



**Universidad  
Europea CANARIAS**

**Efectividad de la Terapia Afirmativa  
en Parejas Adultas del Colectivo  
LGTBQ+ sobre el Bienestar  
Psicológico y Relacional:  
Una Revisión Sistemática**

---

**Trabajo fin de Máster**

---

**Autora:** Aitana Taiane Bispo Pestano

**Tutora:** Andrea Duarte Díaz

6 de febrero de 2026

## Resumen

Las parejas LGBTQ+ pueden experimentar estresores específicos asociados al estigma y al estrés de minorías, lo que puede afectar al bienestar psicológico y relacional. Este trabajo tuvo como objetivo evaluar la efectividad de la terapia afirmativa, en comparación con lista de espera u otras intervenciones psicológicas, sobre el bienestar psicológico y relacional en parejas adultas LGBTQ+.

Se realizó una revisión sistemática siguiendo PRISMA 2020 y recomendaciones del Manual Cochrane. La estrategia de búsqueda preliminar se diseñó en PubMed y se adaptó posteriormente a las bases de datos consultadas, incluyendo PubMed y bases de datos de EBSCOhost (APA PsycInfo, CINAHL with Full Text, MEDLINE Complete y Psychology and Behavioral Sciences Collection), además de búsqueda manual en referencias. Se incluyeron ensayos controlados aleatorizados en adultos ( $\geq 18$  años) LGBTQ+ en contexto de pareja. El riesgo de sesgo se evaluó con RoB 2, complementado con CASPe. Debido a la heterogeneidad clínica y metodológica, se realizó síntesis narrativa sin metanálisis.

Se incluyeron cinco ECA (2009–2025) realizados en Estados Unidos, con un total de 261 parejas. Las intervenciones abarcaron programas de educación relacional adaptados a parejas del mismo sexo y una terapia de pareja adaptada a población LGBTQ+, así como una intervención de prevención del VIH basada en la pareja. En conjunto, se observaron mejoras en variables relacionales y psicológicas (p. ej., comunicación, satisfacción, estrés/distrés) y, según el estudio, beneficios en variables clínicas específicas (p. ej., consumo de alcohol) o reducción del estigma relacional. Ningún estudio evaluó directamente la homofobia interiorizada como desenlace.

La evidencia disponible sugiere beneficios en bienestar relacional y psicológico, pero la heterogeneidad y el riesgo de sesgo global (algunas preocupaciones) limitan conclusiones firmes sobre superioridad de un modelo afirmativo y sobre reducción directa de homofobia interiorizada.

**Palabras clave:** estrés minoritario; parejas LGBTQ+; revisión sistemática; terapia afirmativa; terapia de pareja.

## Abstract

LGBTQ+ couples may face minority stressors related to stigma, which can affect psychological and relational well-being. This systematic review aimed evaluate the effectiveness of affirmative therapy, compared with waiting list or other psychological interventions, on psychological and relational well-being in adult LGBTQ+ couples.

A systematic review was conducted in accordance with PRISMA 2020 and the Cochrane Handbook. The preliminary search strategy was developed in PubMed and subsequently adapted to the databases consulted, including PubMed and EBSCOhost databases (APA PsycInfo, CINAHL with Full Text, MEDLINE Complete, and Psychology and Behavioral Sciences Collection), as well as manual reference searching. Randomized controlled trials involving adult ( $\geq 18$  years) LGBTQ+ couples were included. Risk of bias was assessed using RoB 2, supported by CASPe checklists. Due to substantial clinical and methodological heterogeneity, a narrative synthesis was conducted and no meta-analysis was performed.

Five RCTs published between 2009 and 2025 (United States) were included, comprising 261 couples. Interventions covered culturally adapted relationship education programs for same-sex couples, an LGB-adapted couples therapy, and a couples-based HIV prevention program. Overall, studies reported improvements in key relational and psychological outcomes (e.g., communication, relationship satisfaction, stress/distress), and—depending on the intervention—benefits on specific clinical outcomes (e.g., alcohol use) or reductions in relationship stigma. No study directly assessed internalized homophobia as an outcome.

Available evidence suggests potential benefits for relational and psychological well-being, but heterogeneity and overall “some concerns” risk of bias limit firm conclusions regarding the superiority of affirmative approaches and direct reductions in internalized homophobia.

**Keywords:** affirmative therapy; couple therapy; LGBTQ+ couples; minority stress; systematic review.

## Introducción

Las relaciones de pareja constituyen un aspecto fundamental en la vida adulta y han recibido una amplia atención dentro del campo de la psicología debido a su influencia en el bienestar general. Diversas investigaciones señalan que la calidad de la relación incide de manera directa en la salud emocional, ya que el malestar de pareja se vincula con mayores niveles de ansiedad, depresión y otros problemas asociados. Estudios contemporáneos destacan que la investigación acumulada sobre comunicación, apego y procesos emocionales muestra que los patrones de interacción tienen un papel central en la estabilidad y satisfacción relacional (Lebow y Snyder, 2022). Sin embargo, buena parte de los modelos clásicos de terapia de pareja se desarrollaron dentro de marcos basados en relaciones heterosexuales, lo que generó enfoques centrados en concepciones normativas sobre la vida conyugal (Cabaj y Klinger, 1996). En la actualidad, la evidencia muestra que las intervenciones psicológicas dirigidas a personas LGBTQ+ han sido históricamente limitadas y heterogéneas, lo que ha impulsado la necesidad de adaptar los modelos terapéuticos a las experiencias específicas vinculadas al estrés de las minorías y a los procesos de estigmatización que afectan a esta población (Tudor-Sfetea y Topciu, 2024). En esta misma línea, una revisión sistemática reciente subraya que la salud mental y el funcionamiento relacional no pueden comprenderse de manera aislada, sino como procesos interdependientes donde el afrontamiento en pareja desempeña un papel central, especialmente ante estresores psicológicos (Landolt et al., 2023).

