



Consejo General de
Colegios Oficiales
de Psicólogos

Miembro de la Federación Europea
de Asociaciones de Psicólogos

Conde Peñalver, 45. 5º Izq.
28006 Madrid - España
Tel.: +34 91 444 90 20
Fax: +34 91 309 56 15
secop@cop.es
www.cop.es

EVALUACIÓN DE LA ESCALA

EVP-R

RESUMEN DE LA VALORACIÓN DEL TEST

Descripción general

Características	Descripción
Nombre del test	Escala de predicción del riesgo de violencia grave contra la pareja –Revisada- (EPV-R)
Autor	Echeburúa, E., Amor, P., Loinaz, I y Corral, P.
Autor de la adaptación española	-
Editor de la adaptación española	-
Fecha de la última revisión del test en su adaptación española	2010
Constructo evaluado	Riesgo grave de violencia contra la pareja
Área de aplicación	Psicología Forense
Soporte	Papel y lápiz

Valoración general

Características	Valoración	Puntuación
Calidad de los Materiales y documentación		
Fundamentación teórica	Adecuado	3
Adaptación Española		
Análisis de los ítems	Adecuado	3
Validez de contenido	Adecuada	2.5
Validez de constructo	Adecuada pero con algunas carencias	2
Análisis del sesgo		
Validez predictiva	Adecuado	3
Fiabilidad: equivalencia		
Fiabilidad: consistencia interna	Buena	3.5
Fiabilidad: estabilidad		
Baremos		

NOTA: La escala EPV es de libre distribución y carece propiamente de manual. La revisión se ha realizado a partir de dos artículos (cuya referencia bibliográfica aparece en negrita en el apartado 1.31), por lo que no se ha podido encontrar información para valorar varias características de la tabla precedente.

Comentarios generales

La escala EPV es de libre distribución y carece propiamente de manual. La revisión se ha realizado a partir de dos artículos (sus referencias bibliográficas aparecen en negrita en el apartado 1.31).

La escala EPV pretende predecir la probabilidad de violencia grave de un varón (que ha sido denunciado por agresión) contra su pareja o ex-pareja. En la versión revisada, EPV-R, compuesta por los mismos 20 ítems, cambia la cuantificación de cada uno, se modifican los puntos de corte y se propone un tratamiento de las respuestas omitidas. Presentada la denuncia en la comisaría de policía por parte de la mujer agredida, el policía, tras entrevistar a la víctima y al agresor, aplica la escala al agresor y clasifica su nivel de riesgo de agresión grave en una de 3 categorías (bajo, moderado o alto), según su puntuación en el test y los puntos de corte establecidos. El objetivo final es hacer una predicción del nivel de riesgo, basada en criterios objetivos, que permita la adopción de medidas de protección a las víctimas, cuando se produce la primera denuncia.

Entre sus **puntos fuertes** se podrían destacar los siguientes:

1. La actualidad del tema y la potencial aportación a un campo que necesita disponer de estas técnicas de evaluación. El objetivo que se pretende con su aplicación tiene una especial relevancia social y personal.

2. Es una escala no comercializada, realizada íntegramente en nuestro país y ajustada a nuestra cultura. Las escalas alternativas son de origen americano y algunos de sus ítems no son apropiados, como los relacionados con el acceso a las armas de fuego.

3. Se ha conseguido, como se pretendía, una escala de fácil manejo que puede ser utilizada por policías y personal de los servicios sociales, no especialmente instruidos en la administración e interpretación de pruebas psicológicas.

4. La prueba ha sido incluida en el protocolo de acciones a realizar en muchas comisarías de policía tras una denuncia por agresión contra la mujer. Antes de la aplicación de la escala, las medidas de protección las establecía el policía que recibía la denuncia, tras una evaluación subjetiva del nivel de riesgo del agresor. Desde su inclusión en el protocolo, las medidas son básicamente las establecidas para el nivel de riesgo que se deriva de la aplicación de la escala.

