

**J-40402082-9**

F  
u  
n  
d  
a  
c  
i  
ó  
n  
  
A  
u  
l  
a  
  
V  
i  
r  
t  
u  
a  
l

# Aula Virtual



**Generando Conocimiento**

<http://www.aulavirtual.web.ve>



**ISSN: 2665-0398**

**Deposito Legal: LA2020000026**

**Vol. 7 Nº 15 Año 2026**

Periodicidad Continua



# REVISTA CIENTÍFICA AULA VIRTUAL

## Director Editor:

- Dra. Leidy Hernández PhD.
- Dr. Fernando Bárbara

## Consejo Asesor:

- MSc. Manuel Mujica
- MSc. Wilman Briceño
- Dra. Harizmar Izquierdo
- Dr. José Gregorio Sánchez

**Revista Científica Arbitrada de  
Fundación Aula Virtual**

Email: [revista@aulavirtual.web.ve](mailto:revista@aulavirtual.web.ve)

URL: <http://aulavirtual.web.ve/revista>

La Revista seriada Científica Arbitrada e Indexada **Aula Virtual**, es de acceso abierto y en formato electrónico; la misma está orientada a la divulgación de las producciones científicas creadas por investigadores en diversas áreas del conocimiento. Su cobertura temática abarca Tecnología, Ciencias de la Salud, Ciencias Administrativas, Ciencias Sociales, Ciencias Jurídicas y Políticas, Ciencias Exactas y otras áreas afines. Su publicación es **CONTINUA**, indexada y arbitrada por especialistas en el área, bajo la modalidad de doble ciego. Se reciben las producciones tipo: *Artículo Científico* en las diferentes modalidades cualitativas y cuantitativas, *Avances Investigativos*, *Ensayos*, *Reseñas Bibliográficas*, *Ponencias o publicaciones derivada de eventos*, y cualquier otro tipo de investigación orientada al tratamiento y profundización de la información de los campos de estudios de las diferentes ciencias. La Revista **Aula Virtual**, busca fomentar la divulgación del conocimiento científico y el pensamiento crítico reflexivo en el ámbito investigativo.



ISSN: 2665-0398  
Depósito Legal: LA2020000026  
País: Venezuela  
Año de Inicio: 2020  
Periodicidad: Continua  
Sistema de Arbitraje: Revisión por pares. "Doble Ciego"  
Licencia: Creative Commons [CC BY NC ND](https://creativecommons.org/licenses/by-nc-nd/4.0/)  
Volumen: 7  
Número: 15  
Año: 2026  
Período: Julio 2026 – Diciembre 2026 (continua)  
Dirección Fiscal: Av. Libertador, Arca del Norte, Nro. 52D, Barquisimeto estado Lara, Venezuela, C.P. 3001



# DESEABILIDAD SOCIAL Y VALIDEZ DE LOS AUTOINFORMES DE AGRESIVIDAD EN AGENTES POLICIALES: UNA REVISIÓN SISTEMÁTICA

## SOCIAL DESIRABILITY AND THE VALIDITY OF AGGRESSION SELF-REPORTS IN POLICE OFFICERS: A SYSTEMATIC REVIEW

**Tipo de Publicación:** Artículo Científico

**Área del Conocimiento:** Ciencias Sociales y Aplicadas

**Recibido:** 16/07/2025

**Aceptado:** 19/08/2025

**Publicado:** 22/08/2026

**Código Único AV:** e842

**Páginas:** 1(961-983)

**DOI:** <https://doi.org/10.5281/zenodo.22057838>

### Autores:

**Jhon Paul Trujillo Valer**

Licenciado en Sociología

Doctor en Medio Ambiente y Desarrollo Sostenible

 <https://orcid.org/0000-0002-8179-8645>

**E-mail:** [jtrujillo@unheval.edu.pe](mailto:jtrujillo@unheval.edu.pe)

**Afiliación:** Universidad Nacional Hermilio Valdizán

**País:** República del Perú

**Félix Dulio Sobrado Chávez**

Licenciado en Psicología

Maestro, mención en Intervención Psicológica e Investigación

 <https://orcid.org/0009-0008-5566-9330>

**E-mail:** [fsobrado@unheval.edu.pe](mailto:fsobrado@unheval.edu.pe)

**Afiliación:** Universidad Nacional Hermilio Valdizán

**País:** República del Perú

**Ivan Agui Guillen**

Licenciado en Comunicación Social

Maestro en Educación, mención Investigación y Docencia Superior

 <https://orcid.org/0000-0002-0730-2234>

**E-mail:** [ivan.agui@udh.edu.pe](mailto:ivan.agui@udh.edu.pe)

**Afiliación:** Universidad de Huánuco

**País:** República del Perú

**Pedro Máximo Joaquín Ojeda Gallo**

Médico Cirujano

Doctor en Ciencias de la Salud

 <https://orcid.org/0009-0000-4816-8724>

**E-mail:** [pojedad@unp.edu.pe](mailto:pojedad@unp.edu.pe)

**Afiliación:** Universidad Nacional de Piura

**País:** República del Perú

### Resumen

La evaluación de la agresividad en agentes policiales mediante autoinformes resulta relevante para los procesos de selección, seguimiento y prevención de conductas de riesgo; sin embargo, estas mediciones pueden verse afectadas por la deseabilidad social, especialmente cuando las respuestas pueden generar consecuencias profesionales o disciplinarias. El objetivo de este artículo fue analizar la evidencia científica sobre el efecto de la deseabilidad social en la validez de las puntuaciones de agresividad obtenidas mediante autoinformes en agentes policiales. Se realizó una revisión sistemática, siguiendo criterios explícitos de búsqueda, selección y síntesis de estudios publicados en revistas científicas, con atención a investigaciones que evaluaron agresividad, ira, personalidad, actitudes relacionadas con la violencia y estrategias de control del sesgo de respuesta en población policial. Los resultados mostraron que la evidencia directa fue limitada y heterogénea. El BPAQ-SF fue el principal instrumento identificado para medir agresividad, mientras que el STAXI-2, el NEO-PI-3 y la AMMSA evaluaron dimensiones complementarias. La deseabilidad social fue reconocida como una amenaza probable para la validez, sobre todo en reactivos referidos a amenazas, agresión física, hostilidad y creencias profesionalmente censurables; no obstante, la mayoría de los estudios no utilizó escalas independientes para cuantificarla. También se identificaron estrategias como la incorporación de escalas de simulación y disimulación, la triangulación con registros conductuales, la reformulación de reactivos y el uso de diseños longitudinales. Se concluye que las puntuaciones bajas de agresividad no deben interpretarse automáticamente como ausencia del rasgo, y que la evaluación policial requiere enfoques multimétodo con controles específicos de respuesta válidos.

### Palabras Clave

Deseabilidad social, agresividad, autoinformes, agentes policiales, validez psicométrica.

