

LA CAPACIDAD DE LOS POLICÍAS PARA DETECTAR MENTIRAS

HERNÁN ALONSO, JAUME MASIP Y EUGENIO GARRIDO

Universidad de Salamanca

La elaboración de este trabajo ha sido posible merced a la ayuda recibida del Ministerio de Educación y Ciencia, Proyecto SEJ2006-06606. Los autores también desean expresar su agradecimiento al personal del Centro de Formación de la Policía Nacional en Ávila por haber permitido en diversas ocasiones el acceso a muestras de policías, así como a Guadalupe Lucas por su colaboración en este trabajo.

Resumen: En este trabajo se examina la investigación empírica internacional sobre la capacidad de los policías para diferenciar entre verdades y mentiras. Dicha capacidad es muy limitada y similar a la del resto de las personas, y además los policías tienden a juzgar las declaraciones como falsas. Se explora la influencia de la experiencia policial y del entrenamiento para detectar mentiras sobre los juicios de veracidad y se proponen algunas estrategias para mejorar las capacidades de los policías.

Palabras clave: Policía, Detección de la Mentira, Entrenamiento, Sesgo de Mendacidad.

Abstract: In this paper, the empirical research on police officers' ability to distinguish between truths and lies is reviewed. The officers' detection ability is very poor and similar to that of lay persons. In addition, police officers tend to judge the statements as deceptive. The influence that police experience and deception-detection training may have on veracity judgments is examined, and some strategies to improve police officers' abilities are described.

Key words: Police, Officers, Deception Detection, Lie Detection, Training, Lie Bias.

1. Introducción

Cuando se comete un delito, es labor de la policía identificar y detener al posible culpable (sobre las funciones del Cuerpo Nacional de Policía véanse, por ejemplo, Abrunhosa, Soeiro y Ventura, 2006; Álvarez, 2000; Clemente, 1995; Fernández, 1990; 2003; Rico, 1988). Sin embargo, a menudo la investigación policial de las pruebas físicas no basta para solucionar los casos. De hecho, es frecuente que el análisis científico y el escrutinio de las pruebas materiales no se empleen hasta que la policía ya haya identificado a algún sospechoso mediante una entrevista o un interrogatorio. Además, hay muchos delitos que no dejan evidencias físicas incontrovertibles, o éstas desaparecen al poco tiempo (tal es el caso, por ejemplo, del abuso sexual infantil; véanse Cantón, 1997; Lamb, 1994; Masip y Garrido, 2007). En tales circunstancias, la obtención de información del testigo por parte de la policía adquiere la máxima importancia. Asimismo, toda información sobre los sentimientos, los pensamientos y las intenciones de los testigos, las víctimas o los agresores debe obtenerse también mediante una entrevista policial. Todos estos factores hacen que la entrevista sea de capital importancia para el trabajo de la policía, en especial durante las fases iniciales de la investigación (Bull, 1999; Milne y Bull, 1999; Vrij, 2008). Sin embargo, una entrevista policial es un tipo de situación en la que algunas personas pueden mentir. En consecuencia, una parte sustancial de la labor de la policía consiste en diferenciar entre declaraciones verdaderas y falsas. La detección de la mentira, por lo tanto, es muy importante para el trabajo policial (por ejemplo, Garrido y Masip, 1999).

La investigación sobre la capacidad de las personas para detectar mentiras mediante la observación de la conducta ha arrojado resultados desalentadores (véase Masip, 2005). Así, en un amplio meta-análisis¹ de los juicios de veracidad efectuados por 384 muestras de observadores (que incluían un total de 24.483 personas) que, en su conjunto, tuvieron que juzgar la veracidad de 6.651 mensajes emitidos por 177 muestras de emisores (con un total de 4.435 personas), se

¹ Un meta-análisis es un trabajo en el cual se toman los datos de un amplio conjunto de estudios previos y se analizan conjuntamente utilizando sofisticadas y rigurosas técnicas estadísticas. Al comprender un conjunto más amplio y heterogéneo de participantes y condiciones que un estudio individual, sus hallazgos son más representativos y generalizables. Los meta-análisis también permiten examinar la influencia sobre las medidas de ciertas variables en las cuales difieren distintos conjuntos de estudios. Para más información véanse, por ejemplo, Botella y Gambara, 2002, 2006).

halló que el nivel promedio de aciertos (precisión) fue del 54% (Bond y DePaulo, 2006). Esta cifra se sitúa significativamente por encima del 50% esperado por azar, pero no sólo está muy lejos de la perfección (100%), sino que también se encuentra muy alejada del mínimo requerido para hacer juicios con cierta garantía de éxito. Aamodt y Custer (2006) hicieron un meta-análisis menos amplio con 108 estudios, hallando resultados idénticos. En otras palabras: la capacidad de los seres humanos para discriminar entre verdades y mentiras es extremadamente limitada.

Sin embargo, en la mayoría de los estudios incluidos en estos meta-análisis los participantes eran estudiantes universitarios sin ningún interés especial en la detección del engaño. Cabe preguntarse si las personas que desempeñan trabajos en los cuales la detección de las mentiras es importante (jueces, abogados, policías o guardias civiles, agentes de aduanas, inmigración o inteligencia, psiquiatras o psicólogos clínicos, etc.) también alcanzan niveles tan bajos de precisión. En este trabajo pretendemos responder a esta pregunta centrándonos, específicamente, en los policías. Para ello examinamos la literatura científica relevante.

2. La precisión de los policías

Se han llevado a cabo muchos estudios para examinar la capacidad de los policías para juzgar la veracidad de un conjunto de declaraciones. Sus niveles de precisión se han comparado entonces con los de observadores legos, es decir, personas cuya labor profesional no exige detectar el engaño. En ambos casos, los juicios de veracidad se hacen sin ninguna herramienta, instrucción específica o información sobre qué indicios utilizar; simplemente a partir de la intuición y las creencias de cada cual. Pues bien, quizás sorprenda constatar que tales estudios no muestran ninguna superioridad entre los policías a la hora de discriminar entre declaraciones verdaderas y falsas.

En la Tabla 1 se ofrecen los resultados de los trabajos relevantes que hemos podido localizar. Se incluyen sólo los datos de aquellos grupos de participantes que no recibieron ningún entrenamiento. En concreto, en la Tabla se muestran los datos de 27 estudios que incluyen un total de 45 muestras de policías (con un total de 2.647 personas) y 3 muestras de agentes de la ley de otra naturaleza (CIA, Servicio Secreto y *sheriffs*; 137 personas) que, por sus características

Tabla 1. Porcentajes de precisión en la detección del engaño

Estudio	Participantes	N	Precisión		
			Verdades	Mentiras	Global
<i>Akelhurst, Bull, Vrij y Köhnken (2004)</i> ^a	Policías	26			68
	Trabajadores sociales	14			75
	Estudiantes	18			57
<i>Bond (2007)</i> ^b	Agentes de la ley Estudiantes	112	59	45	52
		122	66	39	52
Prueba 2	Agentes expertos Estudiantes	8	55	63	59
		8	63	33	48
<i>DePaulo y Pfeiffer (1986)</i>	Policías veteranos	144			52
	Policías noveles	114			53
	Estudiantes	161			54
<i>Ekman y O'Sullivan (1991)</i>	Jueces	110			57
	Policías federales	60			56
	Policías	126			56
	Agentes Servicio Secreto	34			64
	Psiquiatras	67			58
	Personas interesadas	73			55
	Estudiantes	39			53
<i>Ekman, O'Sullivan y Frank (1999)</i>	Agentes de la CIA	23			73
	Sheriffs	43			66
	Policías	36			51
	Jueces	84			62
	Psicólogos clínicos interesados ^c	107			68
	Psicólogos clínicos sin interés ^d Psicólogos académicos ^e	209 125			62 58

Tabla 1. (Continuación)

Estudio	Participantes	N	Precisión		
			Verdades	Mentiras	Global
<i>Elaad (2003)</i>	Policías	60	48	51	49
<i>Garrido, Masip y Herrero (2004)</i>	Policías	121	26	69	47
	Estudiantes	146	51	66	59
<i>Hartwig, Granhag, Strömwall y Kronkvist (2006)^f</i>	Policías		57	55	56
<i>Hartwig, Granhag, Strömwall y Vrij (2004)</i>	Policías	30			57
<i>Kassin, Meissner y Norwick (2005)</i> Condición visual	Detectives de la policía	28			42
	Estudiantes	29			53
Condición auditiva	Detectives de la policía	29			55
	Estudiantes	32			64
<i>Köhnken (1987)^g</i>	Policías	20	63	32	47
<i>Leach, Talwar, Lee, Bala y Lindsay (2004)^h</i> Experimento 1	Policías	25			44
	Agentes de aduanas	48			49
	Estudiantes	32			51
Experimento 2	Policías	28			63
	Agentes de aduanas	35			62
Experimento 3	Estudiantes	40			64
	Policías	47			59
	Agentes de aduanas	40			57
	Estudiantes	39			61

Tabla 1. (Continuación)

Estudio	Participantes	N	Precisión		
			Verdades	Mentiras	Global
<i>Mann y Vrij (2006)</i>	Policías	27	70	66	68
<i>Mann, Vrij y Bull (2004)</i>	Policías	99	64	66	65
<i>Mann, Vrij, Fisher y Robinson (2008)</i>					
Condición visual	Policías	37	34	53	43
Condición auditiva	Policías	31	63	69	66
Condición audiovisual	Policías	35	60	70	65
<i>Masip, Alonso, Garrido y Barba (2008)</i>					
Policías veteranos	Policías veteranos	72	48	61	54
Policías noveles	Policías noveles	72	54	61	57
Estudiantes	Estudiantes	72	58	51	55
<i>Masip, Garrido y Herrero (2003)</i>	Policías	224	30	74	52
<i>Meissner y Kassir (2002)</i>					
Policías	Policías	44			50
Estudiantes	Estudiantes	20			56
<i>Porter, Woodworth y Birt (2000)ⁱ</i>					
Policías	Policías	32	20	60	40
Estudiantes	Estudiantes	32	42	42	42
<i>Vrij (1992)</i>	Detectives de la policía	91	51	46	49
<i>Vrij (1994)ⁱ</i>					
Declaraciones espontáneas	Detectives de la policía				50
Declaraciones planificadas	Detectives de la policía				47
Cuerpo entero	Detectives de la policía				51
Sólo manos y sin sonido	Detectives de la policía				44

Tabla 1. (Continuación)

Estudio	Participantes	N	Precisión		
			Verdades	Mentiras	Global
<i>Vrij, Akehurst, Brown y Mann (2006)</i> Emisores niños (5 a 6 años)	Profesores (n = 36), trabajadores sociales (n = 37), policías (n = 37) y estudiantes (n = 40)	150	61	63	62
Emisores adolescentes (14 a 15 años)	Profesores (n = 36), trabajadores sociales (n = 37), policías (n = 37) y estudiantes (n = 40)	150	67	57	62
Emisores adultos	Profesores (n = 36), trabajadores sociales (n = 37), policías (n = 37) y estudiantes (n = 40)	150	58	63	60
<i>Vrij y Mann (2001a)</i>	Policías	65	70	57	64
<i>Vrij y Mann (2001b)</i>	Policías	52			51
<i>Vrij, Mann, Fisher, Leal, Milne y Bull (2008)</i> Orden cronológico	Policías	24	50	42	46
Orden inverso	Policías	31	56	60	58
<i>Vrij, Mann, Kristen y Fisher (2007)</i> Entrevista facilitadora	Policías	26			52
Entrevista BAI	Policías	18			51
Entrevista acusatoria	Policías	24			48
<i>Vrij, Mann, Robbins y Robinson (2006)^k</i>	Policías	37	70	73	72
<i>Vrij y Winkel (1993)</i>	Detectives de la policía	91	51	46	49

Nota: Las celdas se han dejado vacías si la información correspondiente no se ofrece en el artículo original. N: Número de personas participantes.
^a Fase de preentrenamiento, los emisores fueron niños; ^b Entre los agentes de la ley se incluyó a *sheriffs*, policías locales, agentes de fronteras, del FBI, de la CIA y del BIA (*Bureau of Indian Affairs*); ^c psicólogos clínicos en la detección del engaño; ^d psicólogos clínicos sin un interés especial en la detección del engaño; ^e psicólogos académicos con distintas orientaciones profesionales, excepto la clínica; ^f grupo sin entrenamiento; ^g grupo control; ^h los emisores fueron niños; ⁱ grupos sin entrenamiento; ^j grupo sin información, N = 360, pero el autor no indica el número de detectives asignados a cada condición; ^k cada participante hizo cuatro pruebas similares de detección del engaño, los datos son el promedio de precisión en las cuatro pruebas.

específicas, consideraremos aparte². El resto de muestras están formadas por otros profesionales y por estudiantes incluidos como grupos control con los que comparar la ejecución de los profesionales.