5. Las escalas EPV y EPV-R se han aplicado a muestras de varones denunciados por agresión a su pareja de un tamaño aceptable. En un estudio (de 72 casos) la fiabilidad interjueces fue 0.73. Los coeficientes alfas ofrecidos están en el rango 0.66-0.72. La media en el test y en cada uno de los ítems es significativamente diferente en el grupo de casos severos (agresión grave) y en el de casos no severos.

La escala debiera ser mejorada en varios aspectos para incrementar su eficiencia. Entre los **puntos a mejorar** señalamos los que se exponen a continuación:

1. Sería deseable la elaboración de un manual con información detallada sobre el modo de asignar las puntuaciones en los ítems, mostrando ejemplos de cada caso, pues la interpretación de algunos puede ofrecer dudas. En el manual debería incluirse también toda la información sobre diseño, análisis de ítems, fiabilidad, validez e interpretación de las puntuaciones.

2. Debería estudiarse el modo de mejorar la fiabilidad de la escala, pues su principal uso es la estimación individual (no colectiva) del nivel de riesgo. En los artículos revisados no se indica en qué medida la ponderación de los ítems de la escala EPV-R mejora la fiabilidad de las puntuaciones. Cabría explorar otros procedimientos de ponderación que pudieran maximizar la precisión. En la versión revisada los ítems se clasifican en tres subgrupos en función de su correlación ítem-total y se establece la ponderación de cada uno según el subgrupo al que pertenece, ¿sería más apropiado fijar para cada ítem de la escala una ponderación asociada a su poder discriminativo en vez de una común para todos los de cada subgrupo? Los dos primeros ítems tienen un poder discriminativo casi nulo y podrían eventualmente reemplazarse por otros más discriminativos. Convendría informar de la precisión en las clasificaciones de los niveles de riesgo.

3. El estudio de validez predictiva diagnóstica muestra que con el punto de corte propuesto solo el 48% de los pertenecientes al grupo de casos severos es clasificado en la categoría de riesgo alto. En los artículos revisados no se ofrecen datos de un estudio similar con la escala EPV-R, por tanto, se desconoce el impacto que pueda tener en la sensibilidad de la escala los cambios introducidos en la versión revisada. Sería conveniente la realización de estudios adicionales de validez, como el de validez convergente con la escala SARA, que han propuesto los autores de la escala EPV. Podría incorporarse algún estudio de validez de constructo sobre la estructura interna.

4. Convendría prestar más atención al establecimiento de los puntos de corte. Se ha comentado anteriormente que, con el punto corte superior prefijado, la escala EPV tiene una baja sensibilidad. La aplicación del instrumento en muestras sin denuncias por agresión pudiera ayudar a delimitar el punto inferior. Convendría justificar mejor el nuevo punto de corte superior de la versión revisada.

5. Los ítems sin respuesta son ciertamente un problema en la aplicación de la escala. Se propone un procedimiento para tratarlos, se indica cuando la escala no debiera puntuarse y cómo generar la puntuación en la escala si el número de ítems sin respuesta es 8 o menos. Dado que existen procedimientos alternativos disponibles para tratar los ítems sin respuesta, convendría hacer un estudio que justifique la bondad del procedimiento de prorrateo propuesto. Pese al esfuerzo por simplificarlo, el procedimiento resulta complejo y pudiera ser aplicado incorrectamente. Cabría plantearse si no sería mejor opción ofrecer el procedimiento en un programa de ordenador que puntúe al evaluado tras introducir en el programa sus respuestas y omisiones.

En resumen, la escala EPV-R facilita una predicción objetiva de la probabilidad de violencia grave en los varones evaluados, lo que supone un claro avance en relación a la situación anterior, en la que la estimación de la peligrosidad descansaba en el juicio subjetivo del receptor de la denuncia. Futuras versiones de la escala debieran mejorar su sensibilidad (en su estado actual, solo el 48% de los casos severos son calificados como tales).

ANÁLISIS DETALLADO DE LA PRUEBA

1. DESCRIPCIÓN GENERAL DEL TEST

1.1. Nombre del test

Escala de Predicción del Riesgo de Violencia Grave Contra La Pareja- Revisada- (EPV-R).

1.2. Nombre del test en su versión original

-

1.3. Autor

Enrique Echeburúa, Pedro Javier Amor, Ismael Loinaz y Paz de Corral.