### Abstract

Assessing aggression in police officers through self-report instruments is relevant to personnel selection, professional monitoring, and the prevention of risk-related conduct; however, these measurements may be affected by social desirability, particularly when responses may lead to professional or disciplinary consequences. This article aimed to analyse the scientific evidence on the effect of social desirability on the validity of aggression scores obtained through self-reports in police officers. A systematic review was conducted using explicit criteria for searching, selecting, and synthesizing studies published in scientific journals, with particular attention to research assessing aggression, anger, personality, violence-related attitudes, and response-bias control strategies in police populations. The findings showed that direct evidence was limited and heterogeneous. The BPAQ-SF was the main instrument identified for measuring aggression, whereas the STAXI-2, NEO-PI-3, and AMMSA assessed complementary dimensions. Social desirability was recognized as a plausible threat to validity, especially in items addressing threats, physical aggression, hostility, and professionally censured beliefs; nevertheless, most studies did not use independent scales to quantify this bias. The review also identified methodological strategies such as incorporating simulation and dissimulation scales, triangulating self-reports with behavioral records, reformulating sensitive items, and applying longitudinal designs. It is concluded that low aggression scores should not automatically be interpreted as evidence of the absence of aggressive tendencies. Police psychological assessment therefore requires multimethod approaches that integrate specific response-validity indicators, external behavioral criteria, and repeated measurements to distinguish genuine psychological characteristics from strategically favorable self-presentation in high-stakes evaluation contexts across selection, training, promotion, and routine occupational follow-up processes.

**Keywords** Social desirability, aggression, self-reports, police officers, psychometric validity.

## Introducción

La evaluación psicológica de los agentes policiales constituye un campo de creciente interés científico, debido a que las fuerzas del orden desempeñan sus funciones en escenarios caracterizados por una elevada presión operativa, exposición permanente al conflicto y necesidad de adoptar decisiones inmediatas. En tales circunstancias, determinadas respuestas agresivas pueden escalar hasta traducirse en un uso desproporcionado de la fuerza, con serias repercusiones sociales, institucionales y jurídicas.

La agresividad en esta población suele evaluarse predominantemente mediante instrumentos de autoinforme. Sin embargo, este tipo de medición presenta una vulnerabilidad inherente asociada con la deseabilidad social, entendida como la propensión de las personas a emitir respuestas que proyecten una imagen favorable y se ajusten a aquello que consideran social o culturalmente aceptable (Louzán, 2020). Este sesgo adquiere particular relevancia en el ámbito policial, puesto que los evaluados pueden asumir que sus respuestas influirán en decisiones relacionadas con su ingreso, permanencia, promoción profesional o eventual responsabilidad disciplinaria. En consecuencia, las puntuaciones obtenidas podrían no reflejar fielmente sus disposiciones agresivas, sino una presentación estratégicamente favorable de sí mismos.

La deseabilidad social representa uno de los sesgos de respuesta más ampliamente estudiados en la investigación psicológica y constituye una amenaza sustancial para la validez de los estudios desarrollados en psicología y ciencias sociales (Louzán, 2020). En las evaluaciones sustentadas en instrumentos de autoinforme, la aquiescencia y la deseabilidad social figuran entre los estilos de respuesta con mayor capacidad para distorsionar los resultados, pudiendo manifestarse de manera simultánea y alterar la interpretación de las puntuaciones obtenidas (Louzán, 2020).

Desde la perspectiva de la cognición implícita, los autoinformes evidencian limitaciones especialmente relevantes cuando exploran comportamientos, actitudes o preferencias consideradas sensibles. En estos escenarios, las personas pueden ocultar, atenuar o reformular sus verdaderas intenciones con el propósito de generar una impresión positiva y responder de acuerdo con las expectativas sociales percibidas (Parra & Tamayo, 2021). Frente a esta dificultad, se han desarrollado medidas implícitas orientadas a reducir la influencia de sesgos conscientes, entre ellos la deseabilidad social, sobre las respuestas de los participantes (Parra & Tamayo, 2021).

No obstante, la evaluación de la agresividad en agentes policiales enfrenta, además, limitaciones relacionadas con la especificidad de los instrumentos utilizados. Christopher et al., (2024)



sostienen que herramientas tradicionales como el Inventario Multifásico de Personalidad de Minnesota-2 (MMPI-2) fueron concebidas principalmente para evaluar psicopatología general y no para medir de manera específica la conducta agresiva. Por consiguiente, estos instrumentos carecen de indicadores directos de agresión suficientemente validados para la población policial, circunstancia que restringe su capacidad para identificar con precisión determinadas manifestaciones conductuales relevantes para el ejercicio de la función.

Ceschi et al., (2022) compararon a 149 candidatos policiales suizos con 110 participantes civiles. Los resultados mostraron que los aspirantes a la función policial se describieron a sí mismos como menos agresivos, ansiosos e impulsivos que los integrantes del grupo civil. Paradójicamente, también presentaron niveles significativamente más elevados de deseabilidad social. Este hallazgo permite inferir que las diferencias observadas en los autoinformes podrían no responder exclusivamente a características psicológicas reales, sino también a una tendencia de los candidatos a proyectar una imagen compatible con el perfil profesional esperado.

Los autores señalaron, además, que el estilo de afrontamiento represivo identificado en los candidatos policiales podría funcionar como un mecanismo protector en el corto plazo, al facilitar el

control emocional en situaciones de tensión. Sin embargo, sostenido en el tiempo, este patrón también podría convertirse en un factor de riesgo para el equilibrio psicológico, debido a la inhibición persistente de emociones y conflictos internos (Ceschi et al., 2022).

Por su parte, Christopher et al., (2024) analizaron las propiedades psicométricas de la versión abreviada del Cuestionario de Agresión de Buss-Perry (BPAQ-SF) en una muestra de agentes policiales. Los resultados evidenciaron una adecuada consistencia interna ( $\alpha = .84$ ) y una elevada fiabilidad test-retest ( $r = .86$ ). Estos hallazgos respaldaron su utilización como una medida breve y específica de agresividad y representaron uno de los primeros esfuerzos de validación de un instrumento de esta naturaleza en población policial.

Asimismo, Sellbom et al., (2022) examinaron la validez incremental del Multidimensional Personality Questionnaire (MPQ) respecto del MMPI-2-RF en una muestra de 1.687 candidatos policiales. Los investigadores encontraron que el MPQ incrementó la capacidad predictiva en 17 de las 19 variables analizadas. Estos resultados pusieron de manifiesto que la combinación de distintos instrumentos puede proporcionar información complementaria y clínicamente relevante sobre la idoneidad psicológica de los postulantes. En otras palabras, la evaluación

multimétodo permitiría superar algunas de las restricciones derivadas del empleo aislado de pruebas generales de personalidad o psicopatología.

Pese a los avances alcanzados, la literatura especializada todavía presenta vacíos teóricos, metodológicos y psicométricos de considerable importancia. Zakaria et al., (2023), mediante una revisión sistemática sobre personalidad y desempeño laboral en agentes policiales, constataron que los instrumentos empleados en este ámbito suelen privilegiar la identificación de psicopatologías y rasgos generales de personalidad por encima de la evaluación específica de la conducta agresiva. De igual modo, observaron una insuficiente consideración de la forma en que la deseabilidad social puede comprometer la autenticidad de las respuestas y, por extensión, la validez de las puntuaciones.

Louzán (2020) advierte que, aunque existe una amplia diversidad de instrumentos destinados a medir la deseabilidad social —entre ellos, la Marlowe-Crowne Social Desirability Scale y las escalas L y K del MMPI—, la detección y el control de este sesgo suelen ser soslayados por los profesionales encargados de las evaluaciones psicológicas. Esta omisión resulta especialmente preocupante en procesos de selección o permanencia policial, en los cuales el evaluado posee razones evidentes para ofrecer una imagen favorable y minimizar aquellos comportamientos

que podrían ser interpretados como incompatibles con el cargo.