2.1. Precisión global

Aún siendo conscientes de que las condiciones experimentales y los tamaños muestrales de cada estudio son diferentes, si se promedian las tasas de precisión global (la alcanzada al considerar conjuntamente declaraciones verdaderas y falsas) de las 45 muestras de policías que aparecen en la Tabla 1 se obtiene una precisión media del 54,02%, lo cual no dice mucho a favor de su habilidad³. En la Tabla 1 aparecen 17 condiciones en las que se compara la precisión de policías y estudiantes. Si, a partir de estos datos, se calcula la precisión media de policías ($M = 54.68\%$) y estudiantes ($M = 56.06\%$) en este subconjunto de estudios, se observa que fue casi la misma.

En general, estos datos coinciden con los de dos recientes meta-análisis en los que se ha comparado la precisión de profesionales y no profesionales (Aamodt y Custer, 2006; Bond y DePaulo, 2006). En el meta-análisis de Aamodt y Custer (2006) se observó que la precisión de policías ($M = 55,30\%$), detectives ($M = 50,80\%$), agentes aduaneros ($M = 55,30\%$), agentes federales ($M = 54,54\%$), agentes del Servicio Secreto ($M = 64,12\%$) y agentes encargados del control de la libertad condicional de reclusos ($M = 40,42\%$) no era sustancialmente distinta que la de los estudiantes ($M = 54,22\%$). La precisión hallada por Aamodt y Custer (2006) entre los policías es similar a la observada en la presente revisión, y no difería de la de los estudiantes.

En el meta-análisis de Bond y DePaulo (2006), se comparó la precisión de los «expertos» en detección del engaño (policías, agentes de aduanas, detectives, agentes federales, polígrafistas profesionales,

² No consideraremos a estos agentes de la ley conjuntamente con los policías (excepto en el estudio de Gary Bond, 2007, ya que el autor no indica cuántos profesionales de cada tipo había) por dos razones. Primera, porque el interés de este trabajo se centra en policías y no en agentes de inteligencia. Segunda, porque la experiencia laboral y la formación de estos profesionales puede diferir de la del resto de los policías, dadas las características específicas de su trabajo.

³ En el estudio de Vrij, Akehurst, et al. (2006), los autores indican que no hubo diferencias significativas entre la precisión de los profesores, trabajadores sociales, policías y estudiantes, y ofrecen los datos de estos cuatro grupos conjuntamente. Así pues, para el cálculo de la precisión de los trabajos expuestos en la Tabla 1, se atribuyó a los policías en cada condición del estudio de Vrij, Akehurst, et al. (2006) la precisión global que estos autores ofrecen.

jueces, especialistas en interrogatorios, auditores, técnicos de selección y delincuentes) con la de personas que no tenían ningún interés especial por la detección del engaño. Los primeros mostraron una precisión del 53,81%, y los segundos del 53,29%, sin que hubiera diferencias significativas entre ambos grupos.

Si bien los niveles de precisión de los estudios de la Tabla 1 son ya de por sí poco satisfactorios, cabe tener en cuenta que, de hecho, podrían estar incluso sobreestimados, dados los atípicos resultados de algunos estudios, debidos, quizás, a cuestiones metodológicas. En concreto, algunos trabajos del grupo de Aldert Vrij (Akehurst et al., 2004; Mann y Vrij, 2006; Mann et al., 2004; Mann et al., 2008; Vrij, Akehurst, et al., 2006; Vrij, Mann, et al., 2006), sobresalen por haber encontrado unas tasas de precisión muy elevadas ($M = 65,47\%$ para las muestras obtenidas de estos estudios tomadas conjuntamente). Pasamos a examinar más de cerca las particularidades de estos estudios con el fin de explicar tan atípicos resultados.

En el trabajo de Akehurst et al. (2004) se contó con 26 policías que realizaron dos tareas de detección del engaño: una sin entrenamiento, y posteriormente otra tras haber recibido cierto entrenamiento. Nos centraremos aquí en la tarea pre-entrenamiento, ya que hay algunos detalles que podrían haber influido en los resultados de la misma. En primer lugar, cada uno de los 26 policías (un tamaño muestral reducido) tenía que juzgar sólo cuatro declaraciones, y de hecho hubo algunos que, debido a problemas de tiempo, sólo pudieron juzgar tres. Cuando el número de participantes y de declaraciones a juzgar es tan reducido, la posibilidad de obtener resultados por azar aumenta sensiblemente. Frank y Feeley (2003) recogen la opinión de la gran mayoría de investigadores de la detección del engaño cuando aconsejan que, en las pruebas de detección, el número de declaraciones a juzgar esté en torno a 10.

Además, a cada policía se le presentaron declaraciones elegidas al azar, con lo que algunos sólo juzgaron declaraciones verdaderas, otros sólo declaraciones falsas, y otros declaraciones de ambos tipos. Aunque podríamos esperar que unas combinaciones contrabalanceasen las otras, hay que tener en cuenta que un policía que tuviese un sesgo hacia la mentira (véase Meissner y Kassin, 2002) e hiciese cuatro juicios de mentira, obtendría una precisión del 100% al juzgar cuatro mentiras, o nula al juzgar cuatro verdades. Las puntuaciones no reflejarían la precisión real, sino el sesgo de los policías. En ambos casos estaríamos ante puntuaciones extremas que, teniendo en cuenta lo reducido de la muestra, podrían marcar diferencias importantes en uno u otro sentido al calcular la precisión media.

Por último, la precisión del 57% de los estudiantes es similar a la que se ha encontrado en otros estudios; sin embargo, una precisión del 75% como la de los trabajadores sociales se sitúa muy por encima de los resultados habituales. Sorprende que estas puntuaciones sean tan diferentes y, sin embargo, no se encontraran diferencias significativas entre los tres grupos.

Respecto a los trabajos de Mann y Vrij (2006), Mann et al. (2004), Mann et al. (2008) y Vrij, Mann et al. (2006), en todos ellos los investigadores partieron del mismo conjunto inicial de 54 declaraciones, emitidas por un total de 14 personas, para elaborar los vídeos. Estas declaraciones presentan una serie de particularidades, entre las cuales cabe destacar que muestran a delincuentes reales mintiendo o diciendo la verdad a la policía sobre su implicación en delitos reales. Esto tiene una gran relevancia para examinar la mentira y su detección en la vida real. En sus diversas publicaciones, el equipo de Vrij propone varias explicaciones de los elevados índices de precisión alcanzados por los policías con estas declaraciones: (a) que muestran el tipo de personas (sospechosos) con los cuales la policía está más acostumbrada a tratar; (b) que el contexto (una entrevista policial) y los temas (implicación en un delito) son también familiares para la policía, (c) que lo que había en juego en estos casos (ser detenido, juzgado y probablemente condenado) era mucho más importante que lo que se gana o se pierde en los típicos estudios de laboratorio sobre la mentira; y (d) que, quizás, los sospechosos que aparecen en esos vídeos son malos mentirosos cuyas mentiras son fáciles de detectar. Una clara limitación de estos estudios es que no se han podido mostrar los vídeos a personas no policías, ya que los investigadores no disponen permiso para ello. Esto hace que no podemos afirmar sin lugar a dudas que la policía británica examinada por el equipo de Vrij y colaboradores sea especialmente capaz de detectar este tipo de mentiras (explicaciones a y b). Quizás, si observadores no policías hubieran visto las mismas declaraciones, habrían obtenido similares niveles de precisión, lo cual favorecería las explicaciones c y d. Mientras no se disponga de un grupo control de no policías con el cual comparar la precisión de los policías, no se puede afirmar que éstos tengan una mayor precisión.

Hay, además, otros factores que pueden haber alterado los resultados de este conjunto de estudios. En concreto, la duración de las declaraciones varió ostensiblemente (de 6 a 145 segundos). No es recomendable la utilización de declaraciones excesivamente cortas, ya que pueden carecer de elementos en los que basarse para emitir un juicio. Además, no sabemos si las declaraciones verdaderas y falsas diferían en duración, y no es adecuado comparar declaraciones muy

distintas en este aspecto, ya que la duración puede tener un efecto sobre los juicios (por ejemplo, Masip, Garrido y Herrero, 2006, 2009).

En los estudios de Mann et al. (2004) y Vrij, Mann, et al. (2006) se utilizaron todas las declaraciones. Las diferencias en los tiempos propiciaron diferencias en el número de declaraciones contenidas en cada uno de los cuatro vídeos presentados a los participantes en estos estudios (15, 16, 10 y 13 declaraciones, respectivamente). Aunque los cuatro vídeos tenían más o menos la misma duración, no es lo mismo juzgar 10 declaraciones que juzgar 16, ni se tarda el mismo tiempo. Esto pudo dar lugar a más fatiga en unas condiciones que en otras. Tampoco estuvo equilibrado el número de veces que cada declarante podía aparecer (un mínimo de 2 veces y un máximo de 8). Cuando un declarante aparece en más de una ocasión, se corre el peligro de no juzgar la veracidad de su declaración en sí misma, sino por comparación de su conducta con la mostrada en otras declaraciones (Granhag y Strömwall, 1999). Además, hay evidencia de que la familiaridad con la conducta (veraz) de un emisor incrementa la precisión de los juicios (por ejemplo, Feeley, deTurck y Young, 1995).

En el estudio de Vrij, Mann, et al. (2006), los propios autores señalan que la veracidad real era más fácil de reconocer en algunas declaraciones que en otras. Así, la precisión más baja para una declaración fue del 14% y la siguiente del 30%. Para cinco declaraciones la precisión se situó entre el 31% y el 45%, para otras cuatro entre el 46% y el 55%, para otras 10 entre el 56% y el 70%, para otras nueve entre el 71% y el 80%, para otras 11 entre el 81% y el 90% y para otras 13 entre el 91% y el 100% (Vrij, Mann, et al., 2006). Estos datos muestran dos cosas. Primero, que había gran variabilidad entre las declaraciones: en algunas la veracidad era muy fácil de detectar y en otras muy difícil. Segundo, no obstante lo dicho, casi la mitad de las declaraciones fueron juzgadas correctamente por más del 80% de los participantes. Esto puede indicar que los participantes eran muy buenos detectores, pero también que muchas de las declaraciones eran muy «transparentes». En este último caso, la precisión final se habría visto inflada por haber empleado declaraciones cuya veracidad era muy obvia.