1.4. Autor de la adaptación española

1.5. Editor del test en su versión original

1.6. Editor de la adaptación española

1.7. Fecha de publicación del test original

2010 (Revista Psicothema: Vol. 22, nº 4, pp. 1054-1060).

1.8. Fecha de la publicación del test en su adaptación española

1.9. Fecha de la última revisión del test en su adaptación española

2010 (Revista Psicothema: Vol. 22, nº 4, pp. 1054-1060).

1.10. Área general de la variable que pretende medir el test

- Predicción de violencia de género contra la pareja

1.11. Breve descripción de la variable que pretende medir el test

La escala revisada de Predicción del riesgo de violencia grave contra la pareja (EPV-R) pretende predecir la probabilidad de manifestación de un determinado tipo de agresión violenta contra la pareja o ex-pareja y ha sido el resultado de revisar la escala previa EPV (Echeburúa, Fernández-Montalvo, Corral & López-Goñi, 2009). Ambas escalas distinguen entre “violencia grave” o “alta violencia” y “violencia menos grave” o “baja violencia”. Por “violencia grave” entienden cualquier tipo de comportamiento violento físico que intenta causar un daño grave e incluso matar (homicidios o tentativas, empleo de armas u objetos peligrosos o producción de lesiones), y por “violencia menos grave” cualquier otra manifestación física o psíquica que no cumple estos criterios. Ambas escalas tratan de predecir la violencia “grave”.

La versión revisada (EPV-R, Echeburúa, Amor, Loinaz y Corral. 2010) contempla las mismas variables clínicas y está compuesta por los mismos 20 ítems de la escala EPV, si bien cambia la cuantificación de cada uno, los puntos de corte para la toma de decisiones y se proporciona un tratamiento para las respuestas omitidas. Los 20 ítems se agrupan en 5 apartados: I.- Datos personales (1 ítem); II.- Situación de la relación de pareja (2 ítems), III.- Tipo de violencia (7 ítems); IV.- Perfil del agresor (7 ítems) y V.- Vulnerabilidad de la víctima (3 ítems). Cada ítem recibe una determinada ponderación y es cuantificado bien con 0/1, con 0/2 o con 0/3.

Una vez presentada la denuncia en la Comisaría de Policía por parte de la mujer agredida, el policía, tras entrevistar a la víctima y al agresor, aplica la escala al agresor (varones exclusivamente) para obtener el nivel de riesgo que le corresponde (bajo, moderado o alto), según su puntuación en el test y los puntos de corte establecidos. El objetivo final es hacer una predicción que permita la adopción de medidas de protección a las víctimas, cuando se produce la primera denuncia, que sean adecuadas a sus necesidades específicas y que estén basadas en criterios objetivos, en vez de en los criterios subjetivos que han venido aplicándose previamente.

1.12. Áreas de aplicación

- Psicología forense

1.13. Formato de los ítems

- Respuesta dicotómica (sí/no, verdadero/falso, etc)

1.14. Número de ítems

- 20

1.15. Soporte

- Papel y lápiz

1.16. Cualificación requerida para el uso del test de acuerdo con la documentación aportada

- Entrenamiento y Acreditación específica
- Los autores afirman que la escala debe ser cumplimentada por personal de atención directa a agresores y víctimas de violencia contra la pareja (policías, psicólogos o médicos forenses, personal de Justicia, etc.), siempre que hayan recibido una formación específica sobre la evaluación del riesgo de violencia y sobre el uso de la escala. En la actualidad no existe propiamente una acreditación específica, con sus requisitos correspondientes, para el manejo de la escala.

1.17. Descripción de las poblaciones a las que el test es aplicable

El instrumento se aplica tras entrevistar a agresor y víctima (pareja o ex-pareja), después de que ésta interpone una denuncia sobre malos tratos. La escala se aplica a varones, españoles o no, de un rango amplio de edad (entre 16 y 84 años).

