad que del candidato exigía el concurso de méritos. A pesar de esto, no puede decirse que su implantación haya costado mucho, si bien esto puede afirmarse una vez que los detractores “a priori” la conocieron. Queremos decir que el hecho de conocer la rigurosidad en su construcción, su relación con el trabajo para el que se selecciona, su nula invasión de intimidad, y la libertad y objetividad de sus evaluaciones, además de los controles para sus posibles desviaciones, han hecho cambiar de opinión a los que de entrada se oponían a ella. Y a esto han ayudado los propios entrevistados.

El escepticismo inicial de los primeros candidatos se ha tornado en respeto. Tanto es así que se ha producido un fenómeno interesante: la autoexclusión de algunos candidatos “potencialmente inadecuados” antes de la entrevista. Creemos que hay varios elementos para esta autoexclusión. El primero es que la ECE es un instrumento muy potente que deja en evidencia a quien es bueno y a quien no lo es, y además que esta evidencia es demasiado manifiesta tanto para el panel como para el propio entrevistado. Segundo, que el mero hecho de que durante la entrevista se solicite al entrevistado que indique una segunda fuente de información sobre el suceso que está narrando y sobre su conducta en

el mismo (pedir referencias), reduce en gran medida o impide por completo la mentira y, en caso de producirse, comprometen a posteriori al propio candidato. Tercero, que las propias preguntas, reducen mucho la posible "auto-venta" del candidato porque le ubican en realidades desde las que le es muy difícil inventarse hechos, lo cual no ocurre con las entrevistas convencionales donde las habilidades sociales, la inteligencia y la personalidad del candidato influyen de forma notable (Salgado y Moscoso, 2002)

Los entrevistados con los que hemos mantenido conversaciones después de haber pasado por este proceso han dicho frases como las siguientes: "*Ha sido duro e intenso, en realidad me he quedado vacía*", "*Nunca pensé que se pudiera preguntar tanto, ni tan exhaustivamente sobre algo*", "*La verdad es que [la entrevista] estaba bien preparada. Es la primera vez que me hacen una entrevista en la que entiendo la relación que tiene lo que me preguntan con lo que pide el puesto*", "*Ha sido duro, no me podía escapar; siempre me volvían a traer a donde les interesaba, no era capaz de expresarme del todo. Creo que me ha fallado la memoria*", "*Ahora que he superado la entrevista, me siento como si estuviera en esta plaza con todas las bendiciones*", "*¿Y qué quieres que diga si nunca lo he hecho? ¡Pues me he callado! Aunque ya sé que eso no es bueno*", "*Tiene que ser duro preparar una entrevista así. Me he sentido bien medida*".

Por su parte la mayoría de los miembros de los paneles a los que hemos consultado coinciden en señalar varios valores diferenciados, muchos de ellos bajo el paraguas del término "muy interesante".

Sus calificativos fueron: Riguroso, duro, sistemático, objetivo, fácil de evaluar, muy relacionado con el trabajo. También hemos observado respeto por el trabajo del entrevistador, como si supieran que era un trabajo técnicamente muy bien elaborado y que el entrevistador/a era el auténtico/a experto. Algunos han llegado a decir que "*nunca había sospechado que en esto de la selección de personal había tanta técnica, ni tanta rigurosidad*". Este hecho también se manifestaba en los reparos a la hora de preguntar por parte de los miembros del panel. No es que no quisieran hacerlo, simplemente eran conscientes de la cantidad de variables que intervenían en cada pregunta, de la rigurosidad de su formato, de su naturaleza conductual, de la dificultad de convocar las respuestas de ese tipo y tenían reparos a hacerlas. Y cuando las hacían, previamente se declaraban no expertos en la materia. Esto no ocurría en el caso de que las preguntas se refirieran a conocimientos.

Saber que después se calcularía la fiabilidad interjueces y que los estudios meta-analíticos previos indicaban que la fiabilidad oscilaba entre .70 (Conway et al., 1995) y .75 (Salgado y Moscoso, 1995) despertaba cierta inquietud en los miembros del panel a la hora de ser "calificados" como buenos o malos evaluadores. Lo mismo pero de forma más intensa ocurre con el entrevistador si pensamos que, en realidad, este coeficiente tiene mucho que ver con cómo ha conducido la entrevista ya que puntuaciones altas en este coeficiente significan que todos han visto lo mismo y que ese "ver" ha sido provocado por las preguntas. Otro aspecto interesante de los miembros del panel era la seguridad que expresaban a la hora de eva-

luar al candidato. La frase arriba citada: "fácil de evaluar" era una constante.

Lo dicho para los miembros no expertos del panel es válido también para los Directores y Jefes de Servicio. Sin embargo, centrándonos en lo que es la percepción de los cargos y jefes de servicio no evaluadores, éstos se posicionan de modo ambivalente. Para la mayoría, la ECE ha venido a cubrir un hueco técnico que asegurase la adecuación de las personas a los puestos y desde éste punto de vista la aplauden. Para otros la ECE es un obstáculo a la hora de colocar en sus puestos a gente de su interés.

La Homologación de los Entrevistadores

La Dirección de Función Pública, a través del Instituto Vasco de Administración Pública (IVAP), gestor de la formación de la AGPV, ha convocado y finalizado dos cursos de formación en ECE y aquellos participantes que superaron dicho curso (70%) han sido homologados por la Dirección de la Función Pública para llevar a cabo entrevistas conductuales estructuradas de selección de personal en el ámbito de sus competencias. Los cursos fueron diseñados e impartidos por Jesús Salgado, quien también evaluó el grado de competencia técnica de los participantes. El primer curso tuvo lugar entre Diciembre de 2001 y Enero de 2002 y el segundo entre Febrero y Marzo de 2003. La primera parte del curso es de carácter teórico-práctico y la segunda parte es de aplicación y valoración de la entrevistas conductuales "in vivo". La primera parte del curso consiste en la adquisición de conocimientos y destrezas sobre las entrevistas conductuales estructuradas, sus propiedades psicométricas y como estimarlas, la técnica de inci-