Este último estudio, cuyo objetivo principal consistía en observar la consistencia en la precisión de los policías a través de cuatro tareas de evaluación de la veracidad (en cada una de las cuales se empleó uno de los cuatro vídeos), presenta otra importante limitación: los autores no indican si los participantes tuvieron la oportunidad de hablar entre ellos entre una y otra prueba. Si éste fuera el caso, pudiera ser que los juicios de unos se hubieran visto influidos por los juicios de

otros, quizás por los de aquellos policías más respetados y admirados por su capacidad para detectar mentiras. Sin embargo, la precisión global no aumentó de forma continua con cada prueba, y la obtenida en la cuarta prueba no fue significativamente mayor que la obtenida en la primera. Por lo tanto, es poco probable que los contactos entre los participantes inflaran sustancialmente los índices de acierto.

En los estudios de Mann y Vrij (2006) y Mann et al. (2008) se emplearon sólo 14 (una de cada sospechoso) de las 54 declaraciones de los trabajos anteriores. Aunque se presentó a los detectores el mismo número de verdades y mentiras (al menos así se indica en Mann et al., 2008), las otras limitaciones de las declaraciones a las que hemos hecho referencia son las mismas.

Finalmente, la investigación de Vrij, Akehurst et al. (2006) también presenta ciertos inconvenientes. El objetivo de este trabajo era examinar la habilidad de profesores, trabajadores sociales, policías y estudiantes para detectar el engaño en niños de 5 y 6 años, adolescentes de 14 y 15 años y adultos. Para ello se crearon tres vídeos distintos con 12 declaraciones en cada uno: dos verdades y dos mentiras por cada categoría de emisor (niños, adolescentes y adultos). Los participantes realizaron la tarea de detección distribuidos en grupos, cada uno de los cuales tuvo que juzgar las declaraciones contenidas en uno de los tres vídeos. El aspecto problemático es que en la hoja de respuestas en la que los participantes tenían que reflejar sus juicios también se preguntaba si el emisor *parecía nervioso*, si parecía estar haciendo un gran *esfuerzo mental*, si estaba intentando *parecer sincero* y si había en la declaración *suficiente información* para emitir un juicio. Hay razones para pensar que todas estas cuestiones pudieron haber influido en los juicios. Así, en el meta-análisis de DePaulo, Lindsay, Malone, Muhlenbruck, Charlton y Cooper (2003) se halló que la impresión general de nerviosismo es un indicio útil para detectar el engaño. Además, la investigación ha mostrado que mentir es cognitivamente más complejo que decir la verdad (Vrij, Fisher, Mann, y Leal, 2006), lo que puede reflejarse en un mayor esfuerzo mental (Vrij y Heaven, 1999). Finalmente, desde la perspectiva autopresentacional del engaño (véase DePaulo, 1992) se plantea que un empeño excesivo en mostrarse sincero puede ser un indicio de engaño. Pero la verdadera evidencia de que fijarse en estas claves incrementó la precisión de los participantes de los estudios de Vrij, Akehurst et al. (2006) la encontramos en sus propios resultados. Los autores señalan que los participantes indicaron que los emisores parecían más nerviosos, tenían que hacer un mayor esfuerzo mental e intentaban en mayor medida parecer sinceros al mentir que al decir la verdad (especialmente cuando tales emisores eran adolescentes). Asimismo,

ante emisores adultos, los observadores manifestaron en mayor medida haber recibido la suficiente información cuando las declaraciones eran verdaderas que cuando eran falsas. Además, Vrij, Akehurst et al. Encontraron que las puntuaciones en nerviosismo, esfuerzo mental e intentar parecer sincero correlacionaban positivamente con la precisión al juzgar mentiras y negativamente con la precisión al juzgar verdades. En definitiva, las preguntas que cada emisor tenía que contestar tras ver cada declaración pudieron haber centrado la atención de los participantes sobre indicios válidos de engaño, incrementando así la precisión de sus juicios.

Todas estas limitaciones de los estudios del equipo de Vrij invitan a la cautela antes de concluir que los policías británicos que participaron en los mismos son detectores particularmente buenos. En concreto, sería deseable que hubiera observadores no policías que pudieran juzgar la veracidad de las heterogéneas declaraciones empleadas en los trabajos de Mann y Vrij (2006), Mann et al. (2004), Mann et al. (2008) y Vrij, Mann et al. (2006), actuando así de grupo control de las muestras policiales. De hecho, si se promedian los datos de los policías de la Tabla 1 exceptuando las muestras de los seis estudios a los que acabamos de hacer referencia, la precisión media obtenida ($M = 51,97\%$) se acercaría más al nivel del azar.

2.2. *Precisión al juzgar verdades y mentiras: el sesgo de mendacidad*

En general, la investigación realizada con no profesionales ha encontrado mayores tasas de precisión al juzgar verdades que al juzgar mentiras (Levine, Park y McCornack, 1999). Esto se debe a que se emiten más juicios de verdad que de mentira, es decir, los observadores cometen un sesgo hacia la verdad (véanse las revisiones de Bond y DePaulo, 2006; Kalbfleisch, 1985; Vrij, 2000). Es curioso, sin embargo, observar que esta tendencia se invierte en los policías, observándose en la mayoría de ocasiones una precisión superior para las declaraciones falsas que para las verdaderas (véase Meissner y Kassin, 2002). Esto se debe a que los policías tienden a hacer más juicios de mentira que de verdad, es decir, muestran un sesgo de mendacidad. En la Tabla 1 se ofrece por separado la precisión para verdades y mentiras en 22 muestras de policías (en este caso no se tienen en cuenta los datos de Vrij, Akehurst, et al., 2006). Al promediar los datos, se observa que la precisión para las declaraciones verdaderas ($M = 52,77\%$) es sensiblemente inferior que la precisión para las declaraciones falsas ($M = 58,18\%$).

2.3. Conclusiones

En definitiva, de los datos derivados de los estudios de la Tabla 1 se desprenden las siguientes conclusiones generales: (a) las tasas de precisión de los policías a la hora de distinguir verdades de mentiras se encuentran cerca del nivel del azar; (b) en este sentido, la precisión de los policías es igual de pobre que la de no policías; y (c) los policías, al contrario que los no policías, muestran una mayor precisión al juzgar mentiras que al juzgar verdades.

2.4. ¿Por qué es tan baja la precisión?

Vrij (2004), tras constatar la baja precisión en la detección del engaño, sintetiza en diez puntos las posibles causas de la pobre ejecución de los profesionales (Tabla 2).

Tabla 2. Razones por las que la precisión de los profesionales es muy limitada

- | |
|--|
| <ul style="list-style-type: none">a. Ausencia de indicadores absolutos de la mentira.b. Diferencias muy sutiles al mentir vs. al decir la verdad.c. Toma de decisiones de manera heurística.d. Adherencia a reglas conversacionales que dificultan la detección.e. Confianza en indicios no diagnósticos.f. Atención a canales poco reveladores.g. Desestimación de las diferencias interindividuales de los declarantes (error de idiosincrasia).h. Desestimación de la influencia de factores situacionales (diferencias intraindividuales).i. Desestimación de las diferencias interindividuales de los detectores.j. Empleo de técnicas de entrevista e interrogatorio inadecuadas. |
|--|

Ausencia de indicadores absolutos de la mentira. No existe ningún indicio verbal o no verbal cuya presencia se relacione indefectiblemente con la mentira en cualquier persona y situación.

Diferencias muy sutiles al mentir vs. al decir la verdad. Los indicios que probabilísticamente podrían indicar que alguien está mintiendo generan solamente diferencias muy sutiles entre la conducta al mentir y al decir la verdad. Tales diferencias son difíciles de percibir.

Toma de decisiones de manera heurística. Según Vrij (2004), puede que los humanos emitamos nuestros juicios de veracidad basándonos más en reglas muy simples de decisión (heurísticos) que en un análisis pormenorizado de la conducta y la declaración del emisor. De hecho, hay bastante evidencia experimental en este sentido (por ejemplo, Levine et al., 1999; Masip, Garrido y Herrero, 2006, 2009; Stiff, Kim y Ramesh, 1992). Esta forma de juzgar la credibilidad puede llevar a decisiones erróneas. Sin embargo, no está claro que los policías decidan sobre la veracidad de los sospechosos de manera heurística. Así, por ejemplo, Masip, Garrido y Herrero (2003) encontraron que los policías tendían más que los estudiantes a hacer sus juicios hacia el final de la declaración, mientras que los estudiantes tendían más que los policías a hacerlo al principio. Esto podría indicar que los policías no actuaron bajo la influencia de un heurístico de decisión (en cuyo caso hubieran emitido sus juicios precipitadamente al principio de la declaración), sino que estaban procesando la información que iban recibiendo durante la declaración, y no tomaron su decisión sobre la veracidad hasta el final, cuando ya disponían de toda la información (véase Masip, Garrido, Herrero, Antón y Alonso, 2007). Además, una toma de decisiones heurística parece dar lugar no a un sesgo de mendacidad como el de los policías, sino, por el contrario, a juzgar las declaraciones como verdaderas (Stiff et al., 1992).

Adherencia a reglas conversacionales que dificultan la detección. Hay una regla conversacional que establece que hay que mirar a los ojos de la persona con la que se habla. Así pues, un interlocutor que observase otras partes del cuerpo del emisor en busca de indicios de engaño estaría rompiendo esta regla y, quizás, provocando malestar en el emisor. No hay que olvidar que la investigación empírica ha mostrado que la detección del engaño a partir de la observación del cuerpo es superior a la obtenida a partir de la observación del rostro (Ekman, 1992; Kalbfleisch, 1985; Masip y Garrido, 2000; Zuckerman, DePaulo y Rosenthal, 1981).

Confianza en indicios no diagnósticos. Tal como Vrij (2004) indica, las claves en las cuales los policías basan sus juicios están entre aquellas que la investigación ha mostrado que carecen de valor para diferenciar verdades de mentiras (véanse también Garrido y Masip, 1999; Sporer y Schwandt, 2007).

Atención a canales poco reveladores. Los observadores no sólo prestan atención a indicios concretos poco diagnósticos, sino también a canales poco transparentes. Por ejemplo, se presta más atención al contenido no verbal de una declaración que al verbal, pero éste último es más revelador que aquel (Masip y Garrido, 2000; Zuckerman et

al., 1981), especialmente si quien tiene que juzgar la veracidad no está familiarizado con el tema. Sólo en aquellos casos en los que los policías tienen declaraciones discrepantes de distintas personas sobre un mismo hecho parecen centrarse más en la comparación del contenido. De lo contrario, dirigen su atención principalmente a aspectos no verbales de la conducta.

Desestimación de las diferencias interindividuales de los declarantes. Pese a que existen importantes diferencias en las formas y estilos en que la gente puede ofrecer una declaración, tanto a nivel verbal como no verbal, a menudo tales diferencias individuales no se tienen en cuenta, y se espera que todo el mundo muestre conductas similares en situaciones similares. No tener en cuenta la idiosincrasia individual en la comunicación puede hacer que se juzgue erróneamente a un inocente como culpable (por ejemplo, si es una persona que muestra claros síntomas de nerviosismo), o a un culpable como inocente (si se trata, por ejemplo, de una persona muy fría y poco emocional). Ekman (1992) llamó a este fenómeno «riesgo de Brokaw» o «error de idiosincrasia».

Desestimación de la influencia de factores situacionales. Además de la idiosincrasia a nivel personal, hay que tener en cuenta la especificidad de cada situación. Una persona puede comportarse de diferente manera en diferentes situaciones. Pensar que alguien ha de comportarse siempre igual o no tener en cuenta el efecto de las características de la situación puede hacer que se emitan juicios erróneos. Por ejemplo, la conducta verbal y no verbal de una persona no será la misma cuando hace una primera declaración sobre un hecho que cuando lleva varias horas siendo interrogada.

Desestimación de las diferencias interindividuales de los detectores. Vrij (2004) también plantea que es posible que existan amplias diferencias individuales entre detectores, siendo algunos profesionales muy buenos y otros muy malos. Por ejemplo, en el estudio de Mann et al. (2004) la precisión de los profesionales abarca desde el 30% hasta el 90%.