De manera complementaria, Parra & Tamayo (2021) sostienen que las medidas implícitas presentan ventajas potenciales frente a los autoinformes, debido a que son menos susceptibles a la manipulación consciente y a las respuestas condicionadas por la deseabilidad social. Sin embargo, también precisan que dichas herramientas no están exentas de error y deben utilizarse con cautela, dado que sus resultados pueden verse afectados por factores contextuales, cognitivos y procedimentales. A ello se suma que su aplicación en el análisis de la personalidad y la agresividad policial continúa siendo incipiente, lo que impide establecer conclusiones definitivas sobre su eficacia y pertinencia.

En este orden de ideas, se identifica una brecha relacionada con la escasez de estudios que examinen de manera integrada la agresividad, la deseabilidad social y la validez de los autoinformes empleados en población policial. Aunque Ceschi et al., (2022) evaluaron simultáneamente agresividad y deseabilidad social en candidatos policiales y encontraron que estos presentaban menores niveles autoinformados de agresividad y mayores niveles de deseabilidad social en comparación con participantes civiles, la evidencia disponible continúa siendo limitada. Asimismo, son todavía insuficientes las investigaciones dirigidas a

determinar en qué medida la deseabilidad social modifica específicamente las puntuaciones de agresividad y qué estrategias de evaluación permiten distinguir entre características psicológicas auténticas y respuestas condicionadas por la necesidad de proyectar una imagen profesional favorable.

A partir de los vacíos identificados, el presente artículo tiene como objetivo analizar, mediante una revisión sistemática, la evidencia científica disponible sobre el efecto de la deseabilidad social en la validez de las puntuaciones de agresividad obtenidas a través de instrumentos de autoinforme aplicados a agentes policiales.

### Metodología

La presente revisión sistemática se desarrolló siguiendo las directrices de la declaración PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Page et al., 2021), con el propósito de asegurar una presentación transparente y trazable de las etapas de identificación, cribado, elegibilidad e inclusión de los estudios.

La búsqueda bibliográfica se realizó exclusivamente en la base de datos Scopus, seleccionada por su amplia cobertura multidisciplinaria de literatura científica revisada por pares y por disponer de herramientas avanzadas para la aplicación de estrategias de búsqueda mediante operadores booleanos. La búsqueda

estuvo orientada a identificar estudios desarrollados en población policial o en candidatos policiales que abordaran conjuntamente la deseabilidad social o formas relacionadas de distorsión de respuesta y la agresividad o constructos estrechamente vinculados, dentro de procesos de evaluación psicológica, selección, seguimiento o valoración psicométrica.

La estrategia de búsqueda se estructuró mediante cuatro bloques conceptuales independientes: a) población policial, b) deseabilidad social o distorsión de respuesta, c) agresividad o constructos relacionados y d) evaluación psicológica o psicométrica. Los términos pertenecientes a un mismo bloque fueron combinados mediante el operador OR, mientras que los cuatro bloques fueron relacionados mediante AND, con la finalidad de exigir que los registros recuperados abordaran simultáneamente los componentes centrales de la pregunta de investigación.

La ecuación aplicada en Scopus fue la siguiente:

TITLE-ABS-KEY((police OR policing OR "police officer\*" OR "law enforcement officer\*" OR "law enforcement personnel" OR "police personnel" OR "police candidate\*" OR "police applicant\*" OR "police recruit\*") AND ("social desirability" OR "socially desirable responding" OR "impression management" OR

"response distortion" OR "response bias" OR "defensive responding" OR defensiveness OR faking OR dissimulation OR underreport\*) AND (aggress\* OR hostile\* OR anger OR impulsiv\*) AND (assessment OR evaluation OR psychometric\* OR questionnaire\* OR inventory OR inventories OR scale\* OR personality OR selection OR screening OR "self-report\*")) AND PUBYEAR > 2019 AND PUBYEAR < 2027

Si bien la ecuación de búsqueda fue formulada mediante términos en inglés, los criterios de elegibilidad permitieron considerar publicaciones comprendidas dentro del período establecido para la revisión. La búsqueda abarcó estudios publicados entre enero de 2020 y junio de 2026 y fue ejecutada en la base de datos Scopus durante el mes de junio de 2026. Los registros identificados fueron posteriormente sometidos a un proceso sistemático de cribado, recuperación de textos completos, evaluación de elegibilidad e inclusión, conforme con los criterios previamente establecidos.

Para guiar la revisión sistemática, se formularon las siguientes preguntas de investigación: ¿En qué medida la deseabilidad social afecta la validez de las puntuaciones de agresividad obtenidas mediante autoinformes en agentes y candidatos policiales? ¿Qué instrumentos psicométricos de autoinforme se han utilizado para medir simultáneamente agresividad y deseabilidad

social en poblaciones policiales, y cuáles incorporan escalas de control de sesgo de respuesta? ¿Qué estrategias metodológicas se han propuesto en la literatura para mitigar el efecto de la deseabilidad social en la evaluación de agresividad en contextos de selección y seguimiento policial?

| Criterios de inclusión                                       | Criterios de exclusión                               |
|--------------------------------------------------------------|------------------------------------------------------|
| Artículos revisados por pares (2020–2026).                   | Población no policial.                               |
| Estudios empíricos o revisiones.                             | Uso exclusivo de medidas implícitas.                 |
| Población policial o en formación policial.                  | Opiniones, editoriales, congresos o literatura gris. |
| Uso de autoinformes de agresividad o variables relacionadas. | Agresividad solo como factor contextual.             |
| Publicados e indexados en Scopus                             | Estudios duplicados o preliminares.                  |

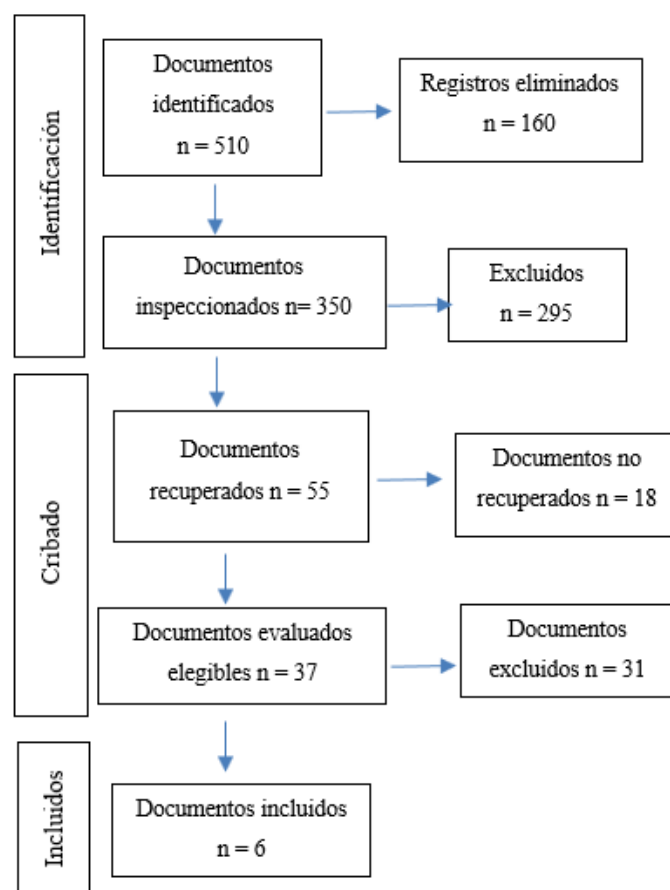
**Tabla 1.** Criterios de inclusión y exclusión

El proceso de selección se desarrolló en dos etapas. En una primera fase se efectuó el cribado de títulos y resúmenes de los registros identificados y, posteriormente, se revisó el texto completo de los estudios potencialmente elegibles. Dos revisores realizaron de manera independiente el proceso de cribado y evaluación de elegibilidad mediante la aplicación de los criterios de inclusión y exclusión previamente establecidos. En los casos en que existió discrepancia entre ambos revisores, el estudio fue sometido a discusión conjunta con los otros dos integrantes del equipo investigador, adoptándose la decisión final mediante consenso de los cuatro autores.