1.18. Existencia o no de diferentes formas del test y sus características

La escala EPV es de hecho la versión reducida de una escala original de 58 ítems (Echeburúa, Fernández-Montalvo & Corral, 2008). Los 20 ítems definitivos son los que mejor diferenciaron los casos de violencia grave de aquellos de violencia menos grave. La escala EPV-R está compuesta por los mismos 20 ítems de la escala EPV, si bien cambia la cuantificación de cada uno, los puntos de corte para la toma de decisiones y se proporciona un tratamiento para las respuestas omitidas.

1.19. Procedimiento de corrección

- Manual mediante plantilla

1.20. Procedimiento de obtención de las puntuaciones directas

Hay que sumar los puntos obtenidos en cada ítem. En la escala EPV cada ítem es cuantificado con 0/1. En la escala revisada, cada ítem es puntuado con 0/1, 0/2 o 0/3, según su discriminación (entre violencia grave y menos grave) sea baja, media o alta.

En la escala revisada, para casos de datos incompletos, se ha sugerido un procedimiento de prorrateo de los ítems omitidos (Echeburúa *et al.*, 2010). El procedimiento consiste básicamente en asignar a cada respuesta omitida la media de la puntuación obtenida en los ítems respondidos.

1.21. Transformación de las puntuaciones

- Característica no aplicable para este instrumento

1.22. Escalas utilizadas

Cada puntuación se interpreta clasificando el riesgo del evaluado en “bajo”, “moderado” o “alto”, a partir de la puntuación del evaluado y los dos puntos de corte establecidos.

1.23. Posibilidad de obtener informes automatizados

- No

1.24. Oferta por parte del editor de un servicio para la corrección y/o elaboración informes

- No

1.25. Tiempo estimado para la aplicación del test

- Los autores indican que, en la mayoría de los casos, el tiempo de aplicación se puede estimar entre 30 y 45 minutos.

1.26. Documentación aportada por el editor

El test es de libre distribución y no tiene manual. Se describe en varios artículos y documentos técnicos. La presente evaluación se ha hecho a partir de los más relevantes, que han sido amablemente proporcionados por los autores.

1.27. Precio de un juego completo de la prueba

1.28. Precio y número de ejemplares del paquete de cuadernillos

1.29. Precio y número de ejemplares del paquete de hojas de respuesta

1.30. Precio de la corrección y/o elaboración de informes por parte del editor

1.31. Bibliografía básica acerca del test aportada en la documentación Bibliografía relevante sobre el test. En negrita los dos artículos principales en los que se ha basado la presente revisión.

Amor, P.J., Echeburúa, E., y Loinaz, I. (2009). ¿Se puede establecer una clasificación tipológica de los hombres violentos contra su pareja? *International Journal of Clinical and Health Psychology*, 8, 355-382.

Andrés-Pueyo, A. (2009). La predicción de la violencia contra la pareja. En E. Echeburúa, J. Fernández-Montalvo y P. Corral (Eds.): *Predicción del riesgo de homicidio y violencia grave en la relación de pareja. Instrumentos de evaluación del riesgo y adopción de medidas de protección* (pp. 21-53). Valencia: Centro Reina Sofía.

Andrés-Pueyo, A., López, S., y Álvarez, E. (2008). Valoración del riesgo de violencia contra la pareja por medio de la SARA. *Papeles de Psicólogo*, 29, 107-122.

Andrés-Pueyo, A., y Echeburúa, E. (2010). Valoración del riesgo de violencia: instrumentos disponibles e indicaciones de aplicación. *Psicothema*, 22, 403-409.

Browne, A., Williams, K.R., y Dutton, D.G. (1999). Homicide between intimate partners. En M.D. Smith y M.A. Zahn (Eds.): *Homicide: A source book of social research* (pp. 149-164). Thousand Oaks, CA: Sage.

Caetano, R., Vaeth, P., y Ramisetty-Milker, S. (2008). Intimate partner violence victims and perpetrator characteristics among couples in the United States. *Journal of Family Violence*, 23, 507-518.

Campbell, J. (1995). *Assessing dangerousness: Violence by sexual offenders, batterers and child abusers*. Thousand Oaks, CA: Sage.

Dutton, D.G. (2006). *Rethinking domestic violence*. Vancouver: UBC Press.

Dutton, D.G., y Kropp, P.R. (2000). A review of domestic violence risk instruments. *Trauma, Violence and Abuse*, 1, 171-182.