Empleo de técnicas de entrevista e interrogatorio inadecuadas. Vrij (2004) también plantea que las bajas tasas de acierto pueden deberse al empleo de unas entrevistas y técnicas de interrogatorio policial que no favorecen la detección, sino que más bien pueden introducir sesgos que generan juicios erróneos (véase, por ejemplo, el trabajo de Kassir y Fong, 1999).

Para otras razones de la escasa precisión de los juicios de veracidad, véase Masip, Garrido y Herrero (2002).

3. La confianza en la precisión de los propios juicios

Ante estos, quizás sorprendentes, resultados respecto a la precisión de los policías, cabría preguntarse si ellos mismos son conscientes de sus propias limitaciones a la hora de juzgar la veracidad; es decir, ¿en qué medida los policías confían en lo acertado de sus juicios de veracidad? DePaulo, Charlton, Cooper, Lindsay y Muhlenbruck (1997), en un meta-análisis sobre la relación entre confianza y precisión, encontraron que los humanos solemos sobreestimar la precisión de nuestros juicios. Además, hallaron que la relación entre confianza y precisión es prácticamente nula. Este último resultado ha sido replicado recientemente en el meta-análisis de Aamodt y Custer (2006).

En diversos estudios hechos con policías se han incluido escalas para medir la confianza en los propios juicios. Generalmente, en estos casos, lo que se ha hecho es pedir a los policías que, tras cada uno de los juicios de veracidad, señalen en una escala tipo Likert la medida en que creen que el juicio que acaban de emitir es correcto. En muchos de estos trabajos se ha encontrado que no hay relación entre la confianza y la precisión de los juicios. Además, los policías suelen sobreestimar la precisión de sus juicios (Akehurst et al., 2004; DePaulo y Pfeiffer, 1986; Ekman y O'Sullivan, 1991; Garrido, Masip y Herrero, 2004; Kassin, Meissner y Norwick, 2005; Mann et al., 2004; Meissner y Kassin, 2002; Vrij y Mann, 2001a, b; Vrij, Mann, et al., 2006).

Entre los estudios que han comparado las estimaciones de los profesionales sobre lo acertado de sus juicios con las estimaciones de participantes no profesionales, es interesante el de Garrido et al. (2004). En este trabajo se halló que la precisión global de los policías no era superior que la de los no policías, sin embargo tanto policías como no policías consideraban que, en general, los miembros del Cuerpo Nacional de Policía tenían una mayor habilidad que el ciudadano medio para discriminar entre verdades y mentiras. Pero, además, esta creencia fue más fuerte entre los policías que entre los no policías. No obstante, no se encontró ninguna correlación significativa entre capacidad estimada y precisión, ni en la muestra de policías ni en la de no policías. Sobra decir que los juicios efectuados con mucha confianza pueden ser, presumiblemente, difíciles de cambiar incluso aunque sean erróneos (véanse las consideraciones de Garrido et al., 2004).

Otros resultados interesantes al respecto son los de Vrij, Mann, Kristen y Fisher (2007), que examinaron el efecto de tres estilos de en-

entrevista policial sobre la cantidad de indicios obtenidos, la precisión de una muestra de policías al juzgar las declaraciones y la confianza en los propios juicios. Los tres estilos de entrevista considerados fueron:

- a) *Information-Gathering Interview*, que podríamos traducir como «entrevista para recoger información», y a la que nos referiremos como «entrevista facilitadora». Este tipo de entrevista se caracteriza por pedir al interrogado que describa los hechos lo más detalladamente posible, utilizando principalmente preguntas abiertas.
- b) *Accusatory Interview*, o «entrevista acusatoria», en la cual los entrevistadores se enfrentan de forma reiterada al entrevistado con acusaciones directas.
- c) *Behavioral Analysis Interview (BAI)*, que podríamos traducir como «entrevista para el análisis de la conducta». Esta entrevista consta de dos partes. En un primer momento se empieza planteando al interrogado distintas preguntas abiertas, para pasar después a una segunda fase en la que se formulan 15 preguntas estandarizadas ante las cuales se espera que quienes mientan muestren una conducta diferente que quienes dicen la verdad.

Los resultados muestran que aunque el tipo de entrevista empleado no influyó sobre la precisión de los juicios, sí que influyó sobre la confianza de los policías en tales juicios. En concreto, la confianza fue mayor cuando la declaración se había obtenido con una entrevista de tipo acusatorio o con una BAI que cuando se había obtenido con una entrevista facilitadora. Además, ante declaraciones obtenidas con una entrevista acusatoria, la confianza de los policías correlacionaba positivamente con el número de juicios de mentira efectuados. Vrij et al. (2007) argumentan que los efectos negativos de la baja precisión de los policías podrían atenuarse si éstos no tuvieran tanta confianza en sus juicios. Plantean que la escasa confianza al emplear entrevistas facilitadoras podría disuadirles de emitir juicios de veracidad antes de tiempo, haciendo que dedicasen más tiempo a investigar el caso más a fondo. En definitiva, la estrategia de disminuir la confianza en los juicios podría, en último término, aumentar la precisión.

Finalmente, Kassir, Leo, Meissner, Richman, Colwell, Leach y La Fon (2008) pasaron un cuestionario a 631 policías que, entre otras cosas, tuvieron que estimar el porcentaje de veces que pensaban que sus juicios sobre la veracidad de los sospechosos eran correctos. Sus estimaciones fueron del 77%, muy superior al nivel de aciertos

del 54,02% señalado anteriormente. Esto muestra una vez más que los profesionales sobreestiman su eficacia en la discriminación de verdades y mentiras.

4. Consideraciones adicionales sobre la precisión

Se ha dejado para el final de este apartado una serie de interrogantes que, de forma adicional, podrían plantearse al observar los datos de la Tabla 1. Estos interrogantes se refieren tanto a ciertas variables que podrían influir de forma diferencial en la precisión de los policías (tales como las características de los emisores, la modalidad de presentación de las declaraciones o el tipo de entrevista con la que se obtuvieron las declaraciones) como a la posibilidad de que haya grupos especialmente dotados dentro de los profesionales.

4.1. *¿Existen profesionales en los que se haya observado una precisión superior?*

En algunas investigaciones se afirma que existen profesionales especialmente capaces de detectar mentiras. Esto es lo que sostienen O'Sullivan y Ekman (2004), que han puesto en marcha el llamado *Wizards Project* («Proyecto de los Magos») para identificar a los «magos» de la detección del engaño, es decir, aquellas personas que tienen una habilidad especial para juzgar la veracidad de otros, de forma que sus tasas de precisión están muy por encima de las del resto de las personas.

Para identificar a estos supuestos «magos» de la detección, los autores sometieron a tres pruebas diferentes a un total de 12.000 personas, entre las que se incluían terapeutas, personal de las fuerzas de la ley, jueces, abogados, mediadores y artistas. En primer lugar, se presentó a los participantes una tarea consistente en juzgar la veracidad de diez declaraciones grabadas en vídeo, cinco verdades y cinco mentiras, en las que el declarante ofrecía su supuesta opinión sobre un tema (que podía ser la prohibición de fumar en lugares públicos o si los delincuentes condenados por asesinato a sangre fría debían ser ejecutados). Se seleccionó entonces a aquellas personas con una precisión en esta prueba igual o superior al 90%, y se les presentaron dos pruebas más, con diez declaraciones filmadas en vídeo (de nuevo, cinco verdaderas y cinco falsas) en cada una. En una prueba tales declaraciones trataban de las emociones que el hablante estaba experimentando, y en la otra sobre un robo de dinero. El criterio inicial que una persona debía cumplir para ser calificada como

«mago» consistía en obtener una precisión igual o superior al 80% en estas dos últimas pruebas. Hubo 14 personas que lo lograron; sin embargo, los autores consideraron su criterio de selección demasiado restrictivo, con lo que posteriormente cambiaron las condiciones, de forma que bastaría con haber alcanzado una precisión del 80% en sólo una de estas dos últimas pruebas (y no necesariamente en las dos). Con este nuevo rasero identificaron a 15 «magos» adicionales.

La segunda fase de la investigación consistió en volver a presentar los vídeos a cada uno de los «magos» para que éstos manifestasen en voz alta las ideas, razonamientos y pensamientos que se les fueran ocurriendo. El objetivo era conocer cómo estas personas elaboran sus juicios de veracidad y en qué se basan para tomar sus decisiones. Siguiendo este método de trabajo, los investigadores llegaron a las siguientes conclusiones: (a) quienes alcanzaron una precisión mayor indicaron haber utilizado indicios no verbales, o indicios verbales y no verbales conjuntamente; (b) una estrategia comúnmente utilizada por los «magos» fue la comparación entre la conducta verbal y no verbal del declarante; (c) sólo 10 de los 29 «magos» eran mujeres, pero los autores indican que la proporción inicial de mujeres en los grupos investigados era todavía menor, de modo que entre los «magos» había más mujeres de lo esperado por azar; (d) la mayoría de los magos no eran ni muy jóvenes ni muy mayores; (e) más curioso es que, según los autores, muchos de los expertos identificados habían tenido algún tipo de experiencia poco usual en la infancia, como ser hijos de alcohólicos o no haber hablado inglés hasta haber superado el periodo de escolaridad elemental; (f) finalmente, según O'Sullivan y Ekman, todos los «magos» mostraron una motivación muy elevada para detectar el engaño, así como un gran entusiasmo por participar en la investigación.

La idea de O'Sullivan y Ekman (2004) de identificar a los verdaderos especialistas en la detección del engaño y posteriormente estudiar sus características y estrategias podría permitir el desarrollo de protocolos útiles para personas menos dotadas que, por las características de su trabajo, tengan que decidir si alguien miente o dice la verdad. El problema es que, en opinión de Bond y Uysal (2007), es muy probable que tales «magos» de la detección no existan, y que las personas identificadas como tales por O'Sullivan y Ekman (2004) lo hayan sido sólo por azar.

Según Bond y Uysal (2007), la consideración de O'Sullivan y Ekman (2004) de que una persona, por azar, tiene un 50% de probabilidades de juzgar correctamente cada una de las declaraciones (modelo al que estos autores se refieren como «lanzar una moneda al

aire») no es una premisa válida. Como prueba de ello, presentan los resultados de distintos trabajos publicados por el grupo de Ekman en los que muestras de estudiantes habían tenido que evaluar los mismos vídeos que luego se emplearían en el *Wizards Project*. La precisión media obtenida por 353, 113 y 464 estudiantes en la detección de mentiras sobre opiniones, el robo y las emociones fue del 56%, del 60% y del 50%, respectivamente. De modo que en dos de estas pruebas los participantes no habrían contestado como si hubieran «lanzando una moneda al aire», pues su precisión está por encima del 50% esperado por azar. La precisión de los magos tenía que haberse comparado con los niveles de precisión del 56% y del 60% obtenidos por estudiantes legos, y no con el nivel de azar. En consecuencia, Bond y Uysal (2007) plantean como alternativa el denominado «modelo basado en la investigación», que permite calcular cuántos individuos de los 12.000 que participaron en el proyecto podrían haber alcanzado los niveles de precisión requeridos para ser considerados «magos» sólo por azar. Para ello tienen en cuenta tanto la cantidad de participantes como los resultados de los trabajos previos. Los cálculos de Bond y Uysal (2007) indican que el número de «magos» identificados sólo por azar debería haber sido de 70. Lo sorprendente es, en consecuencia, que O'Sullivan y Ekman sólo identificaran a 29.

