

# VALIDACIÓN DE PROCEDIMIENTOS DE SELECCIÓN EN LA POLICÍA LOCAL

Alfredo Berges Saldaña

Ayuntamiento de Zaragoza  
abergess@zaragoza.es

RESUMEN. El objetivo de este trabajo es examinar la validez de criterio de los procedimientos de selección de policías locales. Se presentan resultados correspondientes a dos estudios realizados en la Policía Local de Zaragoza. En el primero de los estudios participaron aspirantes admitidos (N=240); se utilizaron como variables predictoras, la habilidad mental general, personalidad medida con el modelo de los cinco factores y una medida de conocimientos relacionados con el trabajo policial. Los criterios a predecir fueron dos, el aprovechamiento de la formación medido a través de la calificación obtenida en la Academia y el desempeño profesional calificado por los mandos. En el segundo estudio (N=226) participaron policías en activo que aspiraban a una unidad operativa. Además de habilidad mental general y personalidad, se utilizaron como predictores una entrevista conductual estructurada, pruebas físicas y una valoración de méritos. El criterio utilizado fue desempeño en el trabajo. Se presentan resultados correlacionales y de regresión múltiple para ambos estudios. Palabras clave: validez de criterio y selección policial.

## INTRODUCCIÓN

La selección de policías es un ámbito de estudio de indudable interés para la Gestión de los Recursos Humanos en las Administraciones Públicas. Sin embargo, pocos son los trabajos empíricos dedicados a examinar la validez de criterio de los procedimientos de selección de policías. En este trabajo se va a examinar la validez de criterio de los procedimientos de selección utilizados en la policía local. Se presentan dos estudios distintos, llevados a cabo en

la Policía Local de Zaragoza, el primero de los estudios corresponde a los procedimientos utilizados para la selección de policías de nuevo ingreso, mientras que el segundo estudio se refiere a los procedimientos para la selección de policías en activo con destino a una unidad operativa de nueva creación.

La selección de policías suele utilizar una estrategia de dos fases. La primera fase trata de descartar a los aspirantes que no reúnen las condiciones adecuadas para desempeñar el trabajo de policía. En el caso de los policías se tiene especial cuidado en evaluar dimensiones de personalidad disfuncional que puedan interferir con el desempeño policial, dadas las especiales características de las interacciones de la policía con la ciudadanía y el hecho de que se trata de un tipo de personal que porta armas. La segunda fase trata de identificar a los mejores aspirantes que, si contamos con instrumentos de selección válidos, serán los que mejor realizarán el trabajo posteriormente. En este caso vamos a centrarnos en esta segunda fase de la estrategia de selección. La cuestión que tratamos de responder es, una vez descartados los aspirantes que se consideran no aptos, ¿qué validez tienen los instrumentos que utilizamos en predecir el desempeño en el trabajo?.

Otra característica de la selección de policías es que, una vez superadas las pruebas de acceso, hay un periodo de formación que suele tener carácter eliminatorio. Desde el punto de vista del examen de la validez de criterio, lo que esto significa es que, además del criterio desempeño en el trabajo, contamos con otro criterio anterior, el aprovechamiento de la formación. Es decir que el proceso de selección de policías comenzaría por identificar a los más aptos para aprender, quienes en última instancia tendrían un desempeño superior en el trabajo.

Un proceso de selección empieza por la identificación de cuáles son las características necesarias de un trabajador para realizar un trabajo. Esas características se derivan de la comprensión del contenido y del cometido del puesto, es decir de lo que hay que hacer y

cómo hacerlo desde el punto de vista de la organización. La metodología adecuada para alcanzar ese conocimiento es el análisis del puesto de trabajo (Yates 1977). Los análisis de puestos en policía son relativamente frecuentes. En la encuesta nacional de prácticas de selección de policías realizada por Cochrane, Tett y Vandecreeck (2003) en los Estados Unidos de América, un 74.5% de los Departamentos de Policía que contestaron la encuesta, afirmaron que habían llevado a cabo un análisis del puesto de trabajo con una evaluación sistemática de sus funciones, y que la información obtenida la ponían a disposición de los evaluadores en procesos de selección en casi el 80% de los casos. En un trabajo reciente, Salgado (2003) identificó qué variables incluir en un procedimiento de selección de policías. Partiendo de trabajos como los de Bernardin (1988), Gorriti (1991), Serrano y colaboradores (1986) y Spilberg (2003), que recogían todos ellos o bien funciones policiales, o bien características personales cognitivas y de personalidad necesarias para su realización, Salgado concluyó que para seleccionar policías es conveniente utilizar, al menos, una prueba de habilidad mental general y una prueba de personalidad que evalúe el modelo de los cinco factores.

Por su parte, además de la habilidad mental general y la personalidad, las Bases que rigen los procesos de selección introducen otros aspectos a evaluar que juegan un papel importante en las decisiones de selección en el caso de policías locales. En los procedimientos de selección que pretendemos validar añadimos conocimientos, aptitud física, méritos de formación y experiencia y entrevista.

## ESTUDIO 1

### Participantes

Participaron en el estudio 240 aspirantes que habían superado las pruebas previas de selección y realizaron con éxito el curso de formación en la Academia. La edad media era de 23 años.

El 87 por ciento eran varones frente al 13 por ciento de mujeres.

Predictores utilizados

### *Habilidad mental general*

Dadas las características de los participantes, todos ellos aspirantes en procesos de selección de personal y para evitar o minimizar los efectos de preparación o conocimiento de las pruebas, se utilizó una prueba de habilidad mental general, el test RLS (Metis, 2001), contrastada técnicamente pero no difundida a través de los canales comerciales de distribución de pruebas de evaluación psicológica.

El test RLS mide la capacidad para razonar utilizando material abstracto, para pensar de un modo lógico y solucionar problemas presentados de un modo esquemático. Los ítems de la prueba también implican la capacidad para establecer relaciones abstractas y planificar y establecer conclusiones a partir de premisas, es decir, razonar deductivamente. La capacidad para encontrar las reglas o conceptos generales a partir de casos o datos específicos, por tanto es capacidad para ir de lo particular a lo general e implica la habilidad para formar hipótesis sobre las relaciones entre los datos, es decir, razonar inductivamente.