Las parejas LGBTQ+ pueden enfrentar desafíos relacionales similares a los de cualquier vínculo afectivo, pero su dinámica está condicionada por factores sociales que operan de manera diferencial. Souza y Santos (2025) documentan que estas parejas están expuestas a diversas formas de prejuicio, discriminación y segregación, sostenidas por discursos y prácticas inscritas en una lógica heterosexista y heteronormativa, lo que contribuye a negar la legitimidad de sus vínculos y a reforzar dinámicas de invisibilización de

sus formas de relación. Esta estructura sociocultural no solo restringe el reconocimiento social, sino que también incide en los procesos vinculares que sostienen el bienestar de la pareja. En consonancia, la revisión sistemática realizada por De Brito Silva et al. (2021) evidencia que muchas intervenciones dirigidas a parejas lesbianas, gays y bisexuales continúan reproduciendo modelos derivados de la conyugalidad heteronormativa, lo que puede generar experiencias de desajuste o alienación. Los autores subrayan la necesidad de deconstruir estos marcos heterocentros e incorporar enfoques terapéuticos que atiendan explícitamente la negociación de expectativas, la clarificación de acuerdos y el impacto del estrés minoritario en la vida relacional. Desde una perspectiva más amplia, Crenshaw (1989) plantea que las experiencias de opresión no pueden comprenderse mediante categorías aisladas, sino a través de la interacción simultánea de múltiples identidades sociales, lo que genera formas específicas de vulnerabilidad que con frecuencia permanecen invisibilizadas. Así, las vivencias relacionales de las parejas LGBTQ+ deben analizarse considerando cómo la orientación sexual se entrecruza con dimensiones como el género, la raza, la clase social o la discapacidad, configurando niveles diferenciados de exposición al estigma y al estrés.

Según Biaggio et al. (2002), en parejas lesbianas se ha descrito el fenómeno de la fusión emocional o *merger*, entendido como un estilo relacional caracterizado por altos niveles de intimidad, cohesión y búsqueda de cercanía. Aunque históricamente fue interpretado desde marcos heteronormativos como un indicador de disfunción, la literatura aclara que esta intensidad afectiva no debe patologizarse, sino comprenderse dentro del contexto sociocultural, afectivo y de socialización compartida entre mujeres, lo cual puede hacer que altos niveles de intimidad sean vividos como deseables y funcionales para muchas parejas. En el caso de parejas de hombres gays, algunos autores señalan que la socialización masculina —centrada en la autosuficiencia, el control emocional y la evitación de la vulnerabilidad— puede dificultar la expresión de necesidades afectivas y, por tanto, la

construcción de intimidad, aspecto observado en intervenciones terapéuticas en las que estos mandatos culturales interfieren en la conexión emocional de la pareja (Lord, 2017). Además, investigaciones recientes muestran que, en el caso de parejas lesbianas y queer que forman una familia, las interacciones con profesionales de la salud, la educación o los servicios sociales pueden estar marcadas por normas heteronormativas y por el modelo del *monomaternalismo* —concepto desarrollado por Park (2013, citado en Levy y Middleton, 2025)—, lo que sitúa a las co-madres, especialmente a aquellas no gestantes o no biológicas, en posiciones de cuestionamiento o invisibilidad respecto a su identidad materna. Estas tensiones entre la identidad que las co-madres construyen y el modo en que los sistemas profesionales las reconocen se analizan mediante el enfoque de Interpretative Phenomenological Analysis (Smith et al., 2009, citado en Levy y Middleton, 2025), mostrando cómo estos marcos normativos pueden afectar su vivencia relacional y familiar. Todo ello evidencia la necesidad de que terapeutas y profesionales incorporen una comprensión crítica de la heteronormatividad en su trabajo (Levy y Middleton, 2025).

Estas experiencias pueden comprenderse mejor si se considera que, en el trabajo con parejas del mismo sexo, uno de los elementos centrales es la gestión del estrés minoritario, entendido como el conjunto de estresores sociales que afectan de manera diferencial a las personas LGB. Este planteamiento se fundamenta en el modelo de estrés de minoría de Meyer (2003), quien señala que las personas lesbianas, gays y bisexuales están expuestas a estresores sociales crónicos —tanto distales, como la discriminación directa, como proximales, entre ellos la vigilancia social, el ocultamiento de la identidad o la homofobia interiorizada— que incrementan la carga psicológica y afectan su bienestar. Desde esta perspectiva, el malestar no se explica por la orientación sexual en sí misma, sino por el contexto social estigmatizante en el que se desarrolla. En esta línea, Green y Mitchell (2015) señalan que el trabajo terapéutico con parejas del mismo sexo requiere ayudar a los miembros

de la pareja a afrontar el estrés minoritario, reducir la ambigüedad en los compromisos y roles de la relación y fortalecer redes de apoyo social, ya que estos aspectos constituyen áreas clave en las que pueden surgir dificultades específicas. De hecho, investigaciones recientes sugieren que el afrontamiento diádico —esto es, las estrategias compartidas que implementa la pareja para gestionar el estrés— constituye un mecanismo fundamental que influye tanto en la salud mental individual como en el ajuste relacional, independientemente del tipo de diagnóstico psicológico (Landolt et al., 2023).