Bond y Uysal (2007) también plantean una serie de críticas metodológicas al *Wizards Project*, tales como haber permitido que fueran los propios participantes los que evaluaran sus propios resultados, o la falta de una definición *a priori* de lo que es un «mago». Bond y Uysal concluyen que O'Sullivan y Ekman (2004) no sólo no presentan evidencia consistente de la existencia de estos magos de la detección, sino que tampoco han especificado un procedimiento estandarizado de diagnóstico de tales individuos (véase Bond y Uysal, 2007).

Ante esta controversia, otro investigador, Gary Bond, publicó en 2007 un trabajo mediante el que trató de arrojar algo de luz sobre el asunto. En este estudio, un grupo de 112 agentes de la ley (policías locales, *sheriffs*, agentes de la CIA y del FBI, agentes de fronteras y agentes del *Bureau of Indian Affairs*) y otro de 122 estudiantes realizaron una primera tarea de detección del engaño en la que tenían que evaluar la credibilidad de 32 declaraciones (16 verdades y 16 mentiras). En ocho de ellas el declarante mentía o decía la verdad sobre la comisión de un robo, en otras ocho participaba en una entrevista de trabajo, en otras ocho expresaba los sentimientos hacia otra persona y en las ocho restantes cada entrevistado mentía o decía la verdad sobre si había visto un vídeo (véase Bond, 2007, para una descripción detallada de la elaboración y contenido de los vídeos estímulo). De los

112 participantes, 11 obtuvieron una precisión del 80% o superior (es decir, acertaron al menos 26 de las 32 declaraciones), siendo seleccionados como posibles expertos y sometidos a una segunda tarea. Ninguno de los estudiantes alcanzó estos niveles de precisión. Para el segundo experimento se pudo contar únicamente con ocho de los 11 profesionales seleccionados, a los que se añadió un grupo control de ocho estudiantes. En este caso se utilizaron 10 declaraciones nuevas (cinco verdades y cinco mentiras) sobre los mismos temas del experimento anterior. De los ocho supuestos expertos, sólo dos mostraron una precisión superior al 80%.

Al final del estudio, el autor se pregunta si los resultados podrían explicarse por azar, y ofrece un cálculo de la probabilidad de acertar por azar 39 del total de los 42 ítems de los dos experimentos (nivel de precisión obtenido por los dos expertos finales identificados). Dicha probabilidad, considerando el total de participantes y las tasas medias obtenidas, sería según G. Bond (2007) de 0,000002. A renglón seguido, el propio investigador presenta las limitaciones principales de este trabajo, entre las que destaca el hecho de que sólo 10 personas generaran todas las declaraciones, con lo que, necesariamente, una misma persona tenía que aparecer en más de una ocasión, y posiblemente en los dos experimentos. Además, G. Bond plantea la necesidad de pasar por más pruebas para ser calificado como experto. Este es un detalle importante, ya que si de los ocho expertos que pasaron la segunda prueba sólo dos obtuvieron puntuaciones superiores al 80%, ¿qué podría haber pasado en una tercera prueba? ¿Cómo decidir cuántas pruebas son necesarias para poder concluir con seguridad de que nos hallamos ante un experto?

Hasta el momento, pues, la pregunta que encabezaba este epígrafe, referida a la existencia entre los profesionales del ámbito legal de algunos que, de forma sistemática, hayan mostrado una precisión superior, no tiene una respuesta clara. Quizás sería necesario, en primer lugar, que los distintos investigadores se pusieran de acuerdo y desarrollasen un protocolo estandarizado para identificar a los verdaderos expertos. Una vez dado este primer paso, habría que esperar a que hubiese evidencia suficiente para establecer una conclusión con cierta solidez.

4.2. Precisión al juzgar la veracidad de los niños

Aunque lo más común es que la precisión de los policías se estudie con grupos de adultos como emisores, en algunos casos se han utilizado niños. La razón es obvia: en años recientes ha aumentado la

preocupación por las denuncias de abuso sexual infantil, y algunos profesionales están preocupados ante la posibilidad de que los niños, ya sea por iniciativa propia o a instancias de un adulto, formulen una acusación falsa (véase Masip y Garrido, 2007). ¿Es la precisión de policías y otros profesionales tan pobre al juzgar la veracidad de los niños como al juzgar la de los adultos? Al contrario de lo que suele pensarse, las mentiras infantiles no son siempre fáciles de detectar (véase Masip y Garrido, 2007, para una revisión). ¿Hay algunos profesionales con una capacidad especial para ello? Los trabajos de Leach et al. (2004) y de Vrij, Akehurst et al. (2006) examinan esta cuestión.

Leach, Talwar, Lee, Bala y Lindsay (2004) presentan tres experimentos en los que policías, agentes de aduana y estudiantes tuvieron que juzgar la credibilidad de un conjunto de declaraciones efectuadas por niños de entre 3 y 11 años. Se empleó un paradigma de resistencia a la tentación en el cual cada niño interactuaba con una experimentadora. En un momento determinado, ésta abandonaba la habitación e indicaba al niño que no debía mirar un juguete determinado. Algunos de los niños obedecían la orden de la experimentadora, y otros no. Cuando ésta regresaba a la habitación, hacía tres preguntas a los niños: (a) «¿mientras yo no estaba, moviste tu cabeza hacia un lado?», (b) «¿te moviste alrededor de la silla?», y (c) «¿curioseaste para ver cuál era el juguete?» Algunos niños mintieron y otros dijeron la verdad sobre si habían mirado el juguete o no. Las declaraciones de los niños se grabaron en vídeo y se mostraron a los detectores. En esta condición, la precisión de los tres grupos (policías, agentes de aduana y estudiantes) fue, respectivamente, del 44%, del 49% y del 51%. Las diferencias entre agentes de aduana y estudiantes no fueron significativas y sus puntuaciones no se diferenciaron significativamente del 50% esperado por azar. La precisión de los policías fue significativamente inferior a la de los otros dos grupos, y además estuvo por debajo de lo esperado por azar.

En el segundo estudio, Leach et al. (2004) utilizaron un paradigma similar al anterior, pero se introdujo una modificación: antes de que los niños contestasen a las preguntas de la experimentadora, se discutió con ellos la moralidad de mentir o decir la verdad. Además, se incluyó una muestra de la conducta del niño con la experimentadora en la que ésta le hacía preguntas que no tenían nada que ver con la posible transgresión. En otras palabras, cada declaración constaba de dos partes: (a) una muestra de la conducta previa del niño con la experimentadora, y (b) las respuestas del niño a las preguntas sobre la posible transgresión. Con esta modificación, las tasas de precisión fueron del 64%, del 62% y del 63% para policías, agentes de aduana y estudiantes, respectivamente. Todas las puntuaciones estu-

vieron por encima del nivel del azar, y las diferencias entre los tres grupos no fueron significativas. Los resultados de este segundo estudio pueden reflejar la utilidad de exponer a los detectores a muestras previas de la conducta veraz de los emisores. Alternativamente, es posible que tras la discusión sobre la moralidad de mentir o decir la verdad los niños emitieran mentiras más transparentes.

En el tercer estudio, el contexto experimental y el formato de las declaraciones fue similar, pero en lugar de discutir con los niños la moralidad de mentir o decir la verdad se les hizo prometer, antes de contestar a las preguntas de la experimentadora, que dirían la verdad. Las tasas de precisión fueron del 59%, del 57% y del 61% para policías, agentes de aduana y estudiantes, respectivamente. De nuevo, las diferencias entre los tres grupos no fueron significativas y todos acertaron por encima del azar.

Vrij, Akehurst, et al. (2006), por su parte, realizan un interesante estudio en el que comparan la precisión de policías, trabajadores sociales, profesores y estudiantes al juzgar las declaraciones de niños (de 5 y 6 años), adolescentes (de 14 y 15 años) y adultos. Los investigadores no encontraron diferencias significativas en función del grupo de pertenencia de los observadores. La precisión estuvo en torno al 60% en todos los casos. Respecto a los emisores, sólo se encontró una precisión superior cuando se juzgaron declaraciones verdaderas ofrecidas por adolescentes (67%).

En definitiva, tanto el trabajo de Leach et al. (2004) como el de Vrij, Akehurst, et al. (2006) parecen indicar que la precisión de profesionales y no profesionales al juzgar la veracidad de declaraciones hechas por niños es similar a la obtenida al juzgar declaraciones de adultos.

4.3. Modalidad de presentación de las declaraciones

Aunque en la mayoría de los casos se ha utilizado un formato audiovisual, algunos estudios han tratado de establecer si hay diferencias en precisión en función de si los policías sólo ven o sólo escuchan las declaraciones que tienen que juzgar. En la condición visual (imagen sin sonido) del estudio de Kassin et al. (2005), un grupo de 28 policías y otro de 29 estudiantes tuvieron que juzgar la veracidad de 10 declaraciones (cinco verdaderas y cinco falsas) sobre distintos delitos. La precisión de los estudiantes fue sólo del 53,4% y la de los policías del 42,1%, inferior a la de aquellos y al 50% esperado por azar. En la condición auditiva (sonido sin imagen), 29 estudiantes y 32 policías juzgaron la veracidad de las mismas 10 declaraciones.

Las tasas de precisión fueron del 64,1% para los estudiantes y del 54,5% para los policías. De nuevo, la precisión de los estudiantes fue superior a la de los policías, que se situó en torno al nivel del azar. Además de mostrar la superioridad de los estudiantes sobre los policías, este estudio mostró que, para ambos grupos, la precisión fue mayor en la condición auditiva que en la condición visual.

En otro estudio posterior, Mann et al. (2008) compararon la precisión de los policías en tres condiciones diferentes: visual (sólo imagen), auditiva (sólo sonido), y audiovisual. En todos los casos se utilizaron los mismos *clips* (siete verdades y siete mentiras), en cada uno de los cuales aparecía una persona haciendo una declaración. Los 37 policías sometidos a la condición visual obtuvieron una precisión global del 43% (34% para las verdades y 53% para las mentiras), los 31 asignados a la condición auditiva una precisión global del 66% (63% para las verdades y 69% para las mentiras) y los 35 de la condición audiovisual una precisión global del 65% (60% para las verdades y 70% para las mentiras). Los autores indican que la precisión fue mayor en la condición auditiva y en la condición audiovisual que en la condición visual. No se hallaron diferencias significativas entre las condiciones auditiva y audiovisual.

En definitiva, la precisión de los juicios de veracidad de los policías mejora cuando, además de ver la declaración, también pueden escuchar lo que dice el emisor. Estos hallazgos van en la misma línea que los de estudios hechos con muestras no profesionales, y el valor de las tasas de precisión es similar. Así, por ejemplo, en el meta-análisis de Kalbfleisch (1985) se observó una precisión del 51% para la modalidad visual, del 58% para la modalidad auditiva y del 57% para la modalidad audiovisual. En el reciente meta-análisis de Bond y DePaulo (2006), la precisión para estas tres modalidades fue del 52%, del 62% y del 56%, respectivamente.

4.4. Conclusiones

Las conclusiones que se pueden derivar de los trabajos revisados en los epígrafes anteriores son las siguientes: (a) no está claro que existan profesionales especialmente capaces de diferenciar entre verdades y mentiras; (b) las tasas de precisión de los policías son similares a las de observadores legos, con independencia de si juzgan la veracidad de declaraciones de adultos o de niños; y (c) en general, parece que la precisión de los policías, como sucede con la de observadores legos, es mayor cuando las declaraciones se presentan en formato auditivo o audiovisual que cuando se presentan en formato visual.