Estos criterios de exclusión se establecieron para garantizar que los estudios seleccionados abordaran directamente la intersección entre discapacidad social, instrumentos de autoinforme y medición de agresividad en el contexto policial, minimizando así el riesgo de incluir evidencia tangencial que pudiera comprometer la coherencia de la síntesis.

Evaluación de la calidad metodológica y riesgo de sesgo. La calidad metodológica de los estudios incluidos fue evaluada mediante las herramientas de evaluación crítica del Joanna Briggs Institute (JBI), seleccionando la lista de verificación correspondiente de acuerdo con el diseño metodológico de cada investigación. La evaluación consideró aspectos relacionados con la selección y caracterización de los participantes, validez y fiabilidad de las mediciones, identificación y tratamiento de posibles factores de confusión, adecuación de los procedimientos estadísticos y coherencia entre el diseño, los resultados y las conclusiones. Cada criterio fue valorado de acuerdo con las categorías establecidas por JBI: “Sí”, “No”, “No está claro” y “No aplica”. La evaluación fue realizada de manera independiente por dos revisores y las discrepancias fueron resueltas mediante consenso. Los resultados de esta valoración se utilizaron para ponderar la solidez de la evidencia durante la síntesis e interpretación de los hallazgos, sin emplearse como criterio automático de exclusión.



**Figura 1.** Flujograma del método PRISMA

## Resultados

| Autor                      | Población                                           | Variable / instrumento                                      | Medición del sesgo | Evidencia principal                                                                                                   | Efecto sobre la validez                                                                                                                            |
|----------------------------|-----------------------------------------------------|-------------------------------------------------------------|--------------------|-----------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| Ceschi et al., (2022)      | 149 candidatos policiales suizos y 110 civiles      | Agresividad (AQ) y deseabilidad social (Marlowe-Crowne SDS) | Sí, directa        | Los candidatos mostraron mayor deseabilidad social y menor agresividad; ambas variables se relacionaron negativamente | La autopresentación favorable puede contribuir a reducir las puntuaciones de agresividad, aunque no explica por sí sola las diferencias observadas |
| Christopher et al., (2024) | 170 agentes policiales de 15 organismos de EE. UU.  | Agresividad (BPAQ-SF)                                       | No                 | Se reconoció posible sesgo de respuesta en ítems sensibles sobre amenazas y agresión física                           | Posible subdeclaración de agresividad; el efecto de la deseabilidad social no fue cuantificado                                                     |
| Posch & Kemme (2025)       | 113 estudiantes policiales y 91 detectives alemanes | Mitos sobre agresión sexual (AMMSA abreviada)               | No                 | Las respuestas sobre actitudes socialmente censuradas podrían estar influidas por presentación favorable              | Puede reducir las puntuaciones observadas, pero el sesgo no fue medido directamente                                                                |

**Tabla 2.** Efecto de la deseabilidad social

La Tabla 2 sintetiza los estudios que abordaron de manera más directa la relación entre agresividad, variables asociadas y deseabilidad social en población policial. Entre ellos, Ceschi et al., (2022) proporcionaron la evidencia más directa al evaluar simultáneamente agresividad y deseabilidad social en candidatos policiales. Sus resultados mostraron que los aspirantes presentaron menores niveles autoinformados de agresividad y mayores niveles de deseabilidad social que los participantes civiles, lo que sugiere que las diferencias observadas podrían estar parcialmente condicionadas por una tendencia a proyectar una imagen compatible con las expectativas asociadas al ejercicio profesional.

Por su parte, Christopher et al., (2024) evaluaron directamente la agresividad mediante el BPAQ-SF y reconocieron que determinados reactivos relacionados con amenazas y agresión física podían ser particularmente sensibles a respuestas defensivas o socialmente favorables. Sin embargo, este estudio no incorporó una escala independiente de deseabilidad social que permitiera cuantificar específicamente dicho efecto. De manera similar, Posch & Kemme (2025) examinaron actitudes relacionadas con la agresión sexual y señalaron que la deseabilidad social podía influir en la reducción de las puntuaciones correspondientes a creencias socialmente censuradas, aunque tampoco utilizaron una medida específica destinada a cuantificar este sesgo.

En conjunto, la Tabla 2 muestra que la evidencia empírica directa continúa siendo limitada. Ceschi et al., (2022) constituyen el principal antecedente identificado en el que agresividad y deseabilidad social fueron evaluadas simultáneamente, mientras que los restantes estudios reconocieron el posible efecto de la presentación favorable sin medirlo de forma independiente. Estos hallazgos respaldan la necesidad de incorporar indicadores específicos de deseabilidad social o validez de respuesta en futuras

evaluaciones de agresividad realizadas en población policial.

Con el propósito de valorar la solidez de la evidencia sintetizada, se efectuó una evaluación crítica de la calidad metodológica de los estudios incluidos. La valoración consideró aspectos relacionados con la selección y caracterización de los participantes, la adecuación y fiabilidad de los instrumentos de medición, el tratamiento de posibles factores de confusión y la pertinencia de los procedimientos analíticos utilizados. Los resultados de esta evaluación se presentan en la Tabla 3.

| Estudio                                   | Diseño                   | Fortalezas metodológicas                                                                                   | Principales limitaciones                                                                            | Valoración           |
|-------------------------------------------|--------------------------|------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------|----------------------|
| Ceschi et al., (2022)                     | Transversal comparativo  | Grupos comparables; instrumentos validados; medición directa de deseabilidad social; análisis multivariado | Diseño transversal, autoinformes y posible sesgo de selección                                       | Riesgo bajo-moderado |
| Christopher et al., (2024)                | Psicométrico             | Buena consistencia interna, estabilidad test-retest y validez convergente del BPAQ-SF                      | No midió directamente deseabilidad social; dependencia del autoinforme                              | Riesgo moderado      |
| Posch & Kemme (2025)                      | Experimental/comparativo | Comparación entre estudiantes policiales y detectives; medición estructurada de actitudes                  | No evaluó directamente deseabilidad social; contexto específico de agresión sexual                  | Riesgo moderado      |
| Tedeholm et al., (2021)                   | Transversal comparativo  | Uso de NEO-PI-3 y comparación con datos normativos                                                         | Muestra pequeña y altamente seleccionada; no controló autopresentación favorable                    | Riesgo moderado      |
| Van Reemst & Jongerling (2024)            | Longitudinal             | Tres mediciones y análisis diferenciado por profesión                                                      | Uso de autoinformes y posible pérdida de participantes; evalúa principalmente exposición a agresión | Riesgo moderado      |
| Wellendorff et al., (2026)                | Transversal comparativo  | Instrumentos validados; control de covariables y bootstrapping                                             | Grupos pequeños, participación voluntaria y ausencia de control de deseabilidad social              | Riesgo moderado      |
| Gutiérrez-Jurado & Romero Martínez (2026) | Validación psicométrica  | Amplia muestra; análisis factoriales; indicadores de fiabilidad y escalas de simulación/disimulación       | Diseño transversal y ausencia de criterio externo de referencia                                     | Riesgo bajo-moderado |