El test consiste en 25 preguntas que incluyen un argumento de razonamiento lógico y el sujeto debe indicar cual es la respuesta a dicho argumento. En cada una de las preguntas, el sujeto dispone de cinco opciones de respuesta, siendo correcta solamente una de ellas. Para completar el test se dispone de 20 minutos.

Las fiabilidades por consistencia interna del test RLS, calculada mediante la fórmula alfa de Cronbach y test-retest fueron .83 y .81 respectivamente (Metis, 2001). La media de respuestas acertadas fue 19 con una desviación típica de 2.6.

### *Dimensiones de personalidad*

La medida de las dimensiones de personalidad se ha obtenido a través del *Inventario de Personalidad IP/5F* (Salgado, 1998). El inventario de Personalidad (IP/5F) tiene un total de 200 ítems que miden las cinco dimensiones de personalidad: estabilidad emocional, extraversión, conciencia, amigabilidad y apertura a la experiencia. Cada una de estas dimensiones se descompone en una serie de compuestos de ítems homogéneos que para toda la prueba alcanzan un total de 34. Para el desarrollo de la prueba se ha seguido el Modelo de los Cinco Factores de Personalidad o *Big Five*. Los datos técnicos arrojan coeficientes de fiabilidad adecuados. Las fiabilidades calculadas mediante el método Alfa de Cronbach que aporta el manual del *IP/5F* van desde .74 para amigabilidad, hasta .90 para estabilidad emocional.

Los estadísticos descriptivos básicos de las puntuaciones de personalidad obtenidas por los sujetos en las dos muestras aparecen en la Tabla 1.

---

INSERTAR AQUÍ LA TABLA 1

---

#### *Prueba de conocimientos técnicos relacionados con el trabajo*

Esta prueba constaba de dos partes. La primera consistía en una prueba objetiva de respuesta múltiple cuyo contenido era la legislación policial básica. La segunda parte consistía en resolver un supuesto práctico con la aplicación de legislación policial básica. La puntuación máxima que los aspirantes podían alcanzar era de 20 puntos. La media de los participantes en esta prueba fue de 13.12 y la desviación típica 1.65. El coeficiente de fiabilidad de Spearman-Brown alcanzó .63

## Medida de los criterios

Se utilizaron dos criterios para examinar la validez, el aprovechamiento de la formación y el desempeño en el trabajo.

### *Aprovechamiento de la formación*

El aprovechamiento de la formación fue medido a través de las calificaciones de los aspirantes en el curso selectivo. El curso tuvo una duración de cuatro meses y los participantes fueron formados en 26 materias diferentes. Para los cálculos utilizamos la calificación media de estas 26 materias obtenida por los participantes.

### *Desempeño en el trabajo*

Como medida de criterio de desempeño, que se ha desglosado en las dos facetas que en la actualidad aglutinan un mayor consenso entre los investigadores en esta materia, el desempeño de tarea (Campbel, 1990) y el desempeño contextual (Borman y Motowidlo, 1993), se ha utilizado la obtenida en el Cuestionario de Evaluación del Desempeño. El cuestionario recoge las calificaciones del personal en activo otorgadas por sus mandos en 10 dimensiones monoitem. Siete de esas dimensiones referidas a desempeño de tarea: (1) Solución de problemas, (2) Toma de decisiones, (3) Organización del trabajo, (4) Conocimiento del puesto, (5) Resistencia a la presión, (6) Prevención de riesgos y (7) Disciplina personal. Y tres con un contenido orientado a evaluar desempeño contextual: (8) Competencia interpersonal, (9) Iniciativa y (10) Desempeño extra rol. Las calificaciones se realizan en escalas de calificación con anclajes conductuales puntuadas de 1 a 5. La puntuación correspondiente a desempeño de tarea, es la suma de las dimensiones 1 a 7. La puntuación correspondiente a desempeño contextual es la suma de las dimensiones 8 a 10. El desempeño global es la suma de las diez dimensiones.

## Procedimiento

La primera prueba que realizaron los aspirantes fue la de conocimientos técnicos que tenía carácter eliminatorio. La prueba de habilidad mental general se aplicó en grupo y en una sola sesión, junto a la prueba de personalidad y al cuestionario de personalidad disfuncional, a partir de cuyos resultados se descartaron a una parte de los aspirantes pero que no forma parte del estudio que aquí presentamos. Las calificaciones del desempeño se recogieron entre año y medio y dos años tras la incorporación de los aspirantes seleccionados. Los mandos que calificaron a los policías recibieron una formación específica en la que se les adiestró en la forma en que debían realizar la calificación del desempeño para disminuir en la medida de lo posible los errores en la evaluación.

#### Análisis de datos

Realizamos dos tipos de análisis, el primero consistió en el cálculo de las correlaciones entre los predictores y los criterios. Además realizamos cálculos de regresión múltiple entre el conjunto de variables predictoras y cada uno de los criterios utilizados. Para contrarrestar los errores artefactuales, las correlaciones se corrigieron por falta de fiabilidad del criterio y por restricción del rango.<sup>1</sup> Para realizar los cálculos utilizamos el programa de software para el cálculo de la validez corregida VALCOR (Salgado, 1997).

#### Resultados

En la Tabla 2 se muestran los coeficientes de correlación corregidos por falta de fiabilidad del criterio y por restricción del rango.

---

INSERTAR AQUÍ LA TABLA 2

---

#### Predicción del aprovechamiento de la formación y del desempeño

---

1 Para conocer los valores utilizados en las correcciones dirigirse al autor.

La correlación observada entre la habilidad mental general y las calificaciones en la Academia ascendió a .32 ( $p < .01$ ). La validez corregida para el error de medida del criterio entre la habilidad mental general y las calificaciones en la Academia, ascendió a .36. Corregidas por restricción del rango, la validez alcanzó .44. Los coeficientes de validez obtenidos son algo más bajos que los obtenidos por los meta-análisis realizados en EEUU con muestras policiales (Aamodt, 2004; Hirsh et al, 1986). En el caso de los meta-análisis realizados en Europa, los resultados de Salgado y sus colaboradores (2003), obtienen un coeficiente de .25, pero los autores señalan que uno de los tres trabajos incluidos puede considerarse un “outlier”. Si se eliminara ese trabajo, el coeficiente pasaría a ser .48, similar al obtenido por nosotros para las calificaciones en la Academia.