La homofobia interiorizada, entendida como la internalización de prejuicios sobre la propia orientación sexual, se ha vinculado a dificultades para expresar necesidades emocionales, aumento del conflicto, menor intimidad y menor percepción de apoyo mutuo en la relación (Pepping y Halford, 2014). Además, este fenómeno tiene un impacto diádico: niveles elevados en uno de los miembros pueden generar tensión, retraimiento emocional o conductas defensivas que afectan la dinámica relacional (Hambrook, 2025). En parejas en las que uno de los miembros transita un proceso de transición de género, estos estresores se combinan con desafíos identitarios y sociales adicionales, lo cual requiere enfoques terapéuticos sensibles que acompañen los cambios personales y relacionales implicados (Boe et al., 2019; Samons, 2009).

En este marco, el bienestar relacional puede entenderse como un proceso influido tanto por dinámicas internas de la pareja como por condiciones externas que afectan su funcionamiento cotidiano. En el caso de parejas lesbianas y queer que forman una familia, aspectos como la forma en que los sistemas profesionales reconocen o cuestionan su legitimidad parental, junto con los marcos heteronormativos presentes en los servicios de salud, educación o apoyo social, influyen de manera significativa en su vivencia relacional y en el modo en que construyen y sostienen sus vínculos familiares (Levy y Middleton, 2025). De forma complementaria, investigaciones recientes muestran que el bienestar de las parejas

LGB también se relaciona con su capacidad para manejar los estresores propios de las minorías sexuales. En este sentido, la adaptación del programa Couples Coping Enhancement Training centrado en el estrés minoritario (CCET-SMS) ha mostrado mejoras en el afrontamiento diádico y en la satisfacción relacional, así como una reducción del estrés minoritario percibido y del malestar general, lo que subraya la importancia de intervenciones que integren explícitamente estos factores en la evaluación y el fortalecimiento del bienestar relacional (Randall et al., 2025).

Si bien enfoques como la terapia cognitivo-conductual de pareja, la terapia sistémica o la terapia focalizada en las emociones han mostrado eficacia en diversos contextos, su aplicación a parejas LGBTQ+ puede resultar limitada cuando no incorporan explícitamente los efectos del estigma, la homofobia (interna y externa), y las experiencias particulares vinculadas a la identidad sexual (Cabaj y Klinger, 1996; Pepping y Halford, 2014). Además, elementos como la salida del armario, la gestión de distintos grados de visibilidad, la falta de apoyo familiar o las normas heteronormativas suelen quedar invisibilizados en los modelos clásicos de intervención, lo que restringe el alcance terapéutico (Malley, 2002). Algunos programas de educación relacional de corte heteronormativo han mostrado beneficios moderados, pero estos efectos se reducen al aplicarse a parejas LGBTQ+ sin adaptaciones específicas que consideren el estrés minoritario y las necesidades propias de estas relaciones (Hatch et al., 2021). Incluso enfoques sistémicos y narrativos que han demostrado utilidad mejoran su eficacia cuando incorporan un análisis crítico del estigma y del contexto sociocultural de las parejas LGBTQ+ (Greenan, 2019; Su y Parker, 2022).

Frente a estas limitaciones, la terapia afirmativa se presenta como un enfoque desarrollado para abordar las necesidades específicas de las personas y parejas LGBTQ+, integrando explícitamente el impacto del estigma y de las presiones socioculturales sobre su bienestar. En coherencia con este planteamiento, las *APA Guidelines for Psychological*

*Practice With Sexual Minority Persons* (2021) recomiendan que los profesionales de la salud mental adopten una práctica culturalmente competente, reconozcan la influencia del estigma estructural, validen las identidades sexuales diversas y promuevan intervenciones que atiendan los determinantes sociales del bienestar. Estas directrices subrayan, además, la importancia de evitar enfoques patologizantes y de considerar cómo la orientación sexual interactúa con otras dimensiones identitarias en la experiencia psicológica y relacional. Desde esta perspectiva, Belous (2015) describe la terapia afirmativa como un modelo que valida la diversidad sexual y de género, contextualiza el malestar psicológico dentro de las dinámicas de discriminación y trabaja directamente sobre creencias y actitudes negativas interiorizadas. Esta orientación terapéutica promueve la afirmación identitaria, la reestructuración de creencias vinculadas al estigma y la deconstrucción de discursos heteronormativos que históricamente han patologizado las relaciones no heterosexuales. En complemento, Niki (2017) subraya la importancia de aplicar estos principios afirmativos al trabajo con personas bisexuales, destacando que la invisibilidad y los estigmas específicos asociados al monosexismo requieren intervenciones que reconozcan y validen su identidad, evitando la reproducción de discursos que refuercen su marginación. De este modo, la terapia afirmativa sitúa el origen del sufrimiento relacional no en la orientación sexual o identidad de género de las personas, sino en los contextos sociales que generan y sostienen el estigma.

Estudios recientes han evidenciado los beneficios de integrar el análisis del estrés minoritario en procesos terapéuticos. Las adaptaciones afirmativas del *Relationship Checkup* han mostrado mejoras en áreas como la comunicación y la intimidad cuando incorporan componentes orientados a cuestionar el estigma interiorizado y validar la identidad (Gray et al., 2022). Asimismo, el programa *CCET-SMS*, enfocado específicamente en parejas que experimentan estrés minoritario, ha demostrado efectos positivos en el afrontamiento diádico, la cohesión emocional y la satisfacción en la relación (Randall et al., 2025). Por su parte, en

el ámbito de parejas trans, tanto los enfoques narrativos como los modelos afirmativos han mostrado ser especialmente útiles para acompañar los procesos de transición, facilitar la redefinición de significados y fortalecer el vínculo relacional (Boe et al., 2019; Samons, 2009).