**Tabla 3.** Evaluación de la calidad metodológica y riesgo de sesgo

La Tabla 3 muestra que los estudios incluidos presentaron, en términos generales, una calidad metodológica aceptable, aunque con limitaciones relacionadas principalmente con el predominio de diseños transversales, el uso de autoinformes y la ausencia de medidas específicas de deseabilidad social. Ceschi et al., (2022) destacaron por evaluar directamente este sesgo junto con la agresividad,

mientras que Gutiérrez-Jurado & Romero Martínez (2026) incorporaron indicadores de simulación y disimulación. Estas limitaciones fueron consideradas al interpretar la solidez de la evidencia disponible.

| Autor                          | Instrumento                             | Constructo                      | Resultado principal                                                                   | Control de respuesta      | Aporte                                               |
|--------------------------------|-----------------------------------------|---------------------------------|---------------------------------------------------------------------------------------|---------------------------|------------------------------------------------------|
| Ceschi et al., (2022)          | AQ, versión corta                       | Agresividad                     | Menor agresividad en candidatos policiales; relación negativa con deseabilidad social | No integrado en la escala | Evidencia directa sobre agresividad autoinformada    |
|                                | Marlowe-Crowne SDS                      | Deseabilidad social             | Mayor deseabilidad social en candidatos policiales                                    | Sí                        | Principal medida directa del sesgo identificada      |
|                                | Short-UPPS                              | Impulsividad                    | La urgencia se relacionó con agresividad                                              | No                        | Variable asociada a conducta agresiva                |
| Christopher et al., (2024)     | BPAQ-SF                                 | Agresividad                     | $\alpha = .84$ ; test-retest $r = .86$                                                | No                        | Instrumento específico de agresividad en policías    |
|                                | PROMIS Anger                            | Ira                             | Evidencia convergente con agresividad                                                 | No                        | Complementa la evaluación del BPAQ-SF                |
| Posch & Kemme (2025)           | AMMSA abreviada                         | Actitudes sobre agresión sexual | Influyó en juicios de credibilidad                                                    | No                        | Evalúa actitudes vinculadas con violencia            |
| Tedeholm et al., (2021)        | NEO-PI-3                                | Personalidad                    | Diferencias en rasgos entre policías y población general                              | No                        | Evalúa características generales de personalidad     |
|                                | Angry Hostility                         | Hostilidad                      | Menor hostilidad en policías antiterroristas                                          | No                        | Dimensión próxima a agresividad                      |
| Wellendorff et al., (2026)     | STAXI-2                                 | Ira                             | $\alpha = .79-.91$                                                                    | No                        | Evalúa estado, rasgo, expresión y control de ira     |
| Van Reemst & Jongerling (2024) | Escala de atribución hostil             | Atribución hostil               | Evaluación longitudinal                                                               | No                        | Mide interpretación de intenciones hostiles          |
|                                | Escala de exposición a agresión laboral | Victimización                   | Agresión verbal, amenazas y agresión física recibida                                  | No                        | Evidencia complementaria sobre exposición a agresión |

| Autor                                     | Instrumento | Constructo     | Resultado principal               | Control de respuesta          | Aporte                                                          |
|-------------------------------------------|-------------|----------------|-----------------------------------|-------------------------------|-----------------------------------------------------------------|
| Gutiérrez-Jurado & Romero Martínez (2026) | VRS-PN      | Riesgo suicida | Adecuada capacidad discriminativa | Sí: simulación y disimulación | Modelo útil para incorporar indicadores de validez de respuesta |

**Tabla 4.** Instrumentos psicométricos utilizados

La Tabla 4 muestra que los instrumentos utilizados fueron heterogéneos. El AQ y el BPAQ-SF evaluaron directamente agresividad, mientras que el STAXI-2 y la faceta Angry Hostility examinaron dimensiones relacionadas con ira y hostilidad. Ceschi et al., (2022) constituyen el antecedente más relevante al incorporar conjuntamente una medida de agresividad y la

Marlowe-Crowne Social Desirability Scale, permitiendo evaluar directamente la deseabilidad social. En contraste, la mayoría de los estudios no empleó controles específicos de respuesta. Gutiérrez-Jurado & Romero Martínez (2026) representaron una excepción al incluir escalas de simulación y disimulación, aunque aplicadas a riesgo suicida.

| Autor                                     | Estrategia                                            | Tipo                 | Finalidad                                                         |
|-------------------------------------------|-------------------------------------------------------|----------------------|-------------------------------------------------------------------|
| Ceschi et al., (2022)                     | Medición de deseabilidad social y control estadístico | Aplicada             | Identificar y controlar la autopresentación favorable             |
|                                           | Uso de medidas implícitas                             | Recomendación futura | Reducir la manipulación consciente de respuestas                  |
| Christopher et al., (2024)                | Validez convergente                                   | Aplicada             | Contrastar la agresividad con variables psicológicas relacionadas |
|                                           | Triangulación con registros conductuales              | Recomendación futura | Comparar autoinforme con conducta observable                      |
|                                           | Reformulación de reactivos sensibles                  | Recomendación futura | Disminuir respuestas defensivas                                   |
| Posch & Kemme (2025)                      | Incorporar escala de deseabilidad social              | Recomendación futura | Cuantificar y controlar el sesgo de respuesta                     |
|                                           | Seguimiento longitudinal                              | Recomendación futura | Examinar cambios durante la trayectoria policial                  |
| Gutiérrez-Jurado & Romero Martínez (2026) | Escalas de simulación y disimulación                  | Aplicada             | Detectar exageración u ocultamiento de respuestas                 |
| Wellendorff et al., (2026)                | Control de covariables y bootstrapping                | Aplicada             | Reducir confusión y mejorar robustez estadística                  |
|                                           | Evaluación multidimensional de la ira                 | Aplicada             | Diferenciar estado, rasgo, expresión y control                    |
|                                           | Diseño longitudinal                                   | Aplicada             | Analizar temporalidad de las relaciones                           |



| Autor                          | Estrategia                                   | Tipo     | Finalidad                                          |
|--------------------------------|----------------------------------------------|----------|----------------------------------------------------|
| Van Reemst & Jongerling (2024) | Análisis por grupo ocupacional               | Aplicada | Evitar comparaciones inadecuadas entre profesiones |
| Tedeholm et al. (2021)         | Instrumento estandarizado y datos normativos | Aplicada | Comparar perfiles policiales con población general |

**Tabla 5.** Estrategias aplicadas y recomendaciones futuras para controlar el sesgo de respuesta

La Tabla 5 diferencia las estrategias efectivamente aplicadas de las recomendaciones metodológicas propuestas para futuras investigaciones. Entre las estrategias utilizadas destacaron la medición directa de discapacidad social, las escalas de simulación y disimulación, el control de covariables, la validez convergente y los diseños longitudinales. Como recomendaciones futuras se identificaron la incorporación de escalas específicas de discapacidad social, la triangulación con registros conductuales, la reformulación de reactivos sensibles y el uso de medidas implícitas. En conjunto, los estudios sugieren que el control del sesgo de respuesta requiere combinar procedimientos psicométricos, estadísticos y multimétodo.