dentos críticos, la identificación de dimensiones conductuales, el diseño de anclajes conductuales (EVAS), la identificación de los diversos tipos de preguntas evaluativas y el diseño de preguntas y repreguntas conductuales, el sistema de evaluación y la fiabilidad interjueces. Esta fase finaliza la elaboración por parte de cada alumno de un diseño completo de ECE de un puesto escogido por dicho alumno. Dicho diseño, que debe incluir tanto las dimensiones, como las preguntas y los anclajes conductuales, además de los incidentes críticos, es objeto de valoración por parte del profesor, no siendo ello excluyente para participar en la segunda fase, aunque es determinante para la valoración final. La segunda parte del curso es de carácter aplicado y en ella cada alumno, a partir del diseño que ha realizado en la primera parte, se entrena la ECE tomando como entrevistado a otro alumno y como panel al resto de alumnos junto con el profesor. La entrevista se graba en audio y el alumno se queda con la copia correspondiente. Finalizada cada entrevista el profesor y todos y cada uno los alumnos dan feedback tanto al entrevistador como al entrevistado y todos valoran al entrevistado. De este modo se puede observar el grado de acuerdo entre las valoraciones del profesor y del alumno y entre las valoraciones de este con las del resto de compañeros. Este momento final permite examinar y corregir los sesgos habituales que se producen durante la entrevista y durante la valoración. Por norma general las entrevistas realizadas "en vivo" suelen durar alrededor de 50 minutos, aunque hay variabilidad dependiendo tanto del entrevistado como del entrevistador.

La superación de las dos fases del curso es condición para que la Dirección de Fun-

ción Pública reconozca la suficiencia en la teoría y práctica de la ECE. Ello implica, administrativa y legalmente, un reconocimiento oficial del órgano competente en materia de Función Pública para el ejercicio de esta práctica selectiva; una homologación como entrevistador. Dicho reconocimiento, dicha homologación, es condición imprescindible para la práctica de la ECE como instrumento que permita la toma de decisiones selectivas respecto del personal de la AGPV y en ella. Es una forma de asegurar un método y unos resultados. En definitiva, una certificación de calidad en esta práctica selectiva.

INVESTIGACIÓN EMPÍRICA DE LA FIABILIDAD DE LOS ENTREVISTADORES Y DE LA DECISIÓN EN LA ADMINISTRACIÓN GENERAL DEL PAIS VASCO

La tercera parte de este artículo tiene como objetivo presentar los resultados de la investigación empírica llevada a cabo para examinar la fiabilidad de la entrevista conductual estructurada incluida en el proceso de selección de personal para los puestos de jefatura de la AGPV. En este sentido, se examinó tanto la fiabilidad de cada entrevistador individualmente como la fiabilidad del proceso de entrevista (fiabilidad del panel de entrevistadores) y se llevaron a cabo tres estudios relacionados que se describen seguidamente. No obstante, antes de presentar los resultados obtenidos es preciso señalar que la investigación ha estado fundamentada en la utilización del meta-análisis psicométrico desarrollado por Hunter y Schmidt (1990) y ampliamente empleado en las investigaciones sobre las propiedades psicométricas de los

instrumentos de selección de personal en Psicología del Trabajo y las Organizaciones. Antes de proceder a la realización de los análisis correspondientes, indicaremos brevemente en qué consiste el meta-análisis psicométrico.

Meta-análisis como Método de Investigación

El meta-análisis es un método de combinar cuantitativamente los resultados de múltiples investigaciones realizados sobre el mismo tema. Su desarrollo actual en la Psicología del Trabajo y las Organizaciones comenzó hace más de veinticinco años (Schmidt y Hunter, 1977), aunque ya existían ejemplos de combinaciones de resultados desde hace 80 años. A mediados de los años setenta, en diversas áreas de la Psicología se hicieron aportaciones de métodos de análisis de datos que colectivamente han sido denominados meta-análisis (p.e. Glass, MacGraw y Smith, 1981; Hedges y Olkin, 1985; Hunter, Schmidt y Jackson, 1982; Rosenthal, 1984). En la Psicología del Trabajo y las Organizaciones y en Selección de Personal en particular, el método de meta-análisis más utilizado ha sido el conocido como meta-análisis psicométrico, desarrollado por Hunter y Schmidt (1990). El meta-análisis psicométrico tiene dos objetivos principales. Por un lado calcular el mejor estimador del tamaño del efecto (p.e. coeficiente de correlación, d) en una relación, sea causal o no causal, a partir de un elevado número de estudios individuales y, en segundo lugar, tratar de explicar si la variabilidad que se observa en los tamaños de los efectos calculados en los estudios individuales es una variabilidad real o es el producto de una serie de errores artefactual que están presentes en

los estudios individuales. El mejor estimador del tamaño del efecto es aquel del que se han eliminado las fuentes de error, habitualmente el error de medida y la restricción en el rango. En el segundo caso, entre los errores artifactuals, que introducen varianza de error cuando no existe en la relación entre dos variables X e Y, se encuentran el error de muestreo, la dicotomización de la variable X, la dicotomización de la variable Y, la fiabilidad de X, la fiabilidad de Y, la restricción en el rango de X, la restricción en el rango de Y, la medida imperfecta del constructo X, la medida imperfecta del constructo Y, y los errores de transcripción y codificación. Para algunos de estos errores existen fórmulas que permiten estimar la cantidad de error debida a tales variables artificiales (p.e. el error de muestreo, el error de medida, la restricción en el rango). Sin embargo, la varianza debida a otros artefactos es imposible de estimar. Por ejemplo este es el caso de la varianza debida a errores tipográficos y de codificación de los datos. Por definición, es imposible saber cuantos errores de este tipo hemos cometido, aunque en la mayoría de las bases de datos se encuentran errores de este tipo.

Al realizar un meta-análisis no siempre es posible corregir el tamaño del efecto por los efectos del error de medida y la restricción en el rango, ni siempre es posible eliminar de la varianza observada la varianza debida a todas las fuentes de error. La varianza que queda después de eliminar la varianza de error es la varianza verdadera. Por tanto, la variabilidad observada (estimada como varianza o como desviación típica) se compone de estas dos partes, varianza verdadera y varianza de error. En selección de personal, el meta-análisis psicométrico ha demostrado en un número

muy elevado de casos que la mayoría de la varianza observada es varianza de error y que la varianza verdadera, en caso de existir, es muy pequeña. Cuando se realiza un meta-análisis en el que no se corrige el tamaño del efecto y sólo se elimina la varianza debida al error de muestreo, a dicho meta-análisis en la literatura científica se le denomina 'bare-bone'. Este tipo de meta-análisis es el que llevaremos a cabo en la presente investigación.