5. El efecto del sesgo del investigador: papel de los años de experiencia y del entrenamiento

Diversos investigadores han tratado de explicar la pobre precisión de los policías, así como su tendencia a juzgar las declaraciones como falsas. Meissner y Kassin (2002) hicieron un pequeño meta-análisis en el cual se muestra que ni el hecho de ser policía ni el haber recibido entrenamiento para detectar mentiras incrementan la precisión de los juicios. Más bien, lo que hacen estos dos factores es sesgar los juicios en dirección a la mentira. En otras palabras, tanto los profesionales de los cuerpos de seguridad como las personas entrenadas para detectar mentiras muestran una marcada tendencia a juzgar las declaraciones como falsas. Meissner y Kassin (2002) etiquetan este fenómeno como *investigator bias effect* (efecto del sesgo del investigador). ¿Qué puede explicar este fenómeno?

5.1. *Experiencia policial y sospecha*

Es posible que la experiencia policial genere la sospecha de que otros mienten, y que esta sospecha dé lugar, en último término, a hacer juicios de mentira. En muchas interacciones de la policía con la población, en especial al tratar con posibles sospechosos o cómplices, el policía a menudo desconfía y cuestiona lo dicho por el entrevistado. Esto puede ser adaptativo, ya que los policías no pueden creer ciegamente todo lo que les dicen; sin embargo, la exposición repetida a este tipo de situaciones podría dar lugar a una cronificación de la sospecha, de modo que los policías acaben creyendo que las personas *normalmente* mienten. Esto encaja con la noción de Sospecha Comunicacional Generalizada (*Generalized Communicative Suspicion* o GCS) de Levine y McCornack (1991), que los autores definen como una «predisposición a creer que los mensajes generados por los demás son falsos» (p. 328). Es plausible la idea de que los policías desarrollen dicha predisposición a lo largo de su carrera profesional, y que esto les lleve a entrar en la situación de evaluación de la credibilidad con la creencia previa de que el testigo va a mentir. En estas condiciones, los policías podrían procesar la información emitida por el emisor de una forma sesgada para confirmar que éste está mintiendo.

Hay evidencia que apunta en esta dirección. Recordemos que los policías del experimento de Garrido et al. (2004) mostraron una mayor tendencia que los no policías a hacer juicios de mentira. Pues bien, ambos grupos tuvieron que señalar los indicadores en que se habían basado para hacer sus juicios (véase Masip, Garrido, Herrero, et

al., 2006). El análisis de estos datos arrojó resultados sorprendentes. En primer lugar, ni los indicadores (por ejemplo, la mirada) enumerados por ninguno de los grupos ni la dirección de tales indicadores (por ejemplo, el emisor sostiene la mirada vs. aparta la mirada) se relacionaban significativamente con la veracidad de las declaraciones. Sin embargo, casi sin excepción estos indicadores (o su dirección) se relacionaban con los juicios (por ejemplo, quienes decían que el emisor apartaba la mirada juzgaban que mentía y quienes decían que la sostenía juzgaban que decía la verdad). Más interesantes todavía son las diferencias entre policías y no policías. Los primeros (que, recordémoslo, hicieron más juicios de mentira que el otro grupo) tendieron a mencionar la dirección de los indicadores asociada a juicios *de mentira* con más frecuencia que la esperada, mientras que los no policías tendieron a mencionar la dirección de los indicadores asociada a juicios *de verdad* con más frecuencia de la esperada. Así, por ejemplo, los policías percibieron que la declarante apartaba su mirada al hablar, y como asociaban esta conducta al engaño, resolvieron que mentía. Por el contrario, los no policías tendieron relativamente a percibir que la declarante sostenía bastante la mirada, y como asociaban esta conducta a la verdad, resolvieron más que el otro grupo que decía la verdad. Nótese que ambos grupos veían las mismas declaraciones, pero percibían cosas opuestas. Es probable que los policías partieran de la idea previa de que la testigo mentía y que siguieran una estrategia confirmatoria consistente en buscar indicadores que creían asociados al engaño, desestimar indicadores que creían asociados a la verdad, e interpretar conductas ambiguas (como la ausencia de conducta nerviosa; véase Masip, Garrido, Herrero, et al., 2006) como indicadores de falsedad y no como indicadores de verdad. Detrás de esta creencia *a priori* de que la testigo iba a mentir podría estar una elevada GCS (véanse Masip, 2002, y Masip, Garrido, Herrero, et al., 2006, para más detalles sobre este estudio).

Decidimos, pues, comprobar si los policías mostraban una mayor GCS que los no policías. Para ello adaptamos al español el cuestionario de GCS diseñado por Levine y McCornack (1991) y lo administramos a 88 policías con una experiencia profesional media de 18 años, a 89 policías noveles (llevaban en el cuerpo menos de dos años) y a 152 estudiantes universitarios. Hallamos que la puntuación media de los policías veteranos en la escala de GCS ($M = 39,42$) fue significativamente superior a la de los policías noveles ($M = 33,13$) y a la de los estudiantes ($M = 35,57$). Sin embargo, estos dos grupos no difirieron significativamente entre sí (Masip, Alonso, Garrido y Antón, 2005). En definitiva, parece que los años en el cuerpo sí que incrementan la sospecha generalizada de los policías. No obstante, que-

daba abierta la cuestión de si esta sospecha estaba detrás del sesgo de los policías a hacer juicios de mentira. En primer lugar, tras llevar a cabo este estudio no pudimos acceder de nuevo a las muestras policiales para someterlas a una tarea de evaluación de la veracidad y examinar las correlaciones entre puntuaciones en GCS y juicios de mentira. En segundo lugar, muchos de los policías del estudio de Garrido et al. (2004) eran noveles, por lo que, presumiblemente, no tenían niveles muy altos de GCS; sin embargo, mostraron un claro sesgo de mendacidad.

Para resolver esta cuestión hicimos otro experimento. En esta ocasión, 72 policías veteranos (experiencia media de 22 años), 72 policías noveles y 72 estudiantes cumplieron el cuestionario de GCS y luego, al cabo de unos meses y como una tarea separada, juzgaron la credibilidad de 12 declaraciones (seis verdades y seis mentiras) presentadas en vídeo (Alonso, Masip, Garrido, Antón y Herrero, 2005; Masip, Alonso, Garrido y Barba, 2008, Estudio 1). Los resultados concernientes a la GCS replicaron los del estudio anterior: las puntuaciones medias de los policías veteranos ($M = 38,81$) fueron significativamente mayores que las de los policías noveles ($M = 33,90$) y marginalmente superiores que las de los estudiantes ($M = 34,94$), y estos dos últimos grupos no difirieron significativamente entre sí. La precisión fue baja en todos los grupos (véase la Tabla 1). Tanto los policías veteranos ($M = 56,6\%$) como los policías nuevos ($M = 53,7\%$) hicieron significativamente más juicios de mentira que los estudiantes ($M = 46,5\%$). Las diferencias entre ambos grupos policiales no fueron significativas. De hecho, los dos grupos de la policía estaban sesgados hacia la mentira y el grupo de estudiantes hacia la verdad. Por otro lado, los policías veteranos mostraron más confianza en sus juicios ($M = 5,0$ en una escala de 1 a 7) que el grupo de policías nuevos ($M = 4,5$) y que el de estudiantes ($M = 4,2$), y estos dos últimos grupos no difirieron entre sí.

Pero la pregunta clave en este estudio era si la GCS explica la tendencia de los policías a hacer juicios de mentira. La respuesta es negativa: ni en ninguno de los tres grupos considerados separadamente, ni para todos los grupos tomados en su conjunto la correlación entre las puntuaciones en la escala de GCS y el porcentaje de juicios de mentira fue significativa; y no lo fue ni al juzgar las declaraciones verdaderas, ni al juzgar las falsas, ni al considerar todas las declaraciones conjuntamente. En otras palabras: si bien la experiencia policial incrementa la GCS, y si bien los juicios de mentira son más frecuentes entre los policías veteranos que entre los no policías, la GCS no está detrás de la tendencia de los policías veteranos a hacer juicios de mentira.

Un dato llamativo de este estudio es que los policías noveles puntuaron tan bajo como los estudiantes en GCS y en confianza; sin embargo, hicieron tantos juicios de mentira como los policías veteranos. Una hipótesis que podría explicar estos resultados es que, en situaciones percibidas como policialmente relevantes (tales como una tarea de evaluación de la veracidad), los policías noveles intentan adoptar la conducta prototípica de un curtido agente de la ley (suspica, poco crédulo...), haciendo muchos juicios de mentira, pero sin ser internamente (a diferencia de sus compañeros veteranos) «curtidos agentes de la ley», por lo que hacen tales juicios sin mucha seguridad y no muestran grandes niveles de GCS. Con el fin de contrastar esta hipótesis, Masip et al. (2008, Estudio 2) construyeron un cuestionario similar al de GCS, pero en el que las preguntas se referían específicamente a situaciones policiales (y no a situaciones de la vida diaria, que es a las que se refiere el cuestionario original). Si la hipótesis es cierta, entonces los policías noveles, al percibir la relevancia policial de las preguntas del cuestionario, responderán mostrando niveles de sospecha tan elevados como los policías veteranos. Se administró el nuevo cuestionario a 74 policías noveles y a 107 policías veteranos (con una experiencia media de 24 años) y los resultados confirmaron la hipótesis: las puntuaciones de los policías noveles fueron elevadas ($M = 53,81$) y virtualmente idénticas que las de los policías veteranos ($M = 53,06$; Masip et al., 2008, Experimento 2).

5.2. *Entrenamiento para detectar mentiras*

Meissner y Kassin (2002) también encontraron que el entrenamiento para detectar mentiras no incrementa la precisión de la detección, sino la tendencia a hacer juicios de mentira (véase también el meta-análisis sobre el entrenamiento de Frank y Feeley, 2003). Aunque la cantidad de policías que recibe este tipo de entrenamiento es escasa (Bull, 1999; Colwell, Miller, Lyons y Miller, 2006; Frank y Feeley, 2003), lo recibe una mayor proporción de policías que de miembros de la población general. Así, pues, la tendencia del entrenamiento a sesgar los juicios puede estar contribuyendo al sesgo de mendacidad de los policías. Es importante saber por qué entrenar a las personas tiene este efecto, ya que entonces se podrían buscar vías para eliminarlo, diseñando programas de entrenamiento realmente eficaces.

Una razón que podría explicar el sesgo hacia la mentira de los juicios debido al entrenamiento podría ser la propia naturaleza de los programas que se ofrecen. El objetivo de los mismos es la detección

de *mentiras*, y dirigen el foco atencional de los entrenandos exclusivamente hacia la búsqueda de indicios de engaño. Como resultado, tras el entrenamiento los participantes van a buscar activamente tales indicios, y van a ser especialmente sensibles a su presencia. No obstante, los indicios del engaño son de naturaleza probabilística; es decir, aunque pueden aparecer más a menudo cuando una persona miente, esto no significa que no puedan estar presentes en algunos casos en los que se está diciendo la verdad. La sensibilidad del individuo entrenado a la aparición de los indicios de la mentira y la búsqueda activa de los mismos podría llevarle a encontrarlos no solamente en declaraciones falsas, sino también en declaraciones verdaderas. Al percibir dichos indicios, el detector podría rápidamente inferir que el emisor está mintiendo, sin preocuparse de buscar otros elementos de contraste que pudieran sugerirle que dice la verdad. Esta estrategia resultaría en un aumento indiscriminado de los juicios de mentira.

Si todo este razonamiento es correcto, entonces un programa de entrenamiento centrado sobre la detección de la verdad (en vez de la mentira) y en el que se enfatizaran los indicadores conductuales de la verdad (en vez de los de la mentira) debería producir un aumento de juicios de verdad. Llevamos a cabo dos estudios para examinar esta cuestión (Masip, Alonso, Garrido y Herrero, en prensa).