La Tabla 2 también muestra las correlaciones obtenidas entre habilidad mental general y el criterio desempeño en el trabajo medido a través de una escala de calificación cumplimentada por los mandos. Como puede apreciarse, las correlaciones fueron cercanas a cero. El valor más elevado correspondió a la predicción del desempeño de tarea cuya validez operativa alcanzó .11. Nuestro resultado, aunque inesperado, no es único. Recientemente Lynn (2006) obtuvo una correlación de -.10 entre habilidad mental general y calificación de mandos realizada un año después de la salida de la Academia. Pero lo cierto es que, tanto los estudios primarios que informan sobre resultados de validez de la habilidad mental general con respecto a desempeño en policía, como los resultados meta-analíticos permiten concluir que hay evidencia suficiente para sostener esa validez. En el caso de los meta-análisis, los coeficientes están en torno a .20. En general los coeficientes obtenidos en el caso del desempeño en el trabajo son inferiores a los obtenidos en el aprovechamiento de la formación. Esto se ha interpretado por algunos autores como una limitación de los mandos para evaluar adecuadamente a los subordinados. En el caso de los puestos policiales los mandos



necesitarían más tiempo que en otros puestos (Hirsh et al., 1986; Schmidt, 2002). El trabajo del policía se realiza con frecuencia en solitario o en presencia de un compañero de patrulla. Consecuentemente el mando no tiene una observación directa de la ejecución del policía, lo que puede disminuir la fiabilidad de la evaluación del desempeño en tanto no transcurra un tiempo suficiente para obtener un conocimiento que permita realizar con más garantía la evaluación de los subordinados.

En nuestro caso, los policías llevaban entre un año y medio y dos años de ejercicio profesional y se habían incorporado un número de ellos más elevado de lo habitual en ambas convocatorias. Diversas circunstancias provocaron que se tratase del número más elevado de policías que se incorporaban en una misma promoción. Este gran volumen de policías incorporados a las unidades pudo impedir el grado de conocimiento y observación necesario para realizar una calificación fiable por parte de los mandos. Una segunda explicación a este hallazgo puede residir en el tiempo de experiencia en el trabajo de los policías de nuestra muestra. Un policía que supera las pruebas de selección, pasa por la Academia y se incorpora a una unidad, atesora un grado muy considerable de conocimiento declarativo pero un grado comparativamente menor de conocimiento procedimental. Si tenemos en cuenta que, de un tiempo a esta parte, en el caso de la Policía Local de Zaragoza, la tradición del entrenamiento de los “novatos” por parte de los “veteranos”, patrullando juntos, se ha perdido a causa de una rápida y numerosa renovación de las plantillas policiales, tal vez podamos concluir que la adquisición de ese conocimiento procedimental esté siendo más lenta y más difícil de evaluar por parte de los mandos. Una tercera explicación es que la habilidad mental general tenga efectos indirectos sobre el desempeño a través del conocimiento del trabajo y el aprovechamiento de la formación. Esto es lo que ponen de manifiesto los modelos explicativos de la relación entre habilidad mental general y desempeño (Hunter, 1983;

Schmidt, Hunter y Outerbridge, 1986).

En la Tabla 2 se muestran las correlaciones corregidas entre las cinco dimensiones de personalidad y las calificaciones en la Academia obtenidas por los aspirantes. Los coeficientes de validez observada iban de .07 para amigabilidad a .11 para conciencia. Los coeficientes de correlación observada se corrigieron por error de medida del criterio y por variabilidad en el rango. Hay que tener en cuenta que en el caso de conciencia y apertura a la experiencia, no existió restricción del rango sino amplificación del rango. La validez corregida más elevada correspondió a la dimensión estabilidad emocional con un coeficiente de .23. A continuación la dimensión extraversión con una validez operativa de .13. Le siguió en cuantía la dimensión conciencia con .11, y por último apertura a la experiencia y amigabilidad que predijeron las calificaciones de la Academia con valideces de .10 y de .08 respectivamente. Dadas las características de la muestra, la validez operativa de conciencia y de apertura a la experiencia fueron inferiores a las valideces corregidas solamente por falta de fiabilidad del criterio. Eso significa que probablemente las correlaciones observadas sobreestimaron el tamaño de la validez de ambas dimensiones.

La cuantía de los coeficientes de validez obtenida por nosotros fue más baja que la obtenida por Aamodt (2004) en su meta-análisis. En ese caso, el mayor poder predictivo correspondía a la dimensión apertura a la experiencia con un coeficiente de .34 seguido de conciencia con .21. En nuestros resultados, la importancia de las dimensiones fue diferente y era la dimensión estabilidad emocional la que presentó una validez más elevada. Por su parte, Extraversión obtuvo una validez ligeramente inferior a la obtenida por Aamodt (2004) y en último lugar, con valideces mucho más bajas, se situaron conciencia y apertura a la experiencia. No debemos olvidar que en estos últimos casos la variabilidad del rango ha provocado una disminución de la validez y que si observamos la validez corregida únicamente por fiabilidad

del criterio los valores eran los más elevados de todas las dimensiones. Por lo tanto, pese a que los valores son pequeños, cuatro de las cinco dimensiones de personalidad, estabilidad emocional, extraversión, conciencia y apertura a la experiencia, son predictores válidos del aprovechamiento de la formación.