Estudio 1. Estimación de los Coeficientes de Fiabilidad por Acuerdo entre Entrevistadores (Valoradores) para cada Diseño (Formato) de Entrevista Realizado

Hasta el momento presente se han elaborado 20 diseños de entrevistas conductuales estructuradas para otros tantos puestos específicos del personal funcionario del Gobierno Vasco. Cada uno de los diseños ha sido elaborado por uno de los entrevistadores homologados por el Gobierno Vasco y constan de tres elementos, las dimensiones a evaluar, las preguntas conductuales que se le formularán al candidato y las escalas de valoración con anclajes conductuales que se utilizarán para valorar la información aportada por cada candidato durante el transcurso de la entrevista y como respuesta a las preguntas que se le formularon. En la Tabla 1 figuran los 20 puestos para los cuales se elaboró una entrevista conductual estructurada específica, así como el número de dimensiones que se valoraron en cada entrevista.

Cada uno de los diseños se elaboró siguiendo el procedimiento estándar de desarrollo de entrevistas conductuales estructuradas que había sido entrenado

Tabla 1 Relación de puestos para los que se diseñó un formato de Entrevista Conductual Estructurada	
Puesto	D
Responsable. Jurídico Administrativo	5
Responsable de Calidad y Procedimientos de OSALAN	7
Responsable de Gestión del Personal	5
Responsable de Gestión Económica Precios Públicos	6
Responsable de Control Económico Normativa	6
Responsable de Planificación y Obras	5
Selección IVAP	6
Gestor deportivo	6
Jefe de Proyectos de la Oficina de Modernización Administrativa	7
Letrado Régimen Jurídico - Dotación 23	5
Letrado Servicios Centrales	5
Coordinador Consejo Vasco	5
Responsable Territorial de Salud	7
Jefe Riesgos y Seguros	3
Jefe Prestaciones Farmacéuticas	4
Jefe Administración	5
Jefe Proyectos de la Oficina de Modernización Administrativa	7
Letrado Régimen Jurídico - Dotación 7	5
Responsable Sectorial de Industria	5
Letrado Régimen Jurídico - Dotación 10	5

Nota. D: número de dimensiones evaluadas en cada entrevista.

durante el curso de formación de los entrevistadores homologados. Por tanto se dieron los siguientes pasos:

1. Elección del Entrevistador Homologado responsable del proceso.
2. El Entrevistador realiza un análisis del puesto y recoge tantos incidentes críticos como le sea posible, acudiendo a todas las fuentes disponibles.
3. Una vez obtenidos los incidentes se agrupan en función de su similitud tanto semántica como conceptual.
4. Se proporciona un nombre a cada una de las dimensiones determinadas por los incidentes críticos.
5. Se escriben las preguntas conductuales que servirán para explorar cada una de las dimensiones del puesto de trabajo identificadas.
6. Se vuelven a utilizar los incidentes críticos para identificar los ejemplos conductuales que se incluirán en las escalas de valoración con anclajes conductuales y que servirán para que los entrevistadores valoren a cada entrevistado en cada dimensión.

Una vez que el entrevistador homologado responsable del proceso de entrevista ha confeccionado el diseño correspondiente, se pone en contacto con los restantes miembros del panel que han sido elegidos para formar parte del mismo y acuerda el momento en que se proporcionará un breve curso de formación sobre la entrevista, las escalas de valoración y el proceso general de entrevista. Una vez impartido el curso de formación, ya puede comenzar el proceso de entrevista, previa convocatoria pública de los concursantes.

La parte de la entrevista correspondiente a la formulación de preguntas es responsabilidad del entrevistador homologado, aunque los restantes miembros del panel tienen la posibilidad de preguntar si algo no ha quedado lo suficientemente claro en las contestaciones de los candidatos. Las repuestas se graban en un cassette, siempre y cuando el entrevistado no tenga inconveniente y dicho cassette se incorpora al expediente administrativo.

Una vez finalizada cada entrevista, de acuerdo con el procedimiento estándar, cada uno de los miembros del panel efectúa la valoración correspondiente a cada una de las dimensiones del puesto de trabajo para cada uno de los entrevistados, de tal modo que al acabar la última entrevista ya se cuenta con la valoración de todos los anteriores candidatos.

A partir de las puntuaciones otorgadas por los evaluadores procedimos a calcular la fiabilidad en cada uno de los diseños. La fiabilidad que interesa en el presente caso es la estimada mediante un coeficiente de acuerdo entre entrevistadores. Hay otras posibilidades como el coeficiente alfa de consistencia interna, pero este coe-

ficiente solo controla la varianza intravalorador, mientras que el coeficiente de acuerdo entre entrevistadores controla además el error entre valoradores. Este coeficiente de fiabilidad se estima como el promedio de los coeficientes de correlación entre las puntuaciones otorgadas a cada candidato por cada uno de los valoradores de la entrevista en cada una de las dimensiones. En el presente caso hay dos coeficientes de fiabilidad interjueces que nos interesan. Por un lado nos interesa conocer la fiabilidad promedio individual de cada uno de los entrevistadores en cada una de las entrevistas. Por otro lado, puesto que la decisión que afecta a los candidatos en la tomada por el panel de entrevistadores en su conjunto, nos interesa también la decisión correspondiente a cada uno de los paneles utilizado en cada uno de los diseños de entrevista aplicados hasta el momento presente. Para calcular la fiabilidad de los paneles hemos utilizado la fórmula profecía de Spearman-Brown y utilizando como 'n' el número de miembros del panel. Puesto que había paneles de 3 y 4 miembros, la fórmula podía adoptar una de estas dos posibilidades:

$$r_{xx \text{ panel}} = (3 * r_{xx \text{ individual}}) / [1 + (2 * r_{xx \text{ individual}})]$$

o bien

$$r_{xx \text{ panel}} = (4 * r_{xx \text{ individual}}) / [1 + (3 * r_{xx \text{ individual}})]$$

En la Tabla 2 figuran los resultados del análisis de la fiabilidad entrejueces para cada uno de los diseños de entrevista elaborados. En esta tabla, la primera columna muestra la fiabilidad promedio de cada juez en cada uno de los diseños elaborados. La segunda columna muestra la fiabilidad correspondiente a cada uno de los paneles

y, por tanto, la fiabilidad correspondiente a cada una de las decisiones tomadas por el correspondiente panel. Por último, las tres columnas finales representan el número de miembros del panel, el número de candidatos entrevistados y en número de dimensiones evaluadas en cada uno de los diseños de entrevista, respectivamente.