La terapia afirmativa también enfatiza la importancia de comprender a las parejas LGBTQ+ dentro de los múltiples sistemas sociales y culturales que influyen en su experiencia relacional. Tal como señala Malley (2002), el trabajo clínico con personas lesbianas y gays requiere situar las dificultades en el marco de los contextos familiares, culturales y sociales que pueden reproducir dinámicas heterosexistas y limitar el reconocimiento de sus vínculos. De manera complementaria, Greenan (2019) destaca que las intervenciones con hombres gays deben considerar cómo los procesos de socialización y los discursos socioculturales inciden en la expresión emocional, la intimidad y la regulación afectiva dentro de la pareja. Este enfoque contextual resulta fundamental para adaptar la intervención a las realidades específicas de las parejas LGBTQ+ y para comprender cómo el entorno contribuye a su bienestar relacional.

En conjunto, la literatura sugiere que la relación entre estigma social, estrés minoritario, homofobia interiorizada y bienestar relacional opera como un proceso encadenado, en el que las experiencias de estigmatización se internalizan y generan dificultades comunicativas, tensiones afectivas y un menor ajuste diádico. La terapia afirmativa interviene precisamente sobre estos mecanismos, al abordar de forma explícita el estigma interiorizado y fortalecer la resiliencia individual y relacional, lo que explicaría su mayor efectividad frente a modelos no adaptados al trabajo con parejas LGBTQ+. No obstante, a pesar del creciente desarrollo de intervenciones afirmativas, la evidencia empírica sobre su efectividad en comparación con otros enfoques psicológicos sigue siendo limitada y heterogénea, especialmente en el ámbito del trabajo con parejas. Por ello, el presente trabajo

tiene como objetivo analizar de forma sistemática la evidencia disponible sobre la efectividad de la terapia afirmativa en parejas adultas LGBTQ+ en relación con el bienestar psicológico y relacional.

### **Método**

El presente trabajo de revisión sistemática se ha desarrollado bajo las directrices de la Declaración PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*, 2020) y siguiendo las recomendaciones metodológicas del Manual Cochrane de Revisiones Sistemáticas de Intervenciones, partiendo de la pregunta PICO de investigación (Tabla 1):

¿Es la terapia afirmativa una intervención efectiva, en comparación con otras intervenciones psicológicas, para la reducción de la homofobia interiorizada y la mejora del bienestar en la pareja en personas adultas LGBTQ+?

#### **Tabla 1**

##### *Pregunta PICO*

- 
- P Personas adultas ( $\geq 18$  años) que se identifiquen como LGBTQ+ y se encuentren en contextos de pareja (parejas del mismo sexo, parejas de minorías sexuales y de género, parejas diversas).
  
  - I Intervenciones o programas psicológicos basados en la terapia afirmativa dirigidos a población LGBTQ+.
  
  - C No intervención. Lista de espera. Otras intervenciones psicológicas no afirmativas o modelos terapéuticos tradicionales aplicados a población LGBTQ+ en contextos de pareja.
  
  - O Reducción de la homofobia interiorizada y mejora del bienestar psicológico y relacional, incluyendo satisfacción de pareja, funcionamiento relacional y afrontamiento del estrés de minoría.
-

## Criterios de elegibilidad

Los criterios de inclusión que se establecieron para la selección de los estudios son:

1. Personas adultas ( $\geq 18$  años) que se identifiquen como LGBTQ+ y se encuentren en contextos de pareja (parejas del mismo sexo, parejas de minorías sexuales y de género, parejas diversas).
2. Personas que se identifiquen como LGBTQ+ (lesbianas, gays, bisexuales, personas transgénero, queer, intersex, otras minorías sexuales y de género).
3. Participantes en contextos de pareja (parejas del mismo sexo, parejas de minorías sexuales y de género, parejas diversas).
4. Estudios que incluyan explícitamente intervenciones o programas dirigidos a personas LGBTQ+ con medición de resultados psicológicos o relacionales (bienestar, satisfacción de pareja, homofobia interiorizada, estrés de minoría).
5. Estudios publicados en revistas revisadas por pares.

Se establecieron, además, los siguientes criterios de exclusión:

1. Menores de 18 años.
2. Parejas heterosexuales cisgénero (sin pertenencia al colectivo LGBTQ+).
3. Estudios centrados en personas LGBTQ+ sin relación de pareja (por ejemplo, intervenciones individuales sin componente de pareja).
4. Estudios en los que la población LGBTQ+ aparece de manera marginal o no diferenciada (ej. estudios generales de población donde no se reportan resultados específicos).
5. Estudios que incluyan únicamente 'terapia de conversión (*conversion therapy* / *reparative therapy*), por razones éticas y por no ser comparables a intervenciones afirmativas.

## **Fuentes de información**

La búsqueda bibliográfica se realizó en bases de datos electrónicas relevantes para el ámbito de la salud y la psicología. Concretamente, se llevó a cabo una búsqueda en PubMed. Asimismo, se realizaron búsquedas en distintas bases de datos accesibles a través de la plataforma EBSCOhost, incluyendo APA PsycInfo, CINAHL with Full Text, MEDLINE Complete y Psychology and Behavioral Sciences Collection.

Adicionalmente, se utilizaron recursos institucionales mediante el Centro de Recursos para el Aprendizaje y la Investigación (CRAI) de la Universidad Europea de Canarias, que facilitó el acceso a diversas bases de datos científicas y recursos bibliográficos especializados. Las búsquedas en las bases de datos se realizaron el 28 de octubre de 2025. La última actualización de la búsqueda bibliográfica se efectuó en esa misma fecha.

## **Estrategia de búsqueda**

Para identificar estudios relevantes relacionados con intervenciones psicológicas en contextos de pareja en población adulta LGBTQ+, se diseñó una estrategia de búsqueda estructurada en bloques conceptuales correspondientes a la población de interés, el contexto relacional y el tipo de intervención, de acuerdo con la pregunta de investigación.