### Discusión de resultados

La presente revisión sistemática tuvo como propósito analizar la evidencia científica disponible acerca de la incidencia de la discapacidad social en la validez de las puntuaciones de agresividad obtenidas mediante instrumentos de autoinforme aplicados a agentes policiales. Los hallazgos

muestran que la discapacidad social constituye una amenaza metodológica plausible y especialmente relevante en este ámbito profesional. Sin embargo, la evidencia empírica que examina directamente su relación con la agresividad autoinformada continúa siendo limitada. Si bien varios estudios reconocieron la posibilidad de subdeclaración, disimulación o autopresentación favorable, solo excepcionalmente incorporaron medidas independientes capaces de cuantificar dicho sesgo.

Esta limitación dificulta establecer si una puntuación reducida de agresividad representa una menor disposición agresiva auténtica o una modificación estratégica de las respuestas ante las consecuencias reales o percibidas de la evaluación. Esta distinción adquiere especial importancia en contextos policiales, puesto que reconocer comportamientos hostiles, impulsivos o violentos puede interpretarse como una señal de menor idoneidad profesional y repercutir en procesos de selección, permanencia, promoción o asignación de funciones.

El hallazgo más directamente relacionado con el objetivo de la revisión procedió de Christopher et

al., (2024), quienes examinaron las propiedades psicométricas del Buss–Perry Aggression Questionnaire–Short Form (BPAQ-SF) en agentes policiales. La consistencia interna global y la estabilidad temporal respaldaron su utilización como medida breve de agresividad. No obstante, el comportamiento menos favorable de algunos reactivos relacionados con amenazas y agresión física sugiere que los contenidos más explícitos, comprometedores y profesionalmente censurados podrían ser particularmente vulnerables a respuestas defensivas.

Este resultado guarda relación con lo sostenido por Louzán (2020), quien advierte que la deseabilidad social y otros estilos de respuesta pueden reducir la autenticidad de los autoinformes cuando los evaluados consideran que determinadas respuestas podrían ocasionarles consecuencias adversas. En población policial, esta posibilidad adquiere especial relevancia porque reconocer hostilidad, impulsividad o conductas agresivas puede percibirse como incompatible con los estándares de autocontrol, disciplina y responsabilidad exigidos profesionalmente.

Los hallazgos de Christopher et al., (2024) también convergen con los de Ceschi et al., (2022). Estos autores observaron que los candidatos policiales se autodescribieron como menos agresivos, ansiosos e impulsivos que los participantes civiles, pero simultáneamente

presentaron niveles superiores de deseabilidad social. Esta combinación indica que los perfiles psicológicos aparentemente favorables no deberían atribuirse exclusivamente a diferencias reales entre poblaciones, puesto que también podrían reflejar una tendencia a responder conforme al prototipo profesional esperado.

La principal diferencia metodológica entre ambas investigaciones radica en que Ceschi et al., (2022) midieron expresamente la deseabilidad social, mientras que Christopher et al., (2024) la plantearon como una posible explicación del funcionamiento diferencial de determinados reactivos. Por ello, el estudio de Ceschi et al., (2022) aporta un indicio más directo de que las puntuaciones reducidas de agresividad pueden coexistir con una elevada motivación por proyectar una imagen favorable.

La comparación entre estos estudios permite plantear que la deseabilidad social no necesariamente invalida un instrumento de agresividad en su totalidad. El BPAQ-SF conservó propiedades psicométricas globalmente adecuadas en la muestra policial analizada por Christopher et al., (2024), pero determinados componentes mostraron mayor vulnerabilidad. Esto sugiere que la distorsión podría operar de manera selectiva, afectando con mayor intensidad aquellos reactivos cuya aceptación implica reconocer conductas

contrarias a la legalidad, la disciplina, el autocontrol o la ética profesional.

En consecuencia, resulta necesario diferenciar entre la calidad psicométrica general de una prueba y la autenticidad de las respuestas obtenidas. Un instrumento puede presentar niveles adecuados de fiabilidad y estabilidad sin que ello descarte la existencia de respuestas estratégicamente modificadas. Por tanto, los indicadores psicométricos favorables deben interpretarse juntamente con las condiciones en las que se administra el instrumento y con la motivación del participante para responder de forma transparente.

Los resultados obtenidos por Posch & Kemme (2025) amplían esta problemática hacia las actitudes relacionadas con la agresión sexual. Aunque la Acceptance of Modern Myths About Sexual Aggression Scale (AMMSA) no evalúa directamente la agresividad personal, sus reactivos exploran creencias mediante las cuales la agresión sexual puede ser justificada, minimizada o normalizada. Los autores identificaron niveles relativamente bajos de aceptación de estos mitos, pero reconocieron que la deseabilidad social podría haber contribuido a reducir las puntuaciones registradas.

Este resultado es coherente con Christopher et al., (2024) en cuanto a que los contenidos sujetos a una intensa censura social o institucional pueden presentar una mayor probabilidad de

subdeclaración. También coincide con Parra & Tamayo (2021), quienes sostienen que los autoinformes presentan especial vulnerabilidad cuando evalúan actitudes sensibles, ya que los participantes pueden ocultar, matizar o reformular sus respuestas para ajustarse a las expectativas sociales predominantes.

El estudio de Posch & Kemme (2025) muestra, además, que la posible distorsión no afectaría únicamente la descripción psicológica individual, sino también el análisis de las relaciones entre variables. Si las actitudes problemáticas son subestimadas debido a la autopresentación favorable, la asociación entre dichas actitudes y determinados comportamientos o decisiones profesionales también podría verse artificialmente reducida. En consecuencia, el sesgo de respuesta puede afectar tanto la medición de un constructo como la estimación de su capacidad explicativa o predictiva.

La revisión también puso de manifiesto la heterogeneidad de los instrumentos utilizados para aproximarse al estudio de la agresividad y de variables asociadas en población policial. El BPAQ-SF fue el instrumento directamente dirigido a la medición de la agresividad, mientras que el STAXI-2 examinó componentes de la ira, el NEO-PI-3 evaluó rasgos generales de personalidad y la AMMSA analizó actitudes vinculadas con la agresión sexual. Esta diversidad exige cautela al

comparar los resultados, ya que ira, agresividad, hostilidad, personalidad y actitudes hacia la violencia constituyen constructos relacionados, pero conceptualmente diferentes.

Wellendorff et al., (2026) demostraron la utilidad del STAXI-2 para diferenciar entre ira-estado, ira-rasgo, expresión interna, expresión externa y control de la ira. Esta delimitación resulta relevante porque impide equiparar automáticamente la experiencia emocional de ira con la conducta agresiva. Un agente puede experimentar ira sin exteriorizarla, mientras que otro puede presentar dificultades en su regulación, aunque registre niveles moderados de dicha emoción.

Los resultados de Wellendorff et al, (2026), según los cuales los integrantes de fuerzas policiales especiales no mostraron niveles superiores de ira en comparación con sus compañeros no desplegados y presentaron indicadores favorables respecto de otros servicios de emergencia, convergen parcialmente con el perfil descrito por Tedeholm et al., (2021). Estos últimos identificaron menores niveles de neuroticismo, hostilidad irritable y vulnerabilidad, así como mayores niveles de extraversión y responsabilidad, entre agentes pertenecientes a unidades antiterroristas frente a la población general.