En ambos casos utilizamos un diseño mixto en el cual tres grupos de participantes hacían dos tareas consistentes en juzgar la veracidad de diez declaraciones (cinco verdaderas y cinco falsas) filmadas en vídeo. La primera de estas tareas la hacían al principio de la sesión y la segunda al final. Los tres grupos diferían en lo que hacían en medio. El grupo de entrenamiento tradicional se sometía a un programa de entrenamiento para detectar mentiras en el cual se enseñaba a identificar supuestos indicadores de la mentira. El grupo de entrenamiento alternativo se sometía a un programa de entrenamiento para detectar verdades en el cual se enseñaba a identificar supuestos indicadores de la verdad. El grupo control no fue sometido a ningún entrenamiento, sino que recibió una clase sobre un tema que no guardaba ninguna relación con el engaño ni su detección. En ambos estudios encontramos que, tras el entrenamiento, el porcentaje de juicios de mentira aumentaba en el grupo tradicional, se reducía en el grupo alternativo y no variaba significativamente en el grupo control. De hecho, los análisis hechos desde la teoría de la detección de señales indicaron que antes del entrenamiento ningún grupo mostraba un sesgo significativo a hacer juicios de verdad ni de mentira, pero que tras dicho entrenamiento el grupo tradicional presentaba un sesgo a hacer juicios de mentira y el grupo alternativo presentaba un

sesgo a hacer juicios de verdad. En cuanto a la confianza, se incrementó en ambos grupos de entrenamiento, pero no en el grupo control. El Estudio 1 mostró, además, que el entrenamiento incrementaba la percepción de los indicios en los que se entrenaba, y que la percepción de los mismos se relacionaba con los juicios. En definitiva, los programas tradicionales de entrenamiento, que centran la atención de los entrenandos hacia las claves de la mentira, tienen el efecto de sesgar las respuestas más que de incrementar la precisión, y ello se debe a que el entrenando busca activamente tales claves, percibiéndolas incluso en las declaraciones verdaderas (Masip et al., en prensa).

De estos dos estudios se desprenden algunas recomendaciones para detectar programas efectivos de entrenamiento. Así, en lugar de centrarse sólo en las claves del engaño, tales programas deberían centrarse también en las claves de la verdad. En este caso, la tendencia a buscar indicios del engaño se vería compensada por la tendencia a buscar indicios de la verdad y, como resultado, los entrenandos serían más receptivos a indicadores conductuales de cualquier signo. Sin embargo, el diseño de cualquier programa de entrenamiento se enfrenta a un problema más básico: no hay claves universales de la verdad o la mentira, claves útiles para cualquier persona en cualquier situación. De hecho, incluso los indicios más discriminativos tienen un valor diagnóstico nulo (véanse los meta-análisis de DePaulo et al., 2003, y de Sporer y Schwandt, 2006, 2007). Una posibilidad consiste en ajustar el programa de entrenamiento a las características de la situación, del posible mentiroso y del tipo de mentira. Los entrenandos deben saber bajo qué condiciones discrimina cada indicador, y bajo cuáles no lo hace. Las iniciativas que se presentan en el siguiente apartado van en esta dirección, pero también dan un paso más: se enseña a los entrenandos a crear las condiciones bajo las cuales ciertas claves discriminan entre veraces y mentirosos.

5.3. Cómo hacer de los policías mejores detectores de mentiras

Hemos visto en este trabajo que la capacidad de los policías para discriminar entre verdades y mentiras es muy limitada y que tales policías tienden a hacer juicios de mentira. Hemos buscado explicaciones de tales fenómenos, observando que la experiencia policial hace a los policías veteranos más suspicaces pero que, no obstante, esta suspicacia (o sospecha generalizada) no explica sus juicios de menti-

ra. También hemos visto que el entrenamiento para detectar mentiras que los policías en ocasiones reciben puede sesgar sus respuestas, y se han ofrecido recomendaciones sobre cómo diseñar un buen programa de entrenamiento: éste debe enseñar no sólo indicios del engaño, sino también de la verdad, y debe indicar cuáles son las condiciones bajo las cuales discrimina cada indicio.

De hecho, un buen programa de entrenamiento podría, en último término, resolver el problema de la baja precisión de los policías. En este sentido, los trabajos de Hartwig, Granhag, Strömwall y Kronkvist (2006; véase también Granhag, Strömwall y Hartwig, 2007, para una visión de conjunto del procedimiento, y Granhag y Hartwig, 2008, para el trasfondo teórico) y Vrij, Mann, Fisher, Leal, Milne y Bull (2008) son particularmente relevantes, ya que sugieren que tales profesionales deben crear las situaciones particulares en las cuales van a aparecer ciertos indicios de engaño en los mentirosos, e indican cómo hacerlo. Hartwig et al. (2006) se centran específicamente en las circunstancias en las cuales los mentirosos muestran inconsistencias. Según los autores, durante el interrogatorio la policía no debería indicar al sospechoso cuáles son las pruebas existentes contra él para que confiese. Por el contrario, debería mantener al sospechoso ignorante de la evidencia que le incrimina. Esto haría que las declaraciones de dicho sospechoso (por ejemplo, que el día del delito viajó con su automóvil fuera de la ciudad) contradijeran cosas que la policía ya supiera (por ejemplo, que su automóvil fue visto el día del delito cerca del lugar de los hechos). En otras palabras: si el sospechoso ignora lo que la policía sabe, es posible que la mentira que invente se contradiga con tales hechos.

Hartwig et al. (2006) entrenaron a un grupo de policías a no develar la evidencia hasta el final, y compararon su ejecución con la de un grupo de policías no entrenados. Entre otros resultados, hallaron lo siguiente. Primero, los sospechosos mentirosos mostraron más inconsistencias al ser interrogados por los policías entrenados que al ser interrogados por los policías no entrenados (no hubo diferencias en los sospechosos veraces, cuyas declaraciones fueron siempre bastante consistentes con los hechos). Segundo, a diferencia del grupo no entrenado, los policías entrenados hicieron más juicios de mentira cuantas más inconsistencias mostraron los entrevistados. Tercero, los policías entrenados alcanzaron una precisión ($M = 85,4\%$) mayor que la de los policías no entrenados ($M = 56,1\%$). Estos resultados son prometedores, ya que muestran que, bajo ciertas circunstancias, el entrenamiento (pero no el clásico tipo de entrenamiento que se ha empleado habitualmente) puede ayudar a los profesionales a detectar mentiras.

El trabajo de Vrij et al. (2008) va en la misma línea. Los autores nos recuerdan que, por lo general, mentir es una labor cognitivamente más compleja que simplemente decir la verdad, y argumentan que si se incrementa artificialmente la carga cognitiva de los declarantes, entonces quienes mientan se van a ver más perjudicados que quienes digan la verdad, mostrando indicios conductuales de sobrecarga cognitiva. Un modo de incrementar la carga cognitiva de los hablantes consiste en pedirles que describan los hechos en orden cronológico inverso. Al hacerlo, Vrij et al. encontraron que las diferencias conductuales entre mentirosos y veraces aumentaron, y cuando estas declaraciones «inversas» se presentaron a una muestra de policías para que evaluaran su veracidad, su precisión ($M = 58\%$) fue superior que la de otra muestra de policías que juzgó las declaraciones en su orden cronológico natural ($M = 46\%$).

La observación de la Tabla 2 invita a ofrecer algunas sugerencias adicionales. Así, también es necesario que los policías hagan sus juicios de veracidad de manera sistemática, sin dejarse llevar por sus intuiciones o corazonadas, que sepan que el canal verbal es más revelador que la comunicación no verbal (Vrij, 2008) y que tengan en cuenta no sólo la influencia de la situación sobre la conducta, sino también la influencia de la personalidad: es posible que personas distintas reaccionen de manera distinta en las mismas situaciones. Esto implica que los indicios del engaño no son necesariamente los mismos para todas las personas.

En definitiva, la reciente investigación psicológica muestra vías prometedoras para facilitar la ardua tarea de discriminar entre verdades y mentiras a la que muchos policías se enfrentan diariamente. Sólo cabe esperar seguir avanzando en esta dirección, y que tales avances lleguen a sus destinatarios últimos (los policías) y éstos los aprovechen.

6. Referencias

- AAMODT, M. G. y CUSTER, H. (2006). Who can best catch a liar? A meta-analysis of individual differences in detecting deception. *Forensic Examiner*, 15, 6-11.
- ABRUNHOSA, R., SOEIRO, C. y VENTURA, J. P. (2006). Psicología de la policía. En E. GARRIDO, J. MASIP y C. HERRERO (Eds.), *Psicología jurídica* (pp. 111-129). Madrid: Pearson.
- AKEHURST, L., BULL, R., VRIJ, A. y KÖHNKEN, G. (2004). The effects of training professional groups and lay persons to use criteria-based content analysis to detect deception. *Applied Cognitive Psychology*, 18, 877-891.

- ALONSO, H., MASIP, J., GARRIDO, E., ANTÓN, C. y HERRERO, C. (2005). *Sesgo de veracidad, sesgo de mendacidad y GCS. Juicios de credibilidad y sospecha en policías y no-policías al evaluar la veracidad de las declaraciones de testigos*. Comunicación presentada en el IX Congreso Nacional de Psicología Social, La Coruña, España.
- ÁLVAREZ, M. (2000, Julio). *Psicología y seguridad pública*. Trabajo presentado en el I Congreso Hispano Alemán de Psicología Jurídica, Pamplona, Navarra.
- BOND, C. F. y DEPAULO, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review*, 10, 214-234.
- BOND, C. F. y UYSAL, A. (2007). On lie detection «wizards». *Law and Human Behavior*, 31, 109-115.
- BOND, G. D. (2007). Deception detection expertise. *Law and Human Behavior*, 32, 339-351.
- BOTELLA, J. y GAMBARA, H. (2002). *Qué es el meta-análisis*. Madrid: Biblioteca Nueva.
- BOTELLA, J. y GAMBARA, H. (2006). El meta-análisis: Una metodología de nuestro tiempo. *Infocop Online*. Encontrado el 12 de Marzo de 2009, en http://www.infocop.es/view_article.asp?id=843.
- BULL, R. (1999). *Police investigative interviewing*. En A. MEMON y R. BULL (Eds.), *Handbook of the psychology of interviewing* (pp. 279-292). Chichester: Wiley.
- CANTÓN, J. (1997). Evaluación del abuso sexual infantil y tratamiento de las víctimas. En J. CANTÓN, y M. R. CORTÉS (Eds.), *Malos tratos y abuso sexual infantil* (pp. 284-362). Madrid: Siglo XXI.
- CLEMENTE, M. (1995). *Fundamentos de la psicología jurídica*. Madrid: Pirámide.
- COLWELL, L., MILLER, H., LYONS, P. y MILLER, R. (2006). The training of law enforcement officers in detecting deception: A survey of current practices and suggestions for improving accuracy. *Police Quarterly*, 9, 275-290.
- DEPAULO, B. M. (1992). Nonverbal behavior and self-presentation. *Psychological Bulletin*, 111, 203-243.
- DEPAULO, B. M., CHARLTON, K., COOPER, H., LINDSAY, J. J. y MUHLENBRUCK, L. (1997). The accuracy-confidence correlation in the detection of deception. *Personality and Social Psychology Review*, 1, 346-357.
- DEPAULO, B. M., LINDSAY, J. J., MALONE, B. E., MUHLENBRUCK, L., CHARLTON, K. y COOPER, H. (2003). Cues to deception. *Psychological Bulletin*, 129, 74-118.
- DEPAULO, B. M. y PFEIFFER, R. L. (1986). On-the-job experience and skill at detecting deception. *Journal of Applied Social Psychology*, 16, 249-267.
- EKMAN, P. (1992). *Cómo detectar mentiras. Una guía práctica para utilizar en el trabajo, la política y la pareja* (L. WOLFSON, Trad.). Barcelona: Paidós. (Trabajo original publicado en 1985).
- EKMAN, P. y O'SULLIVAN, M. (1991). Who can catch a liar? *American Psychologist*, 46, 913-920.
- EKMAN, P., O'SULLIVAN, M. y FRANK, M. (1999). A few can catch a liar. *Psychological Science*, 10, 263-266.
- ELAAD, E. (2003). Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. *Applied Cognitive Psychology*, 17, 349-363.