En relación con la población, se incluyeron términos referidos a minorías sexuales y de género. Para el contexto relacional, se incorporaron términos relacionados con la pareja y la relación. Asimismo, se utilizaron términos vinculados a intervenciones psicológicas aplicadas en contextos de pareja.

Los términos dentro de cada bloque conceptual se combinaron mediante el operador booleano OR, mientras que los distintos bloques se combinaron mediante AND, con el objetivo de maximizar la sensibilidad de la búsqueda y recuperar estudios que integraran simultáneamente la población de interés, el contexto de pareja y la intervención psicológica.

Los resultados (*outcomes*) de interés se definieron previamente en la pregunta de investigación y en los criterios de inclusión, pero no se incluyeron como términos de búsqueda con el fin de evitar la exclusión de estudios relevantes que evaluaran dichos desenlaces sin mencionarlos explícitamente en el título o resumen, siguiendo las recomendaciones del Manual Cochrane.

La estrategia de búsqueda preliminar se diseñó inicialmente para PubMed y se aplicó posteriormente, con los ajustes necesarios, en las demás bases de datos y plataformas consultadas. La estrategia final completa utilizada en las bases de datos consultadas se presenta de forma detallada en el Apéndice A.

Se aplicaron límites de idioma (español, inglés y portugués), población adulta y tipo de publicación (artículos científicos indexados en bases de datos; se excluyeron tesis doctorales y resúmenes de congresos). No se establecieron límites temporales ni geográficos.

Por último, además de las búsquedas electrónicas, se realizó una búsqueda manual mediante la revisión de las listas de referencias de los estudios incluidos. A través de este procedimiento se identificaron dos registros adicionales, correspondientes a revisiones sistemáticas relevantes sobre el tema, que no habían sido recuperadas mediante las bases de datos electrónicas.

### **Proceso de selección de los estudios**

El proceso de selección de los estudios se llevó a cabo en dos fases. En una primera fase, los registros identificados a partir de la búsqueda bibliográfica fueron importados a la plataforma Rayyan, una herramienta diseñada para apoyar el cribado de estudios en revisiones sistemáticas (Ouzzani et al., 2016). En esta fase, se identificaron y eliminaron los registros duplicados. Posteriormente, dos revisoras realizaron de forma independiente el cribado de títulos y resúmenes, seleccionando o descartando los registros en función de los

criterios de inclusión y exclusión previamente establecidos, con el objetivo de minimizar el riesgo de sesgo en el proceso de selección.

En una segunda fase, los artículos considerados potencialmente elegibles fueron evaluados a texto completo. Esta evaluación fue realizada de forma independiente por ambas revisoras. Los estudios seleccionados por al menos una de ellas o aquellos considerados dudosos fueron revisados conjuntamente y discutidos para determinar su inclusión final en la revisión, de acuerdo con los criterios de elegibilidad definidos. Las discrepancias se resolvieron mediante discusión hasta alcanzar consenso. Los motivos de exclusión de los estudios evaluados a texto completo fueron documentados.

El proceso de selección de los estudios se documentó siguiendo las directrices de la declaración PRISMA 2020 (Page et al., 2021).

### **Proceso de extracción de los datos**

El proceso de extracción de datos se realizó a partir de una hoja de extracción previamente definida. Una revisora llevó a cabo la extracción de forma sistemática tras la selección final de los estudios incluidos y una segunda revisora revisó la consistencia de los datos extraídos. Las discrepancias o dudas se resolvieron mediante discusión y consenso.

Los desenlaces fueron definidos a priori y se agruparon en desenlaces psicológicos y relacionales:

#### **Desenlaces psicológicos:**

- Homofobia interiorizada
- Distrés o malestar psicológico.
- Estrés percibido.
- Regulación emocional.
- Consumo de sustancias.

**Desenlaces relacionales:**

- Comunicación de pareja (positiva y negativa; observada y autoinformada).
- Afrontamiento diádico.
- Apoyo social percibido.
- Confianza relacional.
- Metas compartidas o congruencia de metas.
- Estigma relacional.
- Patrones de conflicto (como implicación en el conflicto, resolución positiva de problemas y retirada comunicativa).
- Inestabilidad relacional.
- Satisfacción relacional global y ajuste diádico.

Asimismo, se recogieron desenlaces secundarios adicionales cuando fueron reportados en los estudios incluidos (p. ej., vulnerabilidad al VIH/ITS), siempre que se analizaran en relación con variables psicológicas o relacionales y de acuerdo con los criterios de elegibilidad establecidos.

**Lista de datos**

Se buscaron datos correspondientes a los desenlaces definidos *a priori* en la pregunta de investigación, que fueron organizados en dominios psicológicos y relacionales para facilitar su extracción y síntesis.

En el dominio psicológico, se incluyeron desenlaces relacionados con el bienestar psicológico y la salud mental de las personas participantes, tales como bienestar psicológico general, salud mental, estrés de minoría, homofobia interiorizada y sintomatología depresiva y/o ansiosa cuando fue evaluada como desenlace secundario. En el dominio relacional, se buscaron datos relativos al funcionamiento de la relación de pareja, incluyendo satisfacción de pareja, ajuste diádico, comunicación, funcionamiento relacional, cohesión y apoyo

relacional. Asimismo, se incluyeron desenlaces relacionados con el proceso terapéutico cuando fueron evaluados en los estudios, como el estigma relacional.

Para cada desenlace, se extrajeron todos los resultados compatibles disponibles en los estudios incluidos, independientemente del instrumento de medida utilizado, siempre que el desenlace evaluado se correspondiera conceptualmente con los dominios definidos. Cuando un mismo desenlace fue evaluado mediante múltiples escalas o en diferentes puntos temporales, se priorizaron las medidas postintervención y las escalas principales reportadas por los autores, atendiendo a criterios de relevancia clínica y consistencia entre estudios.