Por lo que se refiere a la fiabilidad de cada uno de los miembros de cada uno de los paneles, ésta es muy satisfactoria en la mayoría de los casos, teniendo un rango de valores que varía desde .41 hasta .96. De los valores obtenidos, el 20% son menores que .70; el 30% se sitúan entre .70 y .80; el 15% están entre .81 y .90; y el 35% son superiores a .90.

Por lo que respecta a la fiabilidad de los distintos paneles y, en consecuencia, a la fiabilidad de las decisiones tomadas por dichos paneles, los coeficientes varían desde .73 a .99, siendo la gran mayoría superiores a .90. De los valores calculados, el 10% tienen una fiabilidad inferior a .80; el 15% tienen un valor inferior a .90; y el 75% de los valores son superiores a .90.

Las conclusiones de este primer estudio son satisfactorias, tanto en lo que respecta a la fiabilidad de los valoradores individuales como en lo que respecta a las decisiones tomadas por los paneles. En el primer caso, las fiabilidades de los valoradores individuales se encuentran en

Tabla 2.
FIABILIDAD PROMEDIO DE CADA ENTREVISTADOR EN CADA PROCESO DE ENTREVISTA Y FIABILIDAD DE LA DECISIÓN EN CADA PROCESO

Puesto	r_{xx} individual	r_{xx} decisión/panell	K	N	D
Resp. Jurídico	.86	.95	3	4	5
Resp. Calidad	.46	.77	4	2	7
Resp. Gestión Personal	.96	.99	3	3	5
Gestión Económica	.95	.98	3	3	6
Resp. Control Económico	.75	.90	3	2	6
Resp. Planificación y obras	.59	.85	4	2	5
Selección IVAP	.41	.73	4	1	6
Gestor deportivo	.70	.90	4	4	6
Oficina Modernización Admin.	.54	.83	4	1	7
Letrado Régimen Jurídico -23	.79	.92	3	2	5
Letrado Servicios Centrales	.98	.99	3	5	5
Coordinador Consejo Vasco	.86	.95	3	4	5
Resp. Territorial Salud	.71	.88	3	4	7
Jefe Riesgos y Seguros	.96	.99	3	1	3
Jefe Prestaciones Farmacéuticas	.79	.95	3	1	4
Jefe Administración	.95	.99	4	5	5
Jefe Proyectos OMA	.89	.97	4	1	7
Letrado Régimen Jurídico - 7	.93	.98	3	3	5
Resp. Sectorial Industria	.79	.92	3	1	5
Letrado Régimen Jurídico- 10	.91	.97	3	2	5

Nota. r_{xx} media = Fiabilidad promedio de los entrevistadores en cada uno de los formatos de entrevista.

r_{xx} decisión = Fiabilidad promedio de la decisión del panel de entrevistadores en cada formato de entrevista.

K= Número de miembros del panel.

N= Número de entrevistados.

D= Número de dimensiones evaluadas en cada entrevista.

el rango de valores obtenidos en los meta-análisis ya comentados sobre la fiabilidad de la entrevista conductual estructurada, pero además estos valores han de juzgarse teniendo en cuenta que tan solo uno de cada tres o de cada cuatro de los miembros de cada panel era un entrevistador entrenado y homologado, siendo los restantes miembros jueces expertos pero no entrevistadores entrenados. Por lo que respecta a la fiabilidad de las decisiones de los paneles analizados ha sido en todos los casos satisfactoria ya que la fiabilidad alcanzada en el peor de los casos ha sido semejante a la fiabilidad obtenida en los meta-análisis de Conway et al. (1995) y de Salgado y Moscoso (1995), pero la gran mayoría de las fiabilidades ha sido extraordinariamente buena, ya que se sitúa por encima de las fiabilidades obtenidas por los instrumentos que mejores características psicométricas han mostrado a lo largo de la historia de la selección de personal (ver Salgado, 1999; Salgado, Viswesvaran y Ones, 2001, para una revisión). En resumen, pues, los resultados mostrados en este primer estudio indican que el proceso de entrevista conductual estructurada y la decisión a que llega el panel en los procesos de selección del Gobierno Vasco que utilizan este procedimiento es plenamente satisfactorio.

Estudio 2. Estimación Meta-analítica de la Fiabilidad Individual de los Entrevistadores de los Procesos de Selección del Gobierno Vasco

Hemos indicado ya que el meta-análisis es la técnica cuantitativa más poderosa para establecer los parámetros del tamaño del efecto de las intervenciones

organizacionales así como de su correspondiente variabilidad. Ya señalamos que el meta-análisis no sólo nos permite establecer cuál es la fiabilidad del proceso de entrevista conductual estructurada sino también cuánto de la variabilidad observada es real y cuánto es de carácter artificial.

En el estudio 1 hemos podido constatar que la fiabilidad de los entrevistadores presentaba un rango de valores que oscilaba entre .41 y .96, lo que significa que existe bastante variabilidad entre las dos puntuaciones. Además, el estudio 1 nos indicaba que las fiabilidades individuales eran bastante satisfactorias, pero no nos indicaba cuál era la fiabilidad más probable (fiabilidad promedio) en el caso de las entrevistas conductuales estructuradas utilizadas en la AGPV. Así pues, teniendo en cuenta los datos precedentes y que el meta-análisis nos permite responder a las dos cuestiones anteriores, a saber, cuál es la fiabilidad promedio de los entrevistadores individuales y cuánto de la variabilidad observada es real, hemos llevado a cabo un estudio meta-analítico para responder a dichas cuestiones.