Los resultados obtenidos, pueden compararse con los de algunos trabajos recientes. Por ejemplo, Meijer, Born, Terlouw y Van der Molen (2008) utilizaron una muestra de policías holandeses en periodo de formación. Obtuvieron medidas de las cinco dimensiones de personalidad para predecir tres criterios de desempeño del periodo de formación. Las medidas del criterio fueron obtenidas a través de escalas de calificación cumplimentadas por policías que supervisaban las prácticas de los aspirantes. El primer criterio se denominó mantenimiento del orden, el segundo ayuda a las víctimas y el tercero era una calificación global de la formación. Las medidas de formación que utilizamos en este trabajo difieren considerablemente de las utilizadas por Meijer et al (2008). En nuestro caso se trata de calificaciones de diferentes materias de conocimientos policiales, mientras que en el de los autores que comentamos, se calificaba la utilización de destrezas profesionales adquiridas en la Academia y desplegadas en la calle. Hecha esta salvedad, los coeficientes de validez más altos obtenidos por los autores, corresponden a la dimensión amigabilidad para predecir mantenimiento del orden (.13) y a la dimensión apertura a la experiencia para predecir ayuda a las víctimas con un coeficiente de -.13. Conciencia alcanzó solamente .10.

Por lo que se refiere a la predicción del desempeño en el trabajo, la única dimensión que alcanzó una validez apreciable aunque escasa, fue apertura a la experiencia. El resto de dimensiones obtuvieron valideces próximas a cero.

La correlación obtenida entre la prueba de conocimientos y las calificaciones de la Academia fue de .44. La validez operativa asciende a .49. Por lo que se refiere a la correlación de la

prueba de conocimientos con el desempeño en el trabajo, la Tabla 2 recoge unos valores obtenidos de .11 para la validez observada que asciende a .25 para la validez operativa en el caso del desempeño de tarea. En el caso del desempeño global la validez es de .24.

Las calificaciones de la Academia nos sirvieron como criterio del aprovechamiento de la formación para examinar la validez de los diferentes instrumentos utilizados en la selección de policías. Pero también utilizamos dichas calificaciones como predictor del desempeño en el trabajo, o lo que es lo mismo, para obtener la relación entre criterios diferentes. Como se muestra en la Tabla 2 las calificaciones de la Academia predijeron el desempeño posterior en el trabajo con una validez operativa de .22 para desempeño de tarea, de .20 para desempeño contextual y de .22 para desempeño global.

### *Predicción múltiple*

Calculamos la correlación múltiple a partir del coeficiente de determinación que proporciona la salida de datos de Lisrel 8.12a, cuando se realiza un path analysis<sup>2</sup> La correlación múltiple de los predictores con el desempeño global que obtuvimos fue de .33. Si pretendemos predecir el aprovechamiento de la formación, la correlación múltiple asciende a .61. En el primer caso estábamos explicando el 11% de la varianza del desempeño global. En el segundo caso, la varianza explicada era del 37%.

## ESTUDIO 2

### Participantes

Los participantes en el Estudio 2 fueron policías en activo aspirantes a ingresar en una unidad operativa. En total contamos con 226 aspirantes. Un 92 por ciento fueron hombres frente a un 8 por ciento de mujeres.

### Medida de los predictores

---

2 Realizamos varios path analysis para examinar relaciones causales entre las variables (ver Berges, 2010)

### *Habilidad Mental General y Personalidad*

En este caso se utilizaron las mismas pruebas comentadas en el Estudio 1. El RLS para medir habilidad mental general y el IP5F para medir personalidad.

### *Entrevista Conductual Estructurada*

Las entrevistas conductuales estructuradas, a diferencia de las tradicionales, requieren un desarrollo específico para su utilización en un procedimiento de selección. Para esta investigación la entrevista conductual estructurada se ha construido siguiendo la metodología propuesta por Salgado y Moscoso (2001). Los expertos que formaron parte del equipo de trabajo, constituido con esta finalidad, estaba compuesto por expertos en el trabajo de la policía local. Estos expertos eran policías y mandos designados como miembros de la Comisión de Selección que evaluaron a los aspirantes a puestos en una unidad operativa de policía local. Además se incorporaron agentes de policía con formación en psicología (licenciados) que voluntariamente se prestaron a participar en la investigación.

Bajo las instrucciones del autor, los miembros de la Comisión de Selección, así como agentes titulados en psicología, llevaron a cabo las tareas siguientes:

- producción de incidentes críticos del puesto de agente de policía en unidades operativas.
- selección de los incidentes críticos relevantes para las dimensiones del perfil profesional a seleccionar.
- elaboración de preguntas a formular a los candidatos.
- elaboración de anclajes conductuales para la valoración de las respuestas de los candidatos.

Para llevar adelante estas tareas, a los miembros del equipo de expertos se les adiestró

convenientemente en aspectos teórico-prácticos de selección de personal y en contenidos de entrevista. Los mandos miembros de la Comisión de Selección intervinieron valorando las entrevistas, para lo que recibieron el entrenamiento adecuado para esta función. Por su parte, los agentes titulados en psicología fueron los responsables de conducir las entrevistas de personal (aunque no intervinieron en su valoración) y recibieron un entrenamiento específico para la realización de la entrevista.

El formato final que adoptó la entrevista, así como el procedimiento general de aplicación, sigue el trabajo desarrollado por Sáez (2007) en la Policía Local de Vitoria. Las modificaciones puntuales en el contenido se realizaron para adecuar la entrevista a nuestra realidad. En el caso de Vitoria los aspirantes carecían de experiencia en trabajo policial, mientras que en Zaragoza los aspirantes eran policías en activo.

#### *Antigüedad*

La antigüedad se valoró a razón de .025 puntos por año completo y reconocido en el Cuerpo de Policía Local de Zaragoza (Media = .22 y Desviación Típica = .13).

#### *Cursos de formación y perfeccionamiento.*

Se valoraron aquellos que estuviesen directamente relacionados con las funciones a cumplir en la Unidad Operativa de destino. Valorados con .0054 puntos por hora de duración del curso (Media = .52 y Desviación típica = .36)

#### *Destrezas o experiencias relacionadas con el puesto de trabajo.*

Se valoraron destrezas como defensa personal o manejo de ordenadores. La puntuación máxima a obtener por este concepto era de .50 puntos. (Media = .18 y Desviación Típica = .09).