Estos perfiles pueden explicarse por procesos de autoselección, selección institucional, formación y entrenamiento especializado que favorecen la

integración de grupos psicológicamente estables y con mayores recursos de regulación emocional. No obstante, la ausencia de medidas específicas de deseabilidad social en ambas investigaciones obliga a mantener una interpretación prudente. Los resultados favorables pueden reflejar características auténticas de los agentes seleccionados, pero también estar parcialmente condicionados por la necesidad de mantener una imagen coherente con las exigencias de unidades profesionales altamente especializadas.

Esta cautela interpretativa coincide con Sellbom et al., (2022), quienes comprobaron que el Multidimensional Personality Questionnaire aportó capacidad predictiva adicional al MMPI-2-RF en la evaluación de candidatos policiales. Su investigación respalda la idea de que ninguna prueba aislada permite captar de manera exhaustiva las características psicológicas relevantes para el desempeño profesional.

No obstante, la combinación de varios instrumentos de personalidad no elimina por sí misma el problema de la deseabilidad social. Cuando las distintas pruebas comparten el formato de autoinforme y se administran en escenarios de consecuencias elevadas, pueden permanecer expuestas a una fuente común de distorsión. Por ello, la diversificación instrumental debería complementarse con escalas de validez, medidas específicas de deseabilidad social y criterios

externos que permitan contrastar la información proporcionada por los participantes.

En esta línea, la VRS-PN estudiada por Gutiérrez-Jurado & Romero Martínez (2026) incorporó indicadores de simulación y disimulación en población policial. Aunque dicho instrumento fue diseñado para evaluar riesgo suicida y no agresividad, demuestra la viabilidad de integrar mecanismos de control de respuesta en pruebas adaptadas al contexto policial. Esta estrategia resulta coherente con Louzán (2020), quien recomienda identificar y controlar los estilos de respuesta mediante escalas especializadas.

La transferencia de esta lógica al estudio de la agresividad permitiría diferenciar con mayor precisión las puntuaciones bajas auténticas de aquellas asociadas con minimización, disimulación o autopresentación excesivamente favorable. Sin embargo, también resulta necesario diferenciar la deseabilidad social de otras amenazas a la validez de los autoinformes.

Van Reemst & Jongerling (2024) señalaron que la interpretación de conceptos como agresión o victimización puede variar entre participantes y grupos profesionales. Su propuesta de utilizar descripciones de comportamientos concretos en lugar de categorías generales contribuye a reducir la ambigüedad conceptual, pero no necesariamente controla la autopresentación favorable. Un agente puede comprender adecuadamente un reactivo

referido a amenazas o comportamientos hostiles y, aun así, negar su ocurrencia para proteger su imagen profesional.

De manera similar, la aplicación estandarizada utilizada por Tedeholm et al., (2021) disminuye los errores de procedimiento y favorece la comparabilidad, pero no elimina la motivación de los participantes para modificar estratégicamente sus respuestas. Por ello, la ambigüedad semántica, los errores de aplicación, la falta de comprensión, el autoengaño, el manejo de impresiones y la disimulación deliberada deben ser tratados como fuentes de error diferentes y requieren mecanismos específicos de control.

Desde una perspectiva integradora, los resultados coinciden con Zakaria et al., (2023), quienes observaron que la literatura sobre personalidad y desempeño policial ha privilegiado el estudio de rasgos generales y psicopatología frente a la medición específica de la agresividad. La presente revisión añade que, incluso cuando se utilizan instrumentos directamente relacionados con el constructo agresivo, continúa siendo insuficiente el control sistemático de la deseabilidad social.

La evidencia disponible permanece fragmentada entre investigaciones sobre agresividad, ira, personalidad, actitudes relacionadas con la violencia y mecanismos de control de respuesta. Esta fragmentación dificulta establecer un modelo explicativo integrado que



permita precisar cuánto y de qué manera la deseabilidad social modifica las puntuaciones de agresividad obtenidas mediante autoinformes.

Entre las principales limitaciones de la revisión destaca precisamente la escasez de estudios que analizaron de forma directa la relación entre deseabilidad social y agresividad autoinformada en agentes policiales. Solo una proporción reducida de los estudios examinó específicamente la agresividad, mientras que los demás abordaron constructos próximos, como ira, hostilidad, personalidad, actitudes frente a la violencia o indicadores de validez vinculados con otras variables psicológicas. Esta heterogeneidad conceptual y metodológica disminuyó la comparabilidad y obligó a interpretar parte de la evidencia como complementaria y no como demostración directa del fenómeno estudiado.

Asimismo, varios estudios reconocieron la posible existencia de sesgos de respuesta sin evaluarlos mediante escalas independientes. En estos casos, las inferencias sobre subdeclaración procedieron del comportamiento psicométrico de determinados reactivos, de puntuaciones particularmente favorables o de consideraciones teóricas. Esta restricción impidió cuantificar la magnitud de la distorsión e identificar con precisión si las respuestas correspondían a manejo consciente de impresiones, autoengaño o percepciones genuinamente positivas.

La utilización exclusiva de Scopus como fuente de recuperación bibliográfica también restringió la exhaustividad de la revisión, dado que estudios pertinentes indexados en otras bases especializadas pudieron quedar fuera del corpus analizado. Del mismo modo, los criterios temporales y lingüísticos pudieron excluir investigaciones potencialmente relevantes. Por ello, los resultados deben entenderse como representativos de los estudios recuperados conforme a los criterios de búsqueda establecidos.

La generalización de los resultados estuvo igualmente limitada por la diversidad de contextos institucionales y poblaciones analizadas. Los estudios procedieron de organismos policiales con diferentes sistemas de selección, formación, culturas organizacionales y regímenes disciplinarios, e incluyeron candidatos, estudiantes de academias, agentes en servicio e integrantes de unidades especializadas. La motivación para ofrecer respuestas favorables puede variar sustancialmente entre una evaluación anónima con fines académicos y un procedimiento de selección, permanencia o promoción.

Otra limitación fue el predominio de diseños transversales y medidas de autoinforme. La escasez de indicadores conductuales independientes dificultó establecer si las puntuaciones obtenidas se relacionaban efectivamente con el uso de la fuerza, las quejas ciudadanas, las sanciones disciplinarias u

otros incidentes documentados. No obstante, estos registros externos tampoco deben considerarse criterios infalibles, puesto que pueden estar afectados por subregistro, ausencia de denuncias, diferencias institucionales de documentación o desigual aplicación de medidas disciplinarias.

A partir de estas limitaciones, las investigaciones futuras deberían evaluar simultáneamente la agresividad, la deseabilidad social y los mecanismos de simulación o disimulación. La combinación de instrumentos específicos permitiría determinar si el sesgo produce una reducción general de las puntuaciones o si afecta selectivamente dimensiones como agresión física, agresión verbal, ira u hostilidad.

Asimismo, convendría diferenciar empíricamente el autoengaño del manejo de impresiones, pues se trata de procesos distintos que requieren interpretaciones y mecanismos de control diferentes. Los modelos de variables latentes, los análisis de invariancia y otros procedimientos psicométricos podrían contribuir a estimar el efecto específico de estos componentes sobre las puntuaciones.