Además de los desenlaces, se extrajeron datos sobre otras variables relevantes, incluyendo las características generales de los estudios (autoría, año de publicación, país y diseño del estudio), características de los/as participantes (tamaño muestral, tipo de pareja y población) y características de la intervención y del comparador (tipo, enfoque, formato y duración).

En los casos en que la información relevante no estuvo claramente disponible o fue incompleta, se utilizaron los datos reportados en los textos completos y materiales suplementarios de los estudios incluidos. No se realizaron imputaciones ni estimaciones de datos ausentes, y los resultados se interpretaron de forma descriptiva cuando la información fue incompleta o incierta.

### **Evaluación del riesgo de sesgo de los estudios individuales**

La evaluación del riesgo de sesgo y de la calidad metodológica de los estudios incluidos (apéndices B, C, D, E y F) se realizó mediante la herramienta RoB 2 de la Colaboración Cochrane y las listas de comprobación del programa CASPe (*Critical Appraisal Skills Programme*), adaptadas al diseño de los ensayos clínicos aleatorizados incluidos en la revisión (CASPe, 2018). Esta herramienta se basa en los principios de la lectura crítica de ensayos clínicos aleatorizados y permite valorar de forma estructurada

aspectos clave relacionados con la validez interna de los estudios, tales como la claridad de la pregunta de investigación, el proceso de aleatorización, la comparabilidad de los grupos, el manejo de las pérdidas durante el seguimiento, la medición de los desenlaces y la comunicación de los resultados (Cabello López et al., 2022).

La evaluación fue realizada por una revisora y posteriormente cotejada de forma independiente por otra revisora en cada uno de los estudios incluidos. Las discrepancias se resolvieron mediante discusión y consenso. No se utilizaron herramientas de automatización en el proceso de evaluación del riesgo de sesgo.

### **Medidas del efecto**

Para cada uno de los desenlaces incluidos en la revisión, las medidas del efecto se extrajeron y presentaron de acuerdo con el tipo de variable evaluada y con la información disponible en los estudios originales. Dado que los desenlaces evaluados fueron mayoritariamente continuos y se midieron mediante instrumentos psicométricos validados, las medidas del efecto se expresaron principalmente como diferencias de medias entre los grupos de intervención y control, acompañadas de sus correspondientes intervalos de confianza del 95 %, cuando esta información fue reportada por los autores.

En aquellos estudios en los que los resultados se presentaron mediante tamaños del efecto estandarizados, se recogieron los valores de  $d$  de Cohen u otras medidas equivalentes, junto con sus estimaciones de precisión, con el fin de facilitar la comparación de la magnitud de los efectos entre estudios que utilizaron escalas diferentes para un mismo dominio de desenlace. Cuando los tamaños del efecto no fueron reportados explícitamente, los resultados se describieron a partir de las estadísticas de resumen disponibles (medias, desviaciones estándar, valores  $t$ , valores  $p$  o coeficientes de regresión reportados), sin realizar estimaciones adicionales ni cálculos derivados.

Para los desenlaces dicotómicos, cuando estuvieron disponibles, se recogieron las medidas de asociación reportadas por los estudios originales (por ejemplo, proporciones o porcentajes), sin calcular razones de riesgos u odds ratios adicionales, debido a la heterogeneidad de los diseños y a la limitada comparabilidad entre estudios.

Dado que no se realizó un metanálisis, las medidas del efecto se utilizaron exclusivamente con fines descriptivos y comparativos, y no se combinaron cuantitativamente. Los resultados se interpretaron atendiendo a la magnitud, dirección y consistencia de los efectos observados entre estudios, considerando además el riesgo de sesgo y la calidad metodológica de cada uno de ellos.

### **Método de síntesis**

El proceso de síntesis de los resultados se llevó a cabo siguiendo las recomendaciones metodológicas del Manual Cochrane para revisiones sistemáticas en las que no se realiza metanálisis. En primer lugar, se determinó la elegibilidad de los estudios para cada síntesis en función de su comparabilidad clínica y metodológica, considerando el tipo de intervención, la población incluida, los desenlaces evaluados y los instrumentos de medida utilizados.

Dada la elevada heterogeneidad clínica y metodológica entre los estudios incluidos —en relación con las intervenciones aplicadas, las características de las poblaciones, los desenlaces evaluados y los puntos temporales de medición—, no fue apropiado realizar una síntesis cuantitativa mediante metanálisis. En consecuencia, se optó por una síntesis narrativa estructurada, orientada a integrar y comparar los resultados de manera sistemática y transparente.

Los estudios se agruparon para la síntesis en función de dominios de desenlace previamente definidos, diferenciando entre desenlaces psicológicos y desenlaces relacionales. Dentro de cada dominio, los resultados se organizaron atendiendo al tipo de intervención y al

desenlace específico evaluado, describiendo la dirección, magnitud y consistencia de los efectos observados en los distintos estudios.

Para la preparación de los datos, se utilizaron exclusivamente las estadísticas de resumen reportadas en los estudios originales. No se realizaron conversiones de escalas ni imputaciones de datos ausentes. En los casos en que un mismo desenlace fue evaluado mediante diferentes instrumentos o en distintos momentos temporales, se priorizaron las medidas postintervención y las escalas principales identificadas por los autores, de acuerdo con criterios de relevancia clínica y consistencia entre estudios.

Los resultados se presentaron mediante una combinación de síntesis narrativa y tablas estructuradas, que permitieron resumir de forma clara las características de los estudios, los desenlaces evaluados y los principales hallazgos. Esta estrategia facilitó la comparación entre estudios sin introducir una combinación cuantitativa inapropiada de los resultados. Las decisiones metodológicas adoptadas se justificaron atendiendo a la heterogeneidad observada y al objetivo de proporcionar una síntesis rigurosa y clínicamente interpretable de la evidencia disponible.