#### *Aptitud física*

Esta prueba consistió en un recorrido realizado en carrera sorteando conos, el transporte de un maniquí con un peso de 35 y 25 kilos para hombres y mujeres respectivamente, una zona de equilibrio y varias zonas de rapidez superando determinados obstáculos. La puntuación obtenida dependía del tiempo invertido en la realización del circuito (Media = 1.64 y Desviación típica = .36).

#### Medida del criterio

Como medida de criterio se utilizó la evaluación del desempeño realizada por los mandos de las unidades de origen de los aspirantes. El instrumento con el que se desarrolló la calificación por parte de los mandos es el mismo que se ha descrito para el Estudio 1.

#### Procedimiento

Las pruebas de habilidad mental general y de personalidad se aplicaron en grupo y en una misma sesión. La segunda prueba fue la de aptitud física y finalmente la entrevista conductual estructurada. La comisión de selección aplicó la valoración de los méritos de los aspirantes y los mandos de procedencia de los aspirantes calificaron su desempeño. Tanto las medidas de los predictores como del criterio se tomaron simultáneamente, por lo tanto en este caso realizamos un estudio de validez concurrente.

#### Análisis de datos

Al igual que en el Estudio 1, realizamos dos tipos de análisis, el primero consistió en el cálculo de las correlaciones entre los predictores y los criterios. Además, realizamos cálculos de regresión múltiple entre el conjunto de variables predictoras y cada uno de los criterios utilizados. Para contrarrestar los errores artefactuales, las correlaciones se corrigieron por falta de fiabilidad del criterio y por restricción del rango. Para realizar los cálculos utilizamos el programa de software para el cálculo de la validez corregida VALCOR (Salgado, 1997).

#### Resultados

### *Predicción del desempeño*

Los coeficientes de correlación corregidos entre habilidad mental general y desempeño en el trabajo alcanzaron unos valores de .24 para desempeño de tarea y desempeño global y .20 para desempeño contextual. Estos valores son muy similares a los encontrados en estudios con muestras de policías. Varios estudios meta-analíticos incluyeron muestras policiales; el de Hirsh et al. (1986) obtuvo, con una combinación de habilidades que se puede considerar una medida de habilidad mental general, un valor de .22. En el de Salgado et al. (2003a) la validez obtenida es la misma, .24. En el caso del meta-análisis de Aamodt (2004) la validez asciende a .27. Hay que tener en cuenta que en los casos de los meta-análisis comentados las correcciones efectuadas han incluido la corrección por restricción del rango. En nuestros datos, al tratarse de un diseño de validez concurrente en el que todos los casos se incluyeron, con independencia del resultado final del proceso de selección, no se aplicó ninguna corrección por restricción del rango. Lo cual no significa que no la hubiera, puesto que los policías que formaban parte de esta muestra superaron en su momento un proceso de selección, sus puntuaciones también podrían estar sometidas a este artefacto de medida.

Hubo una pequeña diferencia entre la validez de la habilidad mental general para predecir el desempeño de tarea y el desempeño contextual. Como sugieren Borman y Motowidlo (1993), el desempeño de tarea tiene que ver con habilidades cognitivas mientras que el desempeño contextual tiene que ver más con variables de personalidad. Este resultado, también puede interpretarse como la potencia que tiene la habilidad mental general para predecir resultados en el entorno laboral, más allá del desempeño de tarea, como han puesto de manifiesto Ones et al (2009).

En cuanto a las dimensiones de personalidad, la validez para desempeño de tarea ascendió a .21 para las dimensiones estabilidad emocional, a .21 para conciencia y a -.20 para



amigabilidad. Apertura a la experiencia obtuvo una validez de -.14. Amigabilidad una validez corregida de -.24 para desempeño contextual y de -.22 para desempeño global. Apertura a la experiencia obtuvo una validez de -.17 y de -.18 para desempeño contextual y desempeño global respectivamente. Con una cuantía algo inferior se situaron estabilidad emocional, .13 y conciencia, .11 en desempeño contextual, y .19 y .18 para desempeño global respectivamente. Lo que estos resultados nos indican es que las correlaciones de estabilidad emocional y conciencia predecían el desempeño de tarea y el desempeño global, pero no el desempeño contextual. La correlación de amigabilidad con desempeño fue negativa. Esta tendencia en la dirección de la correlación se había obtenido ya en la muestra del Estudio 1. Además, amigabilidad era la dimensión de las cinco que presentaba unos valores más elevados. Ambos resultados, la dirección de la correlación y la cuantía de la misma para la dimensión amigabilidad, son difíciles de interpretar. Se podría pensar que la función de mantenimiento del orden y de ejercicio de la autoridad fuese incompatible con puntuaciones altas en amigabilidad. Pero en la actuación policial hay actividades de ayuda que estarían teóricamente relacionadas positivamente con amigabilidad. Lo mismo podríamos decir con respecto al desempeño contextual, una relación inversa con amigabilidad resulta contraintuitiva. Excepto en el caso de amigabilidad, lo que los resultados muestran es que las dimensiones de personalidad predecían mejor el desempeño de tarea y el desempeño global y lo hacían peor en el caso del desempeño contextual. Por otra parte, los resultados no indicaban un poder predictivo mayor del factor conciencia que del resto de las dimensiones. Por lo que se refiere a las cuantías obtenidas, si se tiene en cuenta que no se aplicó corrección por restricción del rango, serían muy parecidas a las de los meta-análisis realizados con muestras policiales (Aamodt, 2004; Barrick y Mount, 1991; Barrick et al, 2001; Salgado 1997 y 2003b).

En la Tabla 4 se muestran los coeficientes de correlación obtenidos entre la puntuación de la

entrevista conductual estructurada y las calificaciones de los mandos de las unidades de origen de los aspirantes. La validez para desempeño de tarea fue de .49. Para desempeño contextual ascendió a .50. Por último, la validez corregida para desempeño global alcanzó .51. A la vista de estos resultados la entrevista conductual estructurada se ha confirmado como un magnífico predictor del desempeño profesional de los policías locales. De hecho las valideces corregidas son las más elevadas de todos los instrumentos de selección incluidos en el procedimiento general de selección utilizado por la Policía Local de Zaragoza.