Echeburúa, E., Amor, P.J., Loinaz, I. y Corral, P. (2010). Escala de predicción del riesgo de violencia grave contra la pareja-revisada- (EPV-R). *Psicothema*, 22, 4, 1054-1060.

Echeburúa, E., Fernández-Montalvo, J., y Amor, P.J. (2003). Psychopathological profile of men convicted of gender violence: A study in the prisons of Spain. *Journal of Interpersonal Violence*, 18, 798-812.

Echeburúa, E., Fernández-Montalvo, J., y Corral, P. (2008). ¿Hay diferencias entre la violencia grave y la violencia menos grave contra la pareja?: un análisis comparativo. *International Journal of Clinical and Health Psychology*, 8, 355-382.

Echeburúa, E., Sarasua, B., Zubizarreta, I., y Corral, P. (2009). Evaluación de la eficacia de un tratamiento cognitivo-conductual para hombres violentos contra la pareja en un marco comunitario: una experiencia de 10 años (1997-2007). *International Journal of Clinical and Health Psychology*, 9, 199-217.

Echeburúa, E., y Fernández-Montalvo, J. (2009). Evaluación de un programa de tratamiento en prisión de hombres condenados por violencia grave contra la pareja. *International Journal of Clinical and Health Psychology*, 9, 5-20.

Echeburúa, E., Fernández-Montalvo, J., Corral, P., y López-Goñi, J.J.(2009). Assessing risk markers in intimate partner femicide and severe violence. A new assessment instrument. *Journal of Interpersonal Violence*, 24, 925-939.

Hilton, N.Z, Harris, G.T., Rice, M.E., Lang, C., Cormier, C.A., y Lines, K.J. (2004). A brief actuarial assessment for the prediction of wife assault recidivism: The Ontario domestic assault risk assessment. *Psychological Assessment*, 16, 267-275.

Hilton, N.Z., Harris, G.T., Rice, M.E., Houghton, R.E. y Eke, A.W. (2008). An indepth actuarial assessment for wife assault recidivism: The domestic violence risk appraisal guide. *Law and Human Behavior*, 32, 150-163.

Hilton, N.Z., Harris, G.T., y Rice, M.E. (2010). *Risk assessment for domestically violent men. Tools for criminal justice, offender intervention and victim services*. Washington, DC: American Psychological Association.

Instituto de la Mujer (2006). *III Macroencuesta sobre la violencia contra las mujeres. Informe de resultados*. Madrid: Ministerio de Trabajo y Asuntos Sociales.

Kerry, G. (1998). *Risk assessment of male batterers in a community setting*. Toronto, ON, Canada. Professional Development Workshop.

Kropp, P.R. (2004). Some questions regarding spousal assault risk assessment. *Violence Against Women*, 10, 676-697.

Kropp, P.R., Hart, S., Webster, C., y Eaves, D. (1999). *Spousal Risk assessment guide user's manual*. Toronto, ON, Canada: Multi-Health Systems and BC Institute Against Family Violence.

Loinaz, I., Echeburúa, E., y Torrubia, R. (2010). Tipologías de agresores contra pareja en prisión. *Psicothema*, 22, 106-111.

Riggs, D.S., Caulfield, M.B., y Street, A.E. (2000). Risk for domestic violence: Factors associated with perpetration and victimization. *Journal of Clinical Psychology*, 56, 1289-1316.

Schumacher, J.A., Feldbau-Kohn, S., Slep, A.M., y Heyman, R.E. (2001). Risk factors for male-to-female partner physical abuse. *Aggression and Violent Behavior*, 6, 281-352.