## **Resultados**

### **Selección de los estudios**

El proceso de búsqueda y selección de los estudios se describió de acuerdo con las recomendaciones de la declaración PRISMA 2020 y se presenta de forma gráfica en el diagrama de flujo correspondiente (ver Figura 1). En total, se identificaron 1.100 registros a través de las búsquedas realizadas en bases de datos electrónicas (*records identified*, n = 1.100). Adicionalmente, se identificaron dos registros mediante búsqueda manual en las listas de referencias (*records identified via other methods*, n = 2). No se identificaron registros procedentes de registros de estudios ni se eliminaron duplicados.

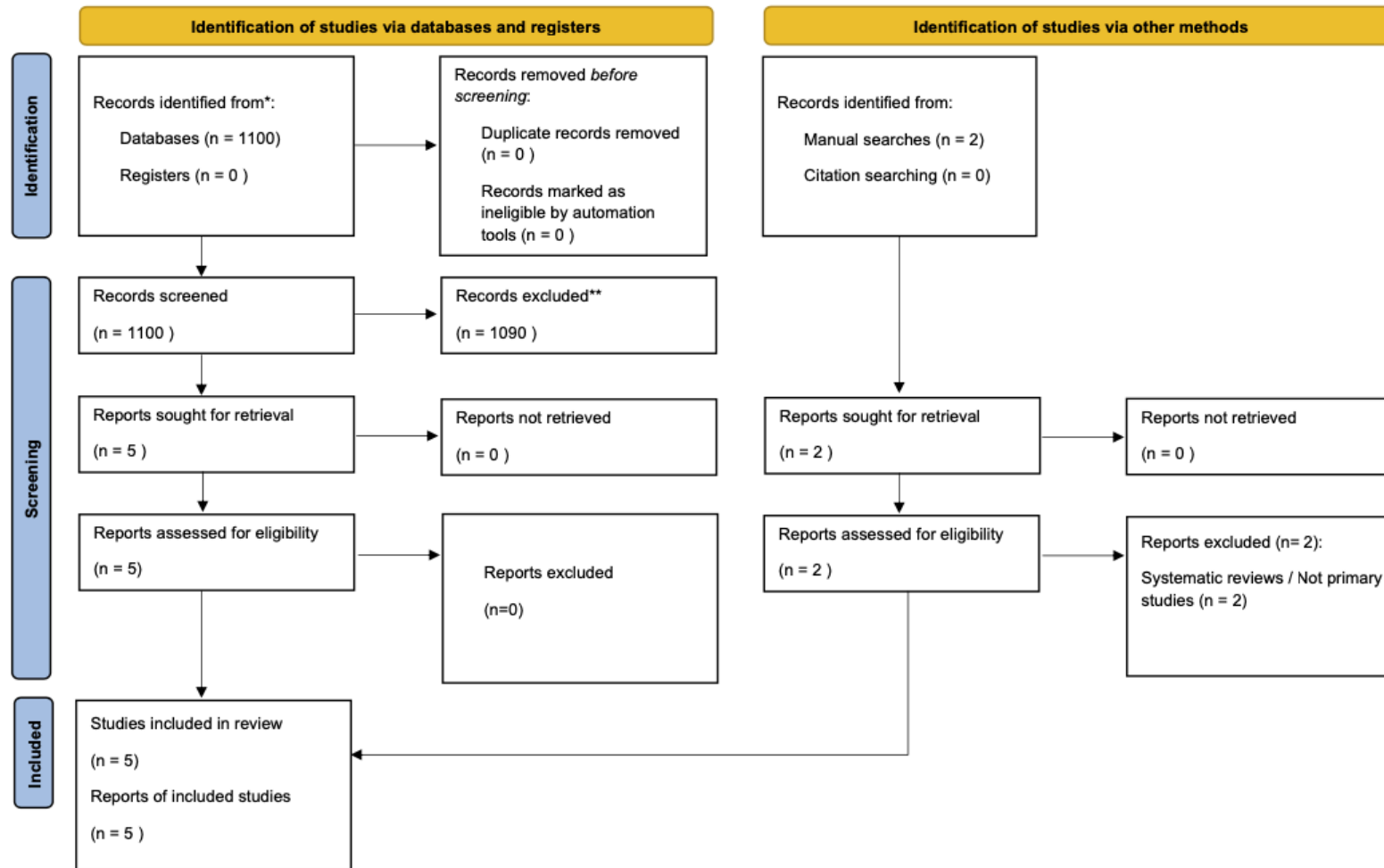
Tras el cribado por título y resumen de los 1.100 registros identificados (*records screened*, n = 1.100), se excluyeron 1.090 registros por no cumplir los criterios de inclusión establecidos (*records excluded*, n = 1.090). Como resultado de este proceso, cinco informes procedentes de las bases de datos fueron seleccionados para su recuperación y evaluación a texto completo (*reports sought for retrieval*, n = 5; *reports assessed for eligibility*, n = 5).

Los cinco informes evaluados a texto completo cumplieron los criterios de elegibilidad y fueron incluidos en la revisión sistemática, sin que se produjeran exclusiones tras la evaluación a texto completo (*reports excluded*, n = 0).

Asimismo, los dos informes identificados mediante búsqueda manual fueron evaluados a texto completo (*reports assessed for eligibility*, n = 2) y excluidos por tratarse de revisiones sistemáticas y no de estudios primarios (*reports excluded*, n = 2) (de Brito Silva et al., 2022; Kousteni y Anagnostopoulos, 2019). Finalmente, se incluyeron cinco estudios, cada uno representado por un único informe, en la presente revisión sistemática (*studies included in review*, n = 5).