- FEELEY, T. H., DETURCK, M. A. y YOUNG, M. J. (1995). Baseline familiarity in lie detection. *Communication Research Reports*, 12, 160-169.
- FERNÁNDEZ, M. (1990). *La profesión de policía*. Madrid: CIS.
- FERNÁNDEZ, M. (2003). La investigación policial del delito: Competencias de la policía en la investigación del delito y métodos policiales de investigación. En J. BURÓN (Ed.), *Psicología médico-forense. La investigación del delito* (pp. 285-347). Bilbao: Desclée de Brouwer.
- FRANK, M. G. y FEELEY, T. H. (2003). To catch a liar: Challenges for research in lie detection training. *Journal of Applied Communication Research*, 31, 58-75.
- GARRIDO, E. y MASIP, J. (1999). How good are police officers at spotting lies? *Forensic Update*, 58, 14-21.
- GARRIDO, E. y MASIP, J. y HERRERO, C. (2004). Police officers' credibility judgments: Accuracy and estimated ability. *International Journal of Psychology*, 39, 254-275.
- GRANHAG, P. A. y HARTWIG, M. (2008). A new theoretical perspective on deception detection: On the psychology of instrumental mind-reading. *Psychology, Crime, & Law*, 14, 189-200.
- GRANHAG, P. A. y STRÖMWALL, L. (1999). Repeated interrogations — Stretching the deception detection paradigm. *Expert Evidence*, 7, 163-174.
- GRANHAG, P. A. y STRÖMWALL, L. y HARTWIG, M. (2007). The SUE technique: The way to interview to detect deception. *Forensic Update*, 88, 25-29.
- HARTWIG, M., GRANHAG, P. A., STRÖMWALL, L. A. y KRONKVIST, O. (2006). Strategic use of evidence during police interviews: When training to detect deception works. *Law and Human Behavior*, 30, 603-619.
- HARTWIG, M., GRANHAG, P. A., STRÖMWALL, L. A. y VRIJ, A. (2004). Police officers' lie detection accuracy: Interrogating freely versus observing video. *Police Quarterly*, 7, 429-456.
- KALBFLEISCH, P. J. (1985). Accuracy in deception detection: A quantitative review. *Dissertation Abstracts International*, 45, 46112B. (UMI No. 8603433).
- KASSIN, S. M. y FONG, C. T. (1999). «I'm innocent!»: Effects of training on judgments of truth and deception in the interrogation room. *Law and Human Behavior*, 23, 499-516.
- KASSIN, S. M., LEO, R. A., MEISSNER, C. A., RICHMAN, K. D., COLWELL, L. H., LEACH, A. M. y LA FON, D. (2008). Police interviewing and interrogation: A self-report survey of police practices and beliefs. *Law and Human Behavior*, 31, 381-400.
- KASSIN, S. M., MEISSNER, C. y NORWICK, R. (2005). I'd know a false confession if I saw one: A comparative study of college students and police investigators. *Law and Human Behavior*, 29, 211-227.
- KÖHNKEN, G. (1987). Training police officers to detect deceptive eyewitness statements: Does it work? *Social Behaviour*, 2, 1-17.
- LAMB, M. E. (1994). The investigation of child sexual abuse: An interdisciplinary consensus statement. *Child Abuse & Neglect*, 18, 1021-1028.
- LEACH, A., TALWAR, V., LEE, K., BALA, N. y LINDSAY, R. C. L. (2004). «Intuitive» lie detection of children's deception by law enforcement officials and university students. *Law and Human Behavior*, 28, 661-685.

- LEVINE, T. R. y McCORNACK, S. A. (1991). The dark side of trust: Conceptualizing and measuring types of communicative suspicion. *Communication Quarterly*, 39, 325-339.
- LEVINE, T. R., PARK, H. S. y McCORNACK, S. A. (1999). Accuracy in detecting truths and lies: Documenting the «veracity effect». *Communication Monographs*, 66, 125-144.
- MANN, S. Y VRIJ, A. (2006). Police officers' judgements of veracity, tenseness, cognitive load and attempted behavioral control in real-life police interviews. *Psychology, Crime and Law*, 12, 307-319.
- MANN, S., VRIJ, A. y BULL, R. (2004). Detecting true lies: Police officers' ability to detect suspects' lies. *Journal of Applied Psychology*, 89, 137-149.
- MANN, S., VRIJ, A., FISHER, P. y ROBINSON, M. (2008). See no lies, hear no lies: Differences in discrimination accuracy and response bias when watching or listening to police suspect interviews. *Applied Cognitive Psychology*, 22, 1062-1071.
- MASIP, J. (2002). *La evaluación de la credibilidad del testimonio a partir de los índices conductuales en el contexto jurídico penal*. Tesis Doctoral. Universidad de Salamanca.
- MASIP, J. (2005). ¿Se pillá antes a un mentiroso que a un cojo? Sabiduría popular frente a conocimiento científico sobre la detección no-verbal del engaño. *Papeles del Psicólogo*, 26, 78-91.
- MASIP, J., ALONSO, H., GARRIDO, E., y ANTÓN, C. (2005). Generalized Communicative Suspicion (GCS) among police officers: Accounting for the investigator bias effect. *Journal of Applied Social Psychology*, 35(5), 1046-1066.
- MASIP, J., ALONSO, H., GARRIDO, E. y BARBA, A. (2008). *Police officers' suspicion and veracity judgments*. Comunicación presentada en la 18th Conference of the European association of Psychology and Law, Maastricht, Holanda.
- MASIP, J., ALONSO, H., GARRIDO, E. y HERRERO, C. (en prensa). Training to detect what? The biasing effects of training to detect deception on veracity judgments. *Applied Cognitive Psychology*.
- MASIP, J. Y GARRIDO, E. (2000). La evaluación de la credibilidad del testimonio en contextos judiciales a partir de indicadores conductuales. *Anuario de Psicología Jurídica*, 10, 93-131.
- MASIP, J. Y GARRIDO, E. (2007). *La evaluación del abuso sexual infantil. Análisis de la validez de las declaraciones del niño*. Alcalá de Guadaíra, Sevilla: Editorial MAD.
- MASIP, J. Y GARRIDO, E. y HERRERO, C. (2002). La detección del engaño sobre la base de sus correlatos conductuales: La precisión de los juicios. *Anuario de Psicología Jurídica*, 12, 37-55.
- MASIP, J. Y GARRIDO, E. y HERRERO, C. (2003). When did you conclude she was lying? The impact of the moment the decision about the sender's veracity is made and the sender's facial appearance on police officers' credibility judgments. *Journal of Credibility Assessment and Witness Psychology*, 4, 1-36.
- MASIP, J. Y GARRIDO, E. y HERRERO, C. (2006). Observers' decision moment in deception detection experiments: Its impact on judgment, accuracy, and confidence. *International Journal of Psychology*, 41, 304-319.

- MASIP, J. Y GARRIDO, E. y HERRERO, C. (2009). *Heuristic versus systematic processing of information in detecting deception: Questioning the truth bias*. Manuscrito enviado y en proceso de revisión.
- MASIP, J. Y GARRIDO, E. y HERRERO, C., ANTÓN, C. y ALONSO, H. (2006). Officers as lie detectors: Guilty before charged. In D. CHADEE y J. YOUNG (Eds.), *Current themes in social psychology* (pp. 187-205). Mona, Jamaica: The University of the West Indies Press.
- MEISSNER, C. A. y KASSIN, S. M. (2002). «He's guilty!»: Investigator bias in judgments of truth and deception. *Law and Human Behavior*, 26, 469-480.
- MILNE, R. y BULL, R. (1999). *Investigative interviewing: Psychology and practice*. Chichester: Wiley.
- O'SULLIVAN, M. y EKMAN, P. (2004). The wizards of deception detection. En P. A. GRANHAG, y L. A. STRÖMWALL (Eds.), *The detection of deception in forensic contexts*. Cambridge: Cambridge University Press.
- PORTER, S., WOODWORTH, M. y BIRT, A. R. (2000). Truth, lies, and videotape: An investigation of the ability of federal parole officers to detect deception. *Law and Human Behavior*, 24, 643-658.
- RICO, J. M. (1988). *Inseguridad ciudadana y policía*. Madrid: Tecnos.
- SPORER, S. L. y SCHWANDT, B. (2006). Paraverbal indicators of deception: A meta-analytic synthesis. *Applied Cognitive Psychology*, 20, 421-446.
- SPORER, S. L. y SCHWANDT, B. (2007). Moderators of nonverbal indicators of deception. *Psychology, Public Policy, and Law*, 13, 1-34.
- STIFF, J. B., KIM, H. J. y RAMESH, C. N. (1992). Truth biases and aroused suspicion in relational deception. *Communication Research*, 19, 326-345.
- VRIJ, A. (1992). Credibility judgments of detectives: The impact of nonverbal behavior, social skills, and physical characteristics on impression formation. *The Journal of Social Psychology*, 133, 601-610.
- VRIJ, A. (1994). The impact of information and setting on detection of deception by police detectives. *Journal of Nonverbal Behavior*, 18, 117-136.
- VRIJ, A. (2000). *Detecting lies and deceit. The psychology of lying and the implications for professional practice*. Chichester, Reino Unido: John Wiley and Sons.
- VRIJ, A. (2004). Why professionals fail to catch liars and how they can improve. *Legal and Criminological Psychology*, 9, 159-181.
- VRIJ, A. (2008). Nonverbal dominance versus verbal accuracy in lie detection. A plea to change police practice. *Criminal Justice and Behavior*, 35, 1323-1336.
- VRIJ, A., AKEHURST, L., BROWN, L. y MANN, S. (2006). Detecting lies in young children, adolescents and adults. *Applied Cognitive Psychology*, 20, 1225-1237.
- VRIJ, A. y HEAVEN, S. (1999). Vocal and verbal indicators of deception as a function of lie complexity. *Psychology, Crime, & Law*, 5, 203-215.
- VRIJ, A., FISHER, R., MANN, S. y LEAL, S. (2006). Detecting deception by manipulating cognitive load. *TRENDS in Cognitive Sciences*, 10, 141-142.
- VRIJ, A. y MANN, S. (2001a). Telling and detecting lies in a high-stake situation: The case of a convicted murderer. *Applied Cognitive Psychology*, 15, 187-203.

- VRIJ, A. y MANN, S. (2001b). Who killed my relative? Police officers' ability to detect real-life high-stake lies. *Psychology, Crime and Law*, 7, 119-132.
- VRIJ, A., MANN, S., FISHER, R., LEAL, S., MILNE, R. y BULL, R. (2008). Increasing cognitive load to facilitate lie detection: The benefit of recalling an event in reverse order. *Law and Human Behavior*, 32, 253-265.
- VRIJ, A., MANN, S., KRISTEN, S. y FISHER, R. (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior*, 31, 499-518.
- VRIJ, A., MANN, S., ROBBINS, E. y ROBINSON, M. (2006). Police officers ability to detect deception in high stakes situations and in repeated lie detection test. *Applied Cognitive Psychology*, 20, 741-755.
- VRIJ, A. y WINKEL, F. W. (1993). Objective and subjective indicators of deception. *Issues in Criminological and Legal Psychology*, 20, 51-57.
- ZUCKERMAN, M., DEPAULO, B. M. y ROSENTHAL, R. (1981). Verbal and nonverbal communication of deception. *Advances in Experimental Social Psychology*, 14, 1-59.