Nuestros resultados están en la línea de los obtenidos por los meta-análisis realizados sobre validez de la entrevista de selección. Por ejemplo la validez obtenida por Wiesner y Cronshaw (1988) fue de .47. Por su parte McDaniel et al. (1994), obtienen una validez de .44 para las entrevistas que incluyen en la categoría de estructuradas en su meta-análisis. Huffcut y Arthur (1994) en los niveles 3 y 4 de estructuración de las entrevistas, obtienen una validez de .56 y .57 respectivamente. Por último Salgado y Moscoso (1995), en un grupo ocupacional mixto donde se encontraban estudios primarios sobre policía, encontraron una validez de .56.

Por lo tanto nuestros resultados no difieren de los hallazgos de los autores mencionados, teniendo en cuenta que la cuantía de la validez encontrada en nuestro caso solamente fue corregida por error de medida y no por restricción del rango. Como los trabajos mencionados no estaban centrados específicamente en la ocupación de policía, es interesante comparar nuestros resultados con referentes de trabajos exclusivamente realizados con policías. El único meta-análisis sobre entrevista y con muestras policiales es el realizado por Aamodt (2004). La validez que encontró, incluyendo ocho estudios, fue de .19. Como el propio autor señala el grado estructura de las entrevistas utilizadas en esos trabajos no se puede determinar con claridad y por lo tanto es difícil comparar con los resultados que se han presentado. Cabe concluir que las entrevistas a las que se hace referencia por parte de Aamodt (2004) no eran ni

estructuradas ni conductuales.

Por el contrario el trabajo de Sáez (2007) es muy similar al que hemos realizado nosotros pero con algunas diferencias que no permiten una comparación perfecta. Su muestra estaba constituida por aspirantes de nueva incorporación mientras que la nuestra incluía policías en activo. En su caso se corrigieron los coeficientes de validez observada por error de medida y por restricción del rango y en el nuestro no se aplicó ninguna corrección por restricción del rango. La validez obtenida por Sáez (2007) con corrección por error de medida fue de: .31 para desempeño de tarea, .18 para desempeño contextual y .29 para desempeño global. Como puede comprobarse los coeficientes obtenidos por nosotros están por encima, especialmente en el caso del desempeño contextual. Por otra parte, Sáez (2007) encontró una diferencia significativa entre la correlación de la entrevista conductual estructurada con desempeño de tarea y con desempeño contextual, lo que confirmaba una de las hipótesis que había establecido en su trabajo. En nuestros datos no aparece esa diferencia, incluso la cuantía de la correlación es ligeramente superior para el desempeño contextual.

Por lo que se refiere a la antigüedad, como pueden verse en la Tabla 4, las valideces obtenidas fueron muy bajas por lo que ésta no predecía el desempeño profesional en nuestra muestra. Este resultado está en la línea de lo obtenido por McEvoy y Cascio (1989) que realizaron un meta-análisis con casi cuarenta mil sujetos correspondientes a 96 estudios para examinar la validez de criterio de la edad. Los resultados mostraron una correlación media de .06 entre edad y desempeño. Las otras dos variables de valoración de méritos, las destrezas y la formación, obtuvieron coeficientes de correlación más altos, especialmente las destrezas profesionales con una validez operativa que alcanzó casi .40 con las calificaciones del desempeño. La validez operativa de la valoración de los cursos de formación estaba en torno a .30, como se muestra en la Tabla 4. Las correlaciones entre la aptitud física y el desempeño

estaban en torno a .20.

### *Predicción múltiple*

Como en el caso del Estudio 1, calculamos los coeficientes de correlación múltiple a partir del coeficiente de determinación en los resultados de un path analysis realizado con Lisrel 8.12a. La correlación múltiple entre los predictores y el criterio asciende a .74. Eso quiere decir que el modelo explica el 54% de la varianza del desempeño global.

## DISCUSIÓN Y CONCLUSIONES

Este trabajo pretendía examinar la validez de los procedimientos utilizados para la selección de policías locales. Los resultados muestran que la habilidad mental general tiene una validez moderada para predecir el aprovechamiento de la formación. Respecto a la predicción del desempeño, el papel de la habilidad mental general está probablemente mediado por la adquisición del conocimiento del trabajo y la formación (Hunter, 1980. Schmidt, Hunter y Outerbridge, 1986)

Por lo que se refiere a las dimensiones de personalidad, en general las valideces obtenidas han sido escasas. Varios autores han realizado estudios primarios (Cortina, Doherty, Schmitt, Kaufman y Smith, 1992; Meiring, Van De Vijver, Rothmann, De Bruin y Foxcroft, 2006; Sanders, 2008 y Meijer, Born, Terlouw y Van der Molen, 2008) que examinan la validez predictiva de las dimensiones de personalidad del Modelo de los Cinco Factores con muestras policiales. La validez obtenida es baja y en algunos casos próxima a cero. Este hallazgo, que no es exclusivo de la selección policial, ha hecho que en la literatura académica se haya generado recientemente un debate respecto a la conveniencia o no, de utilizar medidas de personalidad para adoptar decisiones de incorporación de personas (Morgeson, Campion, Dipboye, Hollenbeck, Murphy y Schmitt, 2007; Murphy y Dzieweczynski, 2005; Ones, Dilchert, Viswesvaran y Judge, 2007; Tett y Christiansen, 2007). Nuestra opinión es que,

aunque la validez de las dimensiones de personalidad sea pequeña, lo cierto es que, en combinación con otros predictores, contribuirán a incrementar la varianza explicada total del desempeño y por lo tanto su utilidad no puede ponerse en duda.