En la Tabla 3 figuran los resultados del meta-análisis llevado a cabo. Como es habitual, hemos realizado los análisis usando dos modalidades. La primera pondera cada uno de los coeficientes de fiabilidad por la correspondiente frecuencia. La segunda modalidad, más apropiada, pondera cada uno de los coeficientes por el tamaño de la muestra utilizada, es decir, por el número de entrevistados. De este modo, se dan pesos diferentes a coeficientes obtenidos con muestras diferentes. En el primer caso, todos los coeficientes se ponderan por igual.

Tabla 3
META-ANÁLISIS DE LOS COEFICIENTES DE FIABILIDAD DE LOS ENTREVISTADORES

$r_{xx \text{ frec}}$	SDr _{frec}	$r_{xx \text{ muestra}}$	SDr _{muestra}	SDr _{verdadera}	K	N	E	% VE
.79	.17	.83	.15	.00	20	51	67	100

Nota. $r_{xx \text{ frec}}$ = Fiabilidad promedio de los entrevistadores ponderada por la frecuencia.
SDr_{frec} = Desviación típica observada de la fiabilidad promedio de los entrevistadores ponderada por la frecuencia.
 $r_{xx \text{ muestra}}$ = Fiabilidad promedio de los entrevistadores ponderada por el tamaño de la muestra.
SDr_{muestra} = Desviación típica observada de la fiabilidad promedio de los entrevistadores ponderada por el tamaño de la muestra.
SDr_{verdadera} = Desviación típica verdadera de la fiabilidad promedio de los entrevistadores.
K = número de puestos para los que se diseñó un formato de entrevista.
N = número de entrevistados.
E = número total de entrevistadores que formaban los paneles de entrevista.
% VE = Porcentaje de variabilidad de los coeficientes de fiabilidad debida a error de muestreo.

Como puede apreciarse, la fiabilidad promedio ponderada por la frecuencia es .79 y la fiabilidad promedio ponderada por el tamaño de la muestra es .83, es decir son dos valores muy semejantes y, en cualquier caso, indicadores de que la fiabilidad promedio de los jueces (entrevistadores) en la AGPV es superior a la encontrada en los estudios meta-analíticos previos (p.e. Conway et al., 1995; Salgado y Moscoso, 1995) y un indicador de que los entrevistadores que usan los diseños elaborados por la AGPV son muy fiables. En nuestra opinión, este resultado se debe principalmente al hecho de que usamos un diseño de entrevista conductual completamente estructurada, es decir, que el mismo patrón de preguntas dimensión y de escalas de valoración se utiliza con todos los entrevistados. A este debe añadirse, naturalmente, que las entrevistas se han elaborado con unos notables principios de exigencia metodológica y que los miembros del panel han sido entrenados previamente. Es importante tener en cuenta que la fiabilidad test-retest promedio de los tests de habilidad mental general, que es el mejor predictor individual

del desempeño en el trabajo, es precisamente .83 (véase, Salgado y Anderson, 2002; 2003; Salgado et al., 2003a,b). Por tanto, el resultado obtenido en el caso de los entrevistadores de los paneles de la AGPV permite situar esta modalidad entrevista entre los mejores instrumentos para la selección de personal.

Por lo que se refiere a la variabilidad observada en los coeficientes de fiabilidad, estimada como una desviación típica, es .17 en el caso de que los coeficientes sean ponderados por la frecuencia y .15 si son ponderados por el tamaño de la muestra. Sin embargo, toda esta variabilidad observada es variabilidad artefactual ya que cuando se considera la varianza atribuible al error de muestreo, puede verse que el 100 de la variabilidad es varianza de error y en consecuencia no hay variabilidad real. En otras palabras, el meta-análisis muestra que la SD de la población es igual a cero y, en consecuencia, que el mejor estimador de la fiabilidad de los entrevistadores de AGPV que usan entrevistas conductuales estructuradas es .83.

Estudio 3. Estimación Meta-analítica de la Fiabilidad de las Decisiones del Panel de Entrevistadores de las Procesos de Selección de Gobierno Vasco

Los datos obtenidos en el estudio 1 pueden ser utilizados como datos de entrada para este tercer estudio que, desde el punto de vista organizacional, tiene todavía mayor interés que el anterior. Si en el caso del estudio 2 examinábamos la fiabilidad de los entrevistadores considerados individualmente, en este tercer estudio, examinamos de la fiabilidad de las decisiones tomadas por el panel de entrevistadores. Desde el punto de vista de la Administración General del País Vasco este es el valor más importante, puesto que determina la eficiencia de las propuestas de cobertura de puestos y, al propio tiempo, si algún candidato desease que se revisara su calificación se haría sobre la totalidad de la puntuación obtenida en la entrevista y no solo sobre la puntuación otorgada por uno de los miembros. Con objeto de garantizar una mayor justicia, equidad, imparcialidad y fiabilidad en las decisiones de la entrevista, la Administración General del País

Vasco, como ya hemos señalado en las secciones precedentes estableció como precepto que en la entrevista participasen entre 3 y 4 personas, representando estamentos diferentes y que la decisión sea tomada colegialmente, no sólo por el entrevistador homologado. Por tanto, para la AGPV lo crucial es la decisión tomada y, por tanto, es crucial conocer cual es la fiabilidad final del proceso de entrevista conductual estructurada.

Los datos del estudio 1 habían mostrado que la decisión tenía una fiabilidad considerable y que los valores oscilaban entre .73 y .99. En definitiva, se volvieron a constatar dos hechos. El primero es que la decisión del panel era cuando menos bastante fiable y el segundo que, a pesar de todo, existía variabilidad en la fiabilidad. Una vez más, el meta-análisis es la técnica necesaria para precisar cual es la fiabilidad real de la decisiones basadas en la entrevista conductual estructurada y, en segundo lugar, si hay variabilidad real en la fiabilidad o si más bien esta es un producto de los artefactos estadísticos con los que se efectúa el cálculo (p.e. error de muestreo).