2. VALORACIÓN DE LAS CARACTERÍSTICAS DEL TEST

Aspectos generales

Contenido	Valoración	Puntuación
2.1. Calidad de los materiales del test		-
2.2. Calidad de la documentación aportada		-
2.3. Fundamentación teórica	Adecuada	3
2.4. Adaptación del test	No aplicable	-
2.5. Calidad de las instrucciones	Adecuada pero con algunas carencias	2
2.6. Facilidad para comprender la tarea	Adecuada pero con algunas carencias	2
2.7. Facilidad para registrar las respuestas	Adecuada pero con algunas carencias	2
2.8. Calidad de los ítems (aspectos formales)	Adecuada	3
2.9. Datos sobre el análisis de los ítems	Adecuados	3

Validez de contenido

Contenido	Valoración	Puntuación
2.10. Calidad de la representación del contenido o dominio	Adecuada	3
2.11. Consultas a expertos	Adecuada pero con algunas carencias	2

- Se ha consultado de manera informal a un pequeño número de expertos

Validez de constructo

Contenido	Valoración	Puntuación
2.12. Diseños empleados	<ul style="list-style-type: none"> Diferencias entre grupos 	
2.13. Tamaño de las muestras	Buena	4
2.14. Procedimiento de selección de las muestras	Véase nota	
2.15. Correlaciones del test con otros tests similares	No se aporta	-
2.16. Calidad de los test empleados como criterio o marcador	No se aporta	-
2.17. Datos sobre el sesgo de los ítems	No se aporta	-

Comentarios sobre la validez de constructo

- Varones que son denunciados por agresión de su pareja o expareja en la Unidad de Violencia de Género de la Policía Autónoma Vasca. En uno de los estudios de validación (Echeburúa *et al.* 2009), el tamaño muestral fue 1.081 casos. En otro (Echeburúa *et al.* 2010), el tamaño fue 450 casos.

Validez predictiva

Contenido	Contenido	Puntuación
2.18. Diseño de selección del criterio	<ul style="list-style-type: none"> Retrospectivo 	
2.19. Tamaño de las muestras	Excelente	5
2.20. Procedimiento de selección de las muestras	-	
2.21. Correlaciones del test con los criterios		

Comentarios sobre validez predictiva

- La EPV se aplicó a dos tipos de muestras, definidas como “de violencia grave” (homicidios, tentativas de homicidio, empleo de armas y producción de lesiones) y “de violencia menos grave” (parejas en las que no se daban los criterios anteriores). Se obtuvieron diferencias de medias estadísticamente significativas entre ambas muestras. Se obtuvieron también los niveles de *sensibilidad* y *especificidad*, derivados de las curvas ROC. Se aconseja un punto de corte de X=10 para diferenciar los casos de violencia grave de los de violencia menos

grave, lo que da lugar a una sensibilidad del 48% (casos con violencia grave que obtienen puntuaciones igual o superior a 10), a una especificidad del 81% (casos con violencia menos grave que obtienen puntuaciones inferiores a 10) y a una eficacia diagnóstica del 73%.

- Los autores indican que no se han hecho propiamente estudios de validez predictiva con la escala. Estos estudios son especialmente complicados en estos contextos, pues las víctimas reciben medidas de protección para reducir el riesgo de conducta violenta por parte de su pareja evaluada
- En el estudio sobre la escala revisada, la muestra está compuesta por 450 casos seleccionados del total de denuncias relacionadas con la violencia contra la pareja o ex-pareja. Los criterios de inclusión son los siguientes: a) ser varón; b) haber sido denunciado por violencia contra su pareja o ex-pareja; y c) tener cumplimentada la escala en su totalidad. Se han excluido 1.412 casos porque la escala estaba incompleta. En otro estudio la EPV se aplicó a dos tipos de muestras, definidas como “de violencia grave” (homicidios, tentativas de homicidio, empleo de armas y producción de lesiones) y “de violencia menos grave” (parejas en las que no se daban los criterios anteriores).