Otra línea relevante consiste en comparar contextos evaluativos de baja y alta consecuencia. La administración de un mismo instrumento en condiciones anónimas y en escenarios de selección, promoción o permanencia permitiría determinar hasta qué punto las consecuencias percibidas

modifican la forma de responder. De igual manera, la comparación entre candidatos, estudiantes de academias, agentes en servicio e integrantes de unidades especializadas podría esclarecer si la intensidad del sesgo cambia a lo largo de la trayectoria profesional.

Los diseños longitudinales también permitirían diferenciar los efectos de la selección, la formación y la socialización institucional, así como evaluar la estabilidad de la deseabilidad social a través del tiempo. Complementariamente, la triangulación de los autoinformes con registros de uso de la fuerza, quejas ciudadanas, sanciones disciplinarias, evaluaciones de supervisores, observaciones estructuradas, pruebas situacionales y, cuando resulte pertinente, medidas implícitas, proporcionaría una base más robusta para la interpretación de los resultados.

Finalmente, resulta necesario desarrollar normas específicas para la población policial y considerar las particularidades de las distintas funciones profesionales. La agresividad, la ira y la hostilidad pueden manifestarse de manera diferente en labores de patrullaje, investigación criminal, operaciones especiales o tareas administrativas. Por ello, la interpretación de las puntuaciones debería considerar el contexto operativo, la experiencia profesional, la exposición a acontecimientos críticos, la cultura organizacional y las demandas emocionales asociadas con cada función.

## Conclusiones

La revisión sistemática permitió concluir que la deseabilidad social representa una amenaza relevante para la interpretación de las puntuaciones de agresividad obtenidas mediante instrumentos de autoinforme en población policial. La evidencia analizada sugiere que esta vulnerabilidad aumenta cuando los reactivos implican el reconocimiento de amenazas, agresión física, hostilidad u otras conductas incompatibles con los estándares de autocontrol, disciplina y responsabilidad asociados con la función policial. En estos casos, las puntuaciones reducidas no deberían interpretarse automáticamente como evidencia de una menor disposición agresiva, especialmente cuando la evaluación puede producir consecuencias profesionales para el participante.

No obstante, la evidencia disponible no permite determinar con precisión la magnitud del efecto de la deseabilidad social sobre las puntuaciones de agresividad. La principal brecha identificada fue la ausencia de estudios que administraran simultáneamente una medida directa de agresividad y una escala específica de deseabilidad social en agentes policiales. En consecuencia, aunque existen indicios de subdeclaración y autopresentación favorable, todavía no es posible establecer cuánto disminuyen las puntuaciones ni diferenciar con suficiente

certeza entre respuestas auténticas, autoengaño, manejo de impresiones y disimulación deliberada.

Desde el punto de vista psicométrico, los resultados muestran que una adecuada consistencia interna, estabilidad temporal u otras propiedades técnicas favorables no garantizan por sí mismas la autenticidad de las respuestas. Por esta razón, la evaluación psicológica de la agresividad en agentes policiales no debería sustentarse exclusivamente en cuestionarios de autoinforme, sino incorporar mecanismos específicos de control de respuesta y fuentes complementarias de información que permitan valorar la coherencia entre las puntuaciones declaradas y otros indicadores psicológicos o conductuales.

En términos aplicados, los hallazgos respaldan la necesidad de avanzar hacia modelos de evaluación multimétodo que integren medidas directas de agresividad, escalas de deseabilidad social y validez de respuesta, junto con criterios conductuales externos pertinentes al contexto policial. El desarrollo de investigaciones que examinen conjuntamente estos componentes permitirá establecer con mayor rigor cuándo una puntuación baja refleja una característica psicológica genuina y cuándo puede corresponder a una estrategia de autopresentación favorable, fortaleciendo con ello la validez de las decisiones derivadas de la evaluación psicológica policial.

## Referencias

- Ceschi, G., Meylan, S., Rowe, C., & Boudoukha, A. H. (2022). Psychological profile, emotion regulation, and aggression in police applicants: A Swiss cross-sectional study. *Journal of Police and Criminal Psychology*, 37(4), 962–971. Documento en línea. Disponible <https://doi.org/10.1007/s11896-022-09548-0>
- Christopher, M., Ferry, M., Simmons, A., Vasquez, A., Reynolds, B., & Grupe, D. (2024). Psychometric properties of the Buss–Perry Aggression Questionnaire–Short Form among law enforcement officers. *Aggressive Behavior*, 50(2), Article e22145. Documento en línea. Disponible <https://doi.org/10.1002/ab.22145>
- Gutiérrez-Jurado, J. E., & Romero Martínez, S. J. (2026). Enhancing suicide risk detection in police officers: Psychometric properties and standardization of the VRS-PN. *Anuario de Psicología Jurídica*, 36, Article e260482. Documento en línea. Disponible <https://doi.org/10.5093/apj2026a15>
- Louzán Mariño, R. (2020). Mejorar la calidad de las evaluaciones de riesgos psicosociales mediante el control de sesgos. *Archivos de Prevención de Riesgos Laborales*, 23(1), 68–81. Documento en línea. Disponible <https://doi.org/10.12961/aprl.2020.23.01.06>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, Article n71. Documento en línea. Disponible <https://doi.org/10.1136/bmj.n71>
- Parra, J. A., & Tamayo, R. M. (2021). Medidas implícitas en cognición: Una aproximación al estudio de la automaticidad. *Revista Colombiana de Psicología*, 30(1), 89–110. Documento en línea. Disponible <https://doi.org/10.15446/rcp.v30n1.85106>
- Posch, L., & Kemme, S. (2025). Rape myth acceptance in German police officers and its influence on the judgment of victim credibility and credibility of the testimony in a real police interview. *Journal of Police and Criminal Psychology*, 40(1), 141–154. Documento en línea. Disponible <https://doi.org/10.1007/s11896-024-09719-1>
- Sellbom, M., Corey, D. M., & Ben-Porath, Y. S. (2022). Incremental validity of the Multidimensional Personality Questionnaire in the preemployment assessment of police officer candidates. *Criminal Justice and Behavior*, 49(7), 1050–1069. Documento en línea. Disponible <https://doi.org/10.1177/00938548211033630>
- Tedeholm, P. G., Sjöberg, A., & Larsson, A. C. (2021). Personality traits among Swedish counterterrorism intervention unit police officers: A comparison with the general population. *Personality and Individual Differences*, 168, Article 110411. Documento en línea. Disponible <https://doi.org/10.1016/j.paid.2020.110411>
- Van Reemst, L., & Jongerling, J. (2024). Exposure to aggression in three contexts of emergency response: A longitudinal examination of the relationship with psychological characteristics of the emergency responder. *Aggressive Behavior*, 50(4), Article e22160. Documento en línea. Disponible <https://doi.org/10.1002/ab.22160>
- Wellendorff, T., Bender, L., Borck, V., Görlich, T., Pahnke, F., Krüger, J. P., & Wesemann, U. (2026). Anger and quality of life in police special forces after a multiple-day hostage-taking involving severe violence. *Journal of Police and Criminal Psychology*. Advance online publication. Documento en línea. Disponible <https://doi.org/10.1007/s11896-026-09813-6>
- Zakaria, N. Z., Osman, I., & Noranee, S. (2023). Personality traits and job performance of police:

---

A systematic literature review. *Information Management and Business Review*, 15(3(I)), 373–384. Documento en línea. Disponible [https://doi.org/10.22610/imbr.v15i3\(I\).3547](https://doi.org/10.22610/imbr.v15i3(I).3547)