La entrevista conductual estructurada se ha mostrado como el mejor predictor del desempeño en el trabajo. En este sentido nuestros resultados no difieren de los encontrados en estudios anteriores. En un sentido, nuestros datos difieren con los obtenidos en otros trabajos similares. En concreto con los de Salgado, Moscoso y Gorriti (2007) y Sáez (2007). En ambos casos se planteaba la hipótesis de que la entrevista conductual estructurada obtendría correlaciones más elevadas con el desempeño de tarea que con el desempeño contextual. Esta hipótesis se establecía de acuerdo con los hallazgos acerca de la validez de constructo de la entrevista de selección (Salgado y Moscoso, 2002). En ambos casos los datos permitieron confirmar la hipótesis. En nuestro caso la entrevista conductual estructurada ha mostrado el mismo grado de validez de criterio para las dos facetas del desempeño. Este hecho, en nuestra opinión, puede tener dos explicaciones. La primera explicación es que en nuestro caso la correlación entre las facetas de desempeño es considerablemente más alta que en el caso de Sáez (2007). En nuestro caso esta correlación es de .83, mientras que en el trabajo de Sáez es de .65. Una segunda explicación de estos resultados puede estar relacionada con el propio perfil de habilidades de los policías en activo y con un grado relativamente alto de experiencia profesional. Es posible que en estas condiciones, las dimensiones evaluadas por medio de la entrevista conductual estructurada, expliquen tanto el desempeño de tarea como el desempeño contextual. A las bondades psicométricas de la entrevista conductual estructurada, hay que añadir un aspecto muy relevante en el entorno de la Administración Pública como es el de su aceptación y la percepción positiva que genera como instrumento de selección de personal y que han sido investigadas por Salgado, Moscoso y Gorriti (2004) en el Gobierno Vasco.

Por último, los resultados indican que las aptitudes físicas y los aspectos valorados como méritos tienen una validez entre escasa y moderada para predecir el desempeño de trabajo en policías locales.

Los estudios que hemos presentado tienen interesantes repercusiones prácticas, pueden contribuir a realizar modificaciones en los procedimientos de selección que se suelen utilizar en la selección de policías locales y, muy especialmente en la forma de valorar y/o ponderar la importancia de unas y otras pruebas.

Desde un punto de vista teórico, el trabajo supone una cierta novedad en nuestro país y permite acumular conocimiento en un ámbito de indudable interés para la psicología jurídica y para la psicología del trabajo y de las organizaciones. Por otra parte el trabajo tiene limitaciones, en especial en lo relativo a la excesiva localización de las muestras lo que dificulta la generalización de los resultados.

## BIBLIOGRAFÍA

- Aamodt, M.G. (2004). *Research in Law Enforcement Selection*. Boca Raton: Brown Walker Press.
- Barrick, M.R., y Mount, M.K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44, 1–26.
- Barrick, M.R., Mount, M.K., y Judge, T.A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next?. *International Journal of Selection and Assessment*, 9, 9–30.
- Berges, A. (2010). Validación de procedimientos de selección de personal en la Policía Local. Tesis Doctoral no publicada. Universidad de Zaragoza.
- Bernardin, H.J. (1988). Police officer. En S. Gael (Ed), *The job analysis handbook for business, industry, and government*. Vol. II (pp 1242-1254). New York: Wiley.
- Borman, W.C. y Motowildo, S.J. (1993) Expanding the criterion domain to include elements of contextual performance. En Schmitt, N. Y Borman, W.C. *Personnel Selection in Organizations*. Pags. 71-98. Jossey Bass.
- Campbell, J.P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. En Dunnette, M Hough, L.M. *Handbook of Industrial and Organizational Psychology*. 1, 2ª ed. Pags. 687-731. Consulting Psychologist Press.
- Cochrane, R.E. Tett, R.P. y Vandecreek, L. (2003) Psychological Testing and The Selection of police Officers. A National Survey. *Criminal Justice and Behavior*, 30, 511-537.
- Cortina, J.M., Doherty, M.L., Schmitt, N., Kaufman, G. y Smith, R.G. (1992) The “Big Five” Personality Factors in the IPI and MMPI: Predictors of police performance. *Personnel Psychology*, 45, 119–140.
- Gorriti, M. (1991). *Un estandar de selección para el curso básico de la Ertzaintza, según el*

- modelo de validez predictiva*. Tesis Doctoral Inédita. Bilbao: Universidad de Deusto.
- Hirsh, H. R., Northrop, L., y Schmidt, F. L. (1986). Validity generalization results for law enforcement occupations. *Personnel Psychology*, 39, 399–420.
- Huffcutt, A. I. y Arthur, W, Jr. (1994). Hunter y Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology*, 79, 184-190.
- Hunter, J.E. (1983). A Causal Analysis of Cognitive Ability, Job Knowledge, Job Performance, and Supervisor Ratings. En Landy, F., Zedeck, S. y Cleveland J. (eds.) *Performance Measurement and Theory*. Hillsdale, NJ: Lawrence Erlbaum Associates, pages 257-266.
- Lynn, J. (2006). *Investigation officer police selection procedures*. Tesis doctoral no publicada. Wichita State University.
- 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.
- McEvoy, G.M. y Cascio, W.F. (1989). Cumulative evidence of the relationship between employee age and job performance. *Journal of Applied Psychology*, 74, 11-17.
- Meijer, L.A.L., Born, M.Ph., Terlouw, Gert y Van der Molen H.T.(2008). Criterion-related validity of dutch police-selection measures and differences between ethnic groups. *International Journal of Selection and Assessment*, 16, 321-332.
- Meiring, D., Van De Vijver, F, Rothmann. S.R., De Bruin, G. P. y Foxcroft, C.(2006). Internal and external bias of cognitive and personality measures in south africa, 18th International Congress of the International Association for Cross-Cultural Psychology (IACCP), Spetses, Greece.
- Metis (2001). Manual técnico de la Batería de Habilidades Cognitivas. Documento no publicado.
- Morgeson, F.P., Campion, M.A., Dipboye, R.L., Hollenbeck, J.R., Murphy, K. y Schmitt, N.