Comentarios generales sobre validez

- Son pocos los estudios de validación aportados en los dos artículos (Echeburúa *et al.* 2009 y Echeburúa *et al.* 2010) sobre los que se basa la revisión. No se informa de la realización de estudios de validez de contenido, si bien se informa del proceso seguido para la creación y selección de ítems, con la utilización de jueces. No se informa de que se hayan correlacionado las puntuaciones en el test con otros similares disponibles (si bien se propone como una tarea a realizar, p. 1059), y tampoco se informa de estudios de validez de constructo, como por ejemplo algún estudio sobre la estructura interna de la escala.
- En ambos trabajos se ofrecen los resultados de un mismo estudio en el que se analiza la capacidad discriminadora de la escala EPV entre dos grupos (casos severos y casos no severos). La diferencia entre ambos grupos en la escala y en cada uno de los 20 ítems es estadísticamente significativa (por cierto, estaría bien indicar los tamaños del efecto). El nivel de sensibilidad (48%) encontrado para el punto de corte recomendado parece bajo, dada la importancia personal y social que tienen los errores en este caso. La especificidad alcanza un valor apreciablemente mayor (81%). La eficacia diagnóstica es del 73% y el área bajo la curva, 0.69. Estos resultados son insuficientes, en especial, por su alta tasa de falsos positivos, dadas las consecuencias personales que pueden derivarse de los errores cometidos al utilizar la escala.

- Se argumenta en los artículos que el test “no busca medir un constructo psicológico o una disposición interna, sino estimar la probabilidad de una conducta violenta” (Echeburúa et al. 2010, p. 1055), lo que haría algunas prácticas psicométricas poco apropiadas. En todo caso, el proceso seguido para la elaboración de la prueba es básicamente el estándar en la elaboración de escalas psicológicas y no parece que los procedimientos de validación al uso no sean de aplicación en este caso.

Fiabilidad

Contenido	Valoración	Puntuación
2.22. Datos aportados sobre fiabilidad	<ul style="list-style-type: none"> Coeficientes de fiabilidad para diferentes grupos de sujetos 	

Equivalencia formas paralelas

Contenido	Valoración	Puntuación
2.23. Tamaño de las muestras	No se aporta	-
2.24. Coeficientes de equivalencia	No se aporta	-

Consistencia interna

Contenido	Valoración	Puntuación
2.25. Tamaño de las muestras	Buena	4
2.26. Coeficientes de consistencia	Adecuada ($0.70 \leq r < 0.80$)	3

Estabilidad (test-retest)

Contenido	Valoración	Puntuación
2.27. Tamaño de las muestras	No se aporta	-
2.28. Coeficientes de estabilidad	No se aporta	-

- Dado que el riesgo de conducta violenta depende de las medidas de protección que se tomen con la víctima, las medidas del nivel de riesgo de una misma persona pudieran ser diferentes en un momento y otro, y no tendría sentido entonces la obtención de los indicadores de estabilidad test-retest en la medida de este constructo.

Comentarios generales sobre fiabilidad

- Los dos artículos aportan poca información sobre la fiabilidad de las puntuaciones. Se informa de un estudio de fiabilidad interjueces: dos policías evaluaron a 72 casos y la fiabilidad interjueces fue de 0,73. Los niveles de consistencia interna en agresores severos (0,69) y no severos (0,66) son algo escasos. No se describen los niveles de error típico de medida ni su incidencia en la estimación de las puntuaciones verdaderas. Tales niveles de precisión son escasos cuando se trata de una escala a aplicar para el diagnóstico individual. No se ofrecen datos de la precisión en la clasificación en las 3 categorías, ni de la precisión de cada medida individual.

Baremación

Contenido	Valoración	Puntuación
2.29. Calidad de las normas		-
2.30. Tamaño de las muestras		-
2.31. Procedimiento de selección de las muestras	Agresores que han sido denunciados en las dependencias policiales por agresión por su propia pareja o ex-pareja	

- Para la interpretación de las puntuaciones no se ofrecen normas, sino dos puntos de corte: 5 y 10 (EPV) y 10 y 24 (EPV-R) que permiten la clasificación del nivel de riesgo de evaluado como "bajo", "moderado" o "alto".

Comentarios generales sobre baremación

- No se ofrecen baremos propiamente en los dos artículos revisados. Para interpretar las puntuaciones se ofrecen dos puntos de corte que permiten obtener el riesgo estimado (bajo, medio o alto). En la escala EPV el punto de corte superior (10) se establece en el estudio con la curva ROC. Los resultados obtenidos, sobre todo el referido a la sensibilidad (48%), es insuficiente. En la escala EPV-R, cuyos valores están comprendidos entre 0 y 48, los puntos de corte se establecen en 10 y 24. No se informa suficientemente del procedimiento seguido para su establecimiento.