Tabla 4
META-ANÁLISIS DE LOS COEFICIENTES DE FIABILIDAD DE LAS DECISIONES BASADAS EN EL PANEL DE ENTREVISTADORES

$r_{xx \text{ frec}}$	SDr_{frec}	$r_{xx \text{ muestra}}$	SDr_{muestra}	$SDr_{\text{verdadera}}$	K	N	E	%VE
.92	.08	.93	.06	.00	20	51	67	100

Nota. $r_{xx \text{ frec}}$ = Fiabilidad promedio de los entrevistadores ponderada por la frecuencia.
 SDr_{frec} = Desviación típica observada de la fiabilidad promedio de los entrevistadores ponderada por la frecuencia.
 $r_{xx \text{ muestra}}$ = Fiabilidad promedio de los entrevistadores ponderada por el tamaño de la muestra.
 SDr_{muestra} = Desviación típica observada de la fiabilidad promedio de los entrevistadores ponderada por el tamaño de la muestra.
 $SDr_{\text{verdadera}}$ = Desviación típica verdadera de la fiabilidad promedio de los entrevistadores.
 K= número de puestos para los que se diseño un formato de entrevista.
 N= número de entrevistados.
 E= número total de entrevistadores que formaban los paneles de entrevista.
 %VE= Porcentaje de variabilidad de los coeficientes de fiabilidad debida a error de muestreo.

En la Tabla 4 pueden verse los resultados del meta-análisis llevado a cabo. Como en el caso del estudio 2, hemos efectuado los cálculos ponderando los coeficientes por la frecuencia y también por el tamaño de la muestra. En el primer caso, la fiabilidad de la decisión ha sido de .92 y en el segundo caso la fiabilidad ha sido de .93, es decir que se ha obtenido un resultado muy semejante para ambos casos. Este dato es revelador de la gran potencia de la fiabilidad de las decisiones basadas en las entrevistas conductuales estructuradas usadas en la AGPV. Existen muy pocos casos en la literatura científica en los que se alcancen estas fiabilidades por acuerdo entre jueces para un procedimiento de entrevista estandarizado (véase la Tabla 2 de Salgado y Moscoso, 1995).

Por lo que respecta a la variabilidad observada de los coeficientes de fiabilidad estimada como la desviación típica de los mismos, ésta es .08 y .06 si se usan respectivamente la frecuencia o el tamaño de la muestra para ponderar los coeficientes. Sin embargo, al calcular la variabilidad que se debe a errores artefactuales (error de muestreo en el presente caso), el resultado indica que toda la variabilidad observada es variabilidad de error y que no existe variabilidad real en la fiabilidad. En otras palabras, el mejor estimador de la fiabilidad de las decisiones del panel es .93.

Una última constatación de la potencia del meta-análisis realizado puede comprobarse si aplicamos la fórmula de Spearman-Brown. Teniendo en cuenta los resultados de la fiabilidad obtenidos en el estudio 2, es decir .79 y .83, si usamos la mencionada fórmula y consideramos que el panel está formado por tres entrevistadores, los resultados predichos por la fórmula

serían .92 y .94, valores redondeados al entero más próximo. En otras palabras, hemos obtenido un resultado prácticamente idéntico al obtenido mediante el meta-análisis (los resultados serían .94 y .95, con paneles de cuatro jueces).

CONCLUSIONES GENERALES

Este artículo tenía tres objetivos. Por un lado, nuestro propósito era presentar una descripción de la entrevistas conductuales estructuradas y una revisión de la investigación reciente sobre su validez predictiva, su validez de constructo y su fiabilidad. En segundo lugar, deseamos presentar la experiencia llevada a cabo durante el último año en la Administración General del País Vasco, en la que se ha estado utilizando la entrevista conductual estructurada para la cobertura de vacantes en puestos de jefatura. Tratábamos con ello de dar cuenta de los trabajos realizados para que pudieran servir de estímulo y de guía para otras administraciones que pudieran estar interesadas en un proceso semejante. El tercer objetivo fue el examen empírico de la fiabilidad de las entrevistas realizadas en la AGPV, teniendo como elementos de examen a los propios entrevistadores y a la decisión tomada por el panel de entrevistadores. Desde un punto de vista formal de investigación, este tercer objetivo es la parte más importante de este artículo y la que permite llegar a algunas conclusiones sobre la adecuación, desde el punto de vista de su fiabilidad, de las entrevistas utilizadas en la AGPV con propósitos de selección.

Entre las aportaciones de la presente investigación cabe mencionar especialmente tres por su relevancia. En primer

lugar que se ha realizado un número importante de diseños de entrevista conductual estructurada y que para cada uno de ellos se ha calculado la fiabilidad promedio de los entrevistadores y la fiabilidad del panel de decisión. En segundo lugar, que se ha utilizado el método más desarrollado para la acumulación de datos de investigaciones, el meta-análisis psicométrico, y se ha aplicado en un contexto particular en el que no se había hecho antes, la administración pública, y para un instrumento para el cual no suelen hacerse este tipo de investigaciones, la entrevista de selección de personal. La tercera aportación relevante en nuestra opinión, es que la investigación llevada a cabo muestra sin ningún género de dudas que se puede investigar científicamente sobre la selección de personal en el ámbito de la administración pública y usando técnicas originalmente concebidas para las organizaciones privadas.

A partir de los resultados obtenidos, también se puede plantear una serie de conclusiones. La primera gran conclusión es que el proceso de selección mediante entrevistas conductuales estructuradas usadas por la Administración General del País Vasco ha sido un éxito total desde el punto de vista de la fiabilidad de los entrevistadores y de la fiabilidad de la decisión final. Entrevistadores homologados y entrevistadores ad hoc han alcanzado unas cotas de fiabilidad que se pueden considerar un hito en los procesos de selección tanto entre las organizaciones públicas como privadas. Un examen de la fiabilidad de la mayoría de los instrumentos de selección de personal muestra que casi ninguno llega a los niveles alcanzados en el presente caso. Por ejemplo, Salgado (1999; Salgado, Viswesvaran y Ones,

2001) en sus revisiones de la investigación efectuada en los últimos diez años, muestra que instrumentos elaborados concienzudamente desde el punto de vista psicométrico como los tests de habilidad mental general tienen una fiabilidad promedio de .83, y que los tests de personalidad basados en el modelo de los cinco grandes factores tienen una fiabilidad de .80. Por tanto, los resultados encontrados en la presente investigación indican que se han superado los estándares más rigurosos, ya que tenemos valores promedios de .83 y .93 para los entrevistadores y para la decisión, respectivamente.