- (2007) Reconsidering the use of Personality Tests in Personnel Selection Contexts. *Personnel Psychology*, 60, 683–729.
- Murphy, K.R. and Dzieweczynski, J.L. (2005). Why Don't Measures of Broad Dimensions of Personality Perform Better as Predictors of Job Performance? *Human Performance*, 18, 343–357.
- Ones, D. S., Dilchert, S., Viswesvaran, C., y Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, 60, 995- 1027.
- Ones, D.S., Dilchert, S., Viswesvaran, C. y Salgado, J. (2009). Cognitive ability. En J.L. Far y N.T. Tippins(Eds.). *Handbook of employe selection*. Mahwah, N.J: Erlbaum.
- Schmidt, F.L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, 15, 187-210.
- Schmidt, F.L., Hunter, J.E., y Outerbridge, A.N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology*, 71, 432–439.
- Sáez, J. (2007). Diseño y validación de una entrevista conductual estructurada para la selección de agentes de policía local. *Revista de psicología del Trabajo y de las Organizaciones*. 23, 57-74.
- Salgado, J.F. (1997a). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, 82, 30–43.
- Salgado, J.F. (1997b). VALCOR: A program for estimating standard error, confidence intervals, and probability of corrected validity. *Behavior Research Methods, Instruments, & Computers*, 29, 464–467
- Salgado, J.F. (1998). *Manual del Inventario de Personalidad de Cinco Factores (IP/5F)*. Tórculo Ediciones. Santiago de Compostela.

- Salgado J.F. (2003b). La Selección de Policías y Personal de Ocupaciones de Seguridad: Recientes Avances de la Investigación. Informe Técnico GRHR03-02. Universidad de Santiago de Compostela.
- Salgado, J.F. (2003c). *La hipótesis de especificidad situacional, la hipótesis de la generalización de la validez y la predicción del desempeño ocupacional: influencia de la personalidad*. Manuscrito no publicado. Universidad de Santiago de Compostela.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., y de Fruyt, F. (2003). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology*, 56, 573–605.
- Salgado, J. F., Gorriti, M. y Moscoso, S. (2007). La entrevista conductual estructurada y el desempeño laboral en la Administración pública Española. *Revista de psicología del Trabajo y de las Organizaciones*, 23, 39-55.
- Salgado, J.F. y Moscoso, S. (1995). Validez de las entrevistas conductuales estructuradas. *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. y Moscoso, S. (2002). Comprehensive meta-analysis of the construct validity of the employment interview. *European Journal of Work and Organizational Psychology*, 11, 299-324.
- Salgado, J.F., Moscoso, S., Gorriti, M. (2004). Investigaciones sobre la Entrevista Conductual Estructurada (ECE) en la Selección de Personal en la Administración General del País Vasco: Meta-Análisis de Fiabilidad. *Psicología del Trabajo y las Organizaciones*, 20, 107-140.
- Sanders, B.A. (2008). Using personality traits to predict police officer performance. *Policing: An*

*International Journal of Police Strategies & Management*, 31, 129-147.

Serrano, P., García-Sevilla, L., Pérez, J.L., Pina, M. y Ruiz, J.R. (1986). Municipal police evaluation: psychometric versus behavioral assessment. En J.C. Yuille (Ed.), *Police selection and training: The role of psychology* (pp. 257-265). Dordrecht: Martinus Nijhoff Publishers.

Spilberg, S.W. (2003). *Development of psychological screening guidelines for peace officers: Background and development of essential traits*. Presentado en la 16 Conferencia Annual de la Society for Industrial and Organizational Psychology, Orlando, FL.

Tett, R. P., y Christiansen, N. D. (2007). Personality tests at the crossroads: A response to Morgeson, Campion, Dipboye, Hollenbeck, Murphy, & Schmitt (2007). *Personnel Psychology*, 60, 967-99

Wiesner, W.H., y Cronshaw, S.F. (1988). The moderating impact of interview format and degree of structure on interview validity. *Journal of Occupational Psychology*, 61, 275-290.

Yates, R. E. (1977). A Job analysis of the FBI special agent position. In Ch. D. Spielberger y H. C. Spaulding (Eds.). *"National Working Conference on the Selection of Law Enforcement Officers"*. Tampa (Fl): University of South Florida. 29-41.

	Media	D. Típica
Estabilidad Emocional	59.61	6.8
Extraversión	56.82	8.9
Apertura a la Experiencia	35.03	13.2
Amigabilidad	43.68	8.1
Conciencia	19.79	14.3

	Aprovechamiento de la formación	Desempeño de Tarea	Desempeño Contextual	Desempeño Global
Habilidad Mental General	<b>.44</b>	.11	.05	.10
Estabilidad Emocional	.23	.05	.16	.08
Extraversión	.13	.00	.00	.00
Apertura a la Experiencia	.10	<b>.12</b>	<b>.12</b>	<b>.12</b>
Conciencia	.11	-.04	.00	-.03
Amigabilidad	.08	-.10	-.04	-.09
Conocimientos	<b>.49</b>	<b>.25</b>	.20	<b>.24</b>
Aprovechamiento de la formación	--	<b>.22</b>	<b>.20</b>	<b>.22</b>

Nota.- Coeficientes de correlación corregidos por falta de fiabilidad del criterio y por restricción del rango. En negrita se indican los coeficientes cuyos límites del intervalo de confianza no incluyen el "0".

	Media	D. Típica
Estabilidad Emocional	53.6	10.6
Extraversión	57.4	9.4
Apertura a la Experiencia	43.8	7.4
Amigabilidad	43.1	13.5
Conciencia	26.6	14.5

	Desempeño de Tarea	Desempeño Contextual	Desempeño Global
Habilidad Mental General	<b>.24</b>	<b>.20</b>	<b>.24</b>
Estabilidad Emocional	<b>.21</b>	.13	<b>.19</b>
Extraversión	-.01	.03	.00
Apertura a la Experiencia	-.14	-.17	-.18
Conciencia	<b>.20</b>	.11	.18
Amigabilidad	<b>-.20</b>	<b>-.24</b>	<b>-.22</b>
Entrevista Conductual Estructurada	<b>.48</b>	<b>.50</b>	<b>.51</b>
Antigüedad	.04	-.09	.00
Destrezas Profesionales	<b>.37</b>	<b>.39</b>	<b>.39</b>
Cursos de Formación	<b>.28</b>	<b>.26</b>	<b>.28</b>
Aptitudes Físicas	<b>.20</b>	<b>.19</b>	<b>.20</b>

Nota.- Coeficientes de correlación corregidos por falta de fiabilidad del criterio. En negrita se indican los coeficientes cuyos límites del intervalo de confianza no incluyen el "0".