Una segunda gran conclusión, es que el proceso sistemático y riguroso empleado por la AGPV para desarrollar y aplicar sus entrevistas ha mostrado que instrumentos de selección de personal, inicialmente concebidos en organizaciones privadas, pueden aplicarse en el ámbito de la Administración Pública sin merma de eficacia en cuanto a sus características de fiabilidad.

Una tercera conclusión es que la decisión tomada por la AGPV de incluir al menos tres miembros en cada panel de entrevista, desde el punto de vista psicométrico ha sido un gran acierto que ha llevado a alcanzar unos niveles de fiabilidad que caen dentro de lo que podría denominarse la excelencia. Como al propio tiempo la presencia en el panel de entrevista de miembros de la propia unidad en la que trabajará el nuevo ocupante garantiza que se están teniendo en cuenta los intereses de dicha unidad, tenemos que señalar que ésta decisión debe mantenerse en el futuro, aun cuando desde el punto de vista psicométrico pudiera aceptarse una menor fiabilidad (.83). Creemos que las ventajas superan con creces a los costes.

Es importante tener en cuenta también la información contextual que hemos obtenido y de la que hemos dado cuenta en las secciones precedentes. El hecho de que tanto los candidatos (evaluados) como los jueces (entrevistadores), como otras personas implicadas en el proceso de desarrollo de la entrevista (sujetos expertos en el puesto) hayan proporcionado un feedback y unos comentarios positivos sobre esta modalidad de entrevista muestra que, por regla general, las reservas iniciales que todo proceso de cambio suscitan han sido superadas satisfactoriamente y que en la actualidad el personal y las personas con responsabilidad en la AGPV ven en la entrevista conductual estructurada un instrumento adecuado e imprescindible para el propósito perseguido, a saber, seleccionar a la persona que más se ajuste al puesto de entre los candidatos que se presentan.

Basándonos en la experiencia pionera de la Administración General del País Vasco que, en cierto modo debe considerarse revolucionaria, consideramos que el uso de la entrevista conductual estructurada debe ampliarse a la totalidad de los puestos de la AGPV. Al propio tiempo, consideramos que otras administraciones públicas, tanto locales como autonómicas como estatales, deberían seguir el ejemplo de la AGPV y comenzar a utilizar entrevistas conductuales estructuradas en sus sistemas de selección y provisión de plazas, siempre y cuando las situaciones organizacionales lo permitan. Naturalmente, un requisito obligatorio sería utilizar dichas entrevistas con todo el rigor metodológico y práctico utilizado en la AGPV.

Queda pendiente un estudio que muestre validez predictiva del instrumento utili-

zado, es decir que muestre si o no las personas seleccionadas por la AGPV mediante la técnica de la entrevista conductual estructurada presentan mejores resultados en términos de desempeño de tarea, de desempeño contextual y, globalmente, consideradas, en su productividad. Estudios meta-analíticos previos en contextos de organizaciones privadas han mostrado que esto es así, por lo cual, teniendo en cuenta la fiabilidad encontrada en la presente investigación y que la fiabilidad y la validez se hayan relacionadas, creemos que no es aventurado plantear como hipótesis que el coeficiente de validez de las entrevistas conductuales estructuradas utilizadas por la AGPV será semejante o inclusive superior a los descritos en la actual literatura o lo que es lo mismo superiores a .45 (véase, Schmidt y Hunter, 1998; Salgado, 1999; Salgado y Moscoso, 1995; Salgado, Viswevaran y Ones, 2001). Sugerimos a los responsables de la AGPV que se inicien las tareas necesarias para poder llevar a efecto tal estudio.

REFERENCIAS BIBLIOGRÁFICAS

Arvey, R.D. y Campion, J.E. (1982). The employment interview: A summary and review of recent. *Personnel Psychology*, 35, 281-322.

Campion, M.A., Palmer, D.K. y Campion, J.E. (1997). A review of structure in the selection interview. *Personnel Psychology*, 50, 1-64.

Campion, M. A., Pursell, E. D. y Brown, B.K. (1988). Structured interviewing: Raising the psychometric properties of the employment interview. *Personnel Psychology*, 41, 25-42.

- Conway, J.M., Jako, R.A. y Goodman, D.F. (1995). A meta-analysis of interevaluator and internal consistency reliability of selection interviews. *Journal of Applied Psychology*, 80, 565-579.
- Fear, R.A. (1984). *The evaluation interview*. 3ª Edición. New York: McGrawHill.
- Flanagan, J.C. (1954). The critical incident technique. *Psychological Bulletin*, 51, 327-358.
- Glass, G.V., McGraw, B. y Smith, M.L. (1981). *Meta-analysis in social research*. Beverly Hills, CA: Sage.
- Ghiselli, E.E. (1966). The validity of a personnel interview. *Personnel Psychology*, 19, 389-394.
- Gorriti, M. (2002). Análisis Funcional de los Grupos A y B de la Administración General del País Vasco. *Revista Vasca de Administración Pública*, 63, 205-242.
- Gorriti, M. (en prensa). Análisis Ocupacional y de Estructuras Organizativas a partir de Componentes de Función y Análisis de Clusters: La Administración General del País Vasco un caso concreto de Aplicación. *Revista de Psicología del Trabajo y las Organizaciones*.
- Harris, M.M. (1998). The Structured Interview: What Constructs Are Being Measured?. En R. Eder y M. Harris (Eds.), *The Employment Interview: Theory, Research and Practice*. 2 Edición. Thousand Oaks, CA: Sage Publications.
- Hedges, L.V. y Olkin, I. (1995). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- Huffcutt, A. I. y Arthur, W, Jr. (1994). Hunter & Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79, 184-190.
- Hunter, J.E. y Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Hunter, J.E. y Schmidt, F.L. (1990). *Methods of meta-analysis. Correcting error and bias in research findings*. Newbury Park, CA: Sage.
- Hunter, J.E., Schmidt, F.L. y Jackson, G.B. (1982). *Meta-analysis: cumulating research findings across studies*. Beverly Hills, CA: Sage.
- Janz, T. (1982). Initial comparisons of patterned behavior description interviews versus unstructured interviews. *Journal of Applied Psychology*, 67, 577-580.
- Janz, T., Hellervik, L. y Gilmore, D.C. (1986) *Behavior Description Interviewing New, Accurate, Cost effective*. Newton MA: Allyn & Bacon.
- Latham, G. (1989). The reliability, validity, and practicability of the situational interview. En R.W. Eder y G. R. Ferris (eds). *The employment interview. Theory, research, and practice (pp 169-182)*. Newbury Park, CA: Sage.
- Latham, G.P., Saari, L.M., Pursell, E.D y Campion, M.A. (1980). The situational interview. *Journal of Applied of Psychology*, 65, 442-431.
- Locke, E.A. (1968). Toward a theory of task motivation and incentives. *Organiza-*

tional Behavior and Human Performance, 3, 157-189.

Locke, E.A. y Latham, G.P. (1984). *Goal setting: A motivational technique that works*. Englewood Cliffs, NJ: Prentice-Hall.

McDaniel, M.A., Whetzel, D.L., Schmidt, F.L. y Maurer, S. (1994). The validity of employment interview: A comprehensive review and meta-analysis. *Journal of Applied Psychology*, 79, 599-617.

Mayfield, E.C. (1964). The selection interview: A re-evaluation of published research. *Personnel Psychology*, 17, 239-260.

Morgan, H.H. y Cogger, J.W. (1973). *The interviewers manual*. New York: The Psychological Corporation

Motowidlo, S.J., Carter, G.W., Dunnette, M.D., Tippins, N., Werner, S., Burnett, J.R. y Vaughan, J. (1992). Studies of behavioral structured interview. *Journal of Applied Psychology*, 77, 571-588.

Quijano, S.D. de (1992). *Sistemas efectivos de evaluación del rendimiento: resultados y desempeños*. Barcelona: PPU.

Reilly, R.R. y Chao, G.T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1-62.

Rosenthal, R.R. (1984). *Meta-analytic procedures for social research*. Newbury Park, CA: Sage.

Salgado, J.F. (2002). *Curso de Formación en Entrevista Conductual Estructura-*

da. Vitoria-Gasteiz: Instituto Vasco de Administración y Pública y Dirección de la Función Pública, Consejería de Hacienda y Administración Pública, Gobierno Vasco.

Salgado, J.F. y Anderson, N. (2002). Cognitive and GMA testing in the European Community: Issues and Evidence. *Human Performance*, 15, 75-96.

Salgado, J.F. y Anderson, N. (2003). Validity generalization of GMA tests across countries in the European Community. *European Journal of Work and Organizational Psychology*, 12, 1-17.

Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C. y de Fruyt, F. (2003a). International Validity Generalization of GMA and Cognitive Abilities as Predictors of Work Behaviors. *Personnel Psychology*, in press.

Salgado, J.F., Anderson, N., Moscoso, S., Bertua, C., de Fruyt, F. y Rolland, J.P. (2003b). Criterion validity of General Mental Ability measures for different occupations in the European Community. *Journal of Applied Psychology*, in press.

Salgado, J.F. y Moscoso, S. (1995). Validez de las entrevistas conductuales estructuradas. [Validity of the behavior structured interviews]. *Psicología del Trabajo y las Organizaciones*, 11, 9-24.

Salgado, J.F. y Moscoso, S. (2001). *Entrevista conductual estructurada de selección de personal. Teoría, práctica y rentabilidad*. Madrid: Pirámide.

Salgado, J.F., Viswesvaran, C. y Ones, D.S. (2001). Predictors used for personnel selection: an overview of construct, met-

- hods and technichs. En N. Anderson, D.S. Ones, H.K. Sinangil y C. Viswesvaran (Eds.), *Handbook of Industrial, Work and Organizational Psychology. Vol 1* (pp 165-199). Londres: Sage
- Schmidt, F.L. y Hunter, J.E. (1977). Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*, 62, 529-540.
- Schmidt, F.L. y Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124, 262-274.
- Schmitt, N. (1976). Social and situational determinants of interview decisions: Implications for the employment interview. *Personnel Psychology*, 29, 79-101.
- Schneider, B. y Schmitt, N. (1986) *Staffing organizations*. 2ª Edición. Glenview, IL.: Scott, Foresman.
- Schüler, H. (1989). Construct validity of a multimodal employment interview. En Fallon, B.J., Pfister, H.P. y Brobner, J. (Eds.). *Advances in industrial and organizational psychology*. New York: North Holland.
- Schüler, H. y Funke, U. (1989). The interview as a multimodal procedure. En E.W. Eder y G.R. Ferris (Eds.). *The Employment interview: Theory, research and practice*. Newbury Park, CA: Sage.
- Smith, P. y Kendall, L.M. (1963). Retranslation of expectations: An approach to the construction of unambiguous anchors for rating scales. *Journal of Applied Psychology*, 47, 149-155.
- Ulrich, L. y Trumbo, D. (1965). The selection interview since 1949. *Psychological Bulletin*, 63, 100116
- Taylor, P.J. y Small, B. (2002). Asking applicants what they would do versus what they did do: A meta-analytic comparison of situational and past behaviour employment interview questions. *Journal of Occupational and Organizational Psychology*, 75, 277-294.
- Wagner, R. (1949). The employment interview: A critical review. *Personnel Psychology*, 2, 17-46.
- Wiesner, W.H. y Cronshaw, S.F. (1988). The moderating impact of interview format & degree of structure on interview validity. *Journal of Occupational Psychology*, 61, 275-290.
- Wright, O.R. (1969). Summary of research on the selection interview since 1964. *Personnel Psychology*, 22, 391-413.
- Wright, P.M., Lichtenfelds, P.A., Pursell, E.D. (1989). The structured interview: additional studies and meta-analysis. *Journal of Occupational Psychology*, 62, 191-199.