Figura 1

*Proceso de selección de estudios*



## Características de los estudios incluidos

Los cinco estudios incluidos en la presente revisión sistemática fueron publicados entre 2009 y 2025 y se realizaron exclusivamente en Estados Unidos. Todos emplearon un diseño experimental, concretamente ensayos controlados aleatorizados (ECA), incluyendo un ensayo piloto con grupo en lista de espera. En conjunto, los estudios analizaron una muestra total de 261 parejas pertenecientes a poblaciones LGBTQ+, que incluyeron parejas del mismo sexo (lesbianas y gays), parejas de mujeres, parejas de hombres gays y mujeres transgénero y sus parejas. El tamaño muestral de los estudios osciló entre 20 y 100 parejas, lo que refleja una notable variabilidad en el alcance de las intervenciones evaluadas. Las características detalladas de los estudios incluidos se detallan en la Tabla 2 y se describen a continuación.

En relación con la población intervenida y el tipo de intervención aplicada, dos de los estudios realizados con parejas del mismo sexo se orientaron a la promoción de habilidades relacionales y al fortalecimiento de la relación de pareja mediante programas de educación relacional culturalmente adaptados. En concreto, Within Our Reach (WOR) es un programa de educación relacional basado en el *Prevention and Relationship Education Program* (PREP), centrado en el fortalecimiento de habilidades de pareja como la mejora de la comunicación, la gestión del estrés y la promoción de relaciones saludables, con especial énfasis en la seguridad relacional a nivel físico, emocional, de compromiso y comunitario. Este programa fue aplicado a parejas del mismo sexo, incluyendo parejas lesbianas y gays (Barden et al., 2021). Asimismo, el programa *Strengthening Same-Sex Relationships* (SSSR) es una intervención de educación relacional culturalmente sensible adaptada específicamente para parejas de hombres gays, orientada al fortalecimiento del vínculo y de las habilidades relacionales (Whitton et al., 2016).

Un estudio se dirigió exclusivamente a parejas del mismo sexo compuestas por mujeres, en relaciones comprometidas de al menos seis meses, evaluando la versión femenina del programa *Strengthening Same-Sex Relationships* (SSSR-F). Este programa consistió en una intervención grupal de educación relacional, culturalmente sensible, con una duración total de 10 horas distribuidas en 11 unidades, centrada en el desarrollo de habilidades de comunicación, el afrontamiento del estrés minoritario, la negociación de expectativas y el fortalecimiento del vínculo relacional (Whitton, 2017).

Por su parte, un estudio incluyó parejas del mismo sexo con un miembro de la pareja con trastorno por uso de alcohol, en el que se evaluó una intervención terapéutica específica, la *Behavioral Couples Therapy* (BCT) adaptada a población LGB. Esta intervención incorporó diversos componentes orientados a apoyar la recuperación y el funcionamiento relacional, incluyendo un *Recovery Contract*, consistente en un acuerdo de sobriedad supervisado por la pareja; entrenamiento en comunicación positiva y expresividad emocional, junto con la reducción de interacciones negativas; el aumento de actividades gratificantes compartidas; el establecimiento de rutinas diarias que apoyan la recuperación; la resolución de conflictos orientada a la abstinencia; y el apoyo mutuo en el mantenimiento de los cambios. La intervención de pareja se comparó con un tratamiento individual sin participación de la pareja (Fals-Stewart et al., 2009).

Asimismo, un estudio se centró en mujeres transgénero y sus parejas, evaluando el programa *It Takes Two*, una intervención de prevención del VIH basada en la pareja, orientada a mejorar la comunicación, promover la congruencia de metas y reducir el estigma relacional, con un enfoque en la salud sexual y relacional (Gamarel et al., 2025).

Respecto a las condiciones comparadoras, tres estudios utilizaron grupos en lista de espera (Barden et al., 2021; Whitton et al., 2016; Whitton et al., 2017), un estudio empleó vídeos educativos como condición control (Gamarel et al., 2025), y un estudio comparó la

intervención de pareja con un tratamiento individual sin participación de la pareja (Fals-Stewart et al., 2009).

La duración, intensidad y formato de las intervenciones variaron entre estudios y, en varios casos, el intervalo temporal del grupo control no coincidió con el periodo de intervención, sino con un tiempo de espera o una condición alternativa. En Barden et al. (2021), el grupo de intervención recibió Within Our Reach (WOR) (12 horas) a lo largo de 4 semanas, comenzando entre 1 y 2 semanas tras el enrolamiento, y la recogida de datos se realizó en la evaluación inicial y tras completar la intervención ( $\approx 30$  días post-aleatorización). El grupo control fue asignado a una lista de espera y recibió una adaptación abreviada de la intervención tras completar las medidas de resultado a los 6 meses.

En Whitton et al. (2016), el programa *Strengthening Same-Sex Relationships* (SSSR) (10 horas) se administró típicamente a lo largo de 3 semanas, con una sesión inicial intensiva de 6 horas en fin de semana seguida de dos sesiones de 2 horas entre semana. La condición control fue una lista de espera, con tratamiento diferido tras un periodo de espera de 4 semanas. Las evaluaciones se realizaron en línea base y a las 4 semanas (postprograma para el grupo de intervención inmediata y post-waitlist para el grupo control), y se incluyó además una evaluación de seguimiento a los 3 meses postprograma.

De forma equivalente, en Whitton (2017) la versión femenina del programa *Strengthening Same-Sex Relationships* (SSSR-F) (10 horas) se impartió a lo largo de 3 semanas, con una sesión inicial intensiva de 6 horas en fin de semana seguida de dos sesiones de 2 horas entre semana. El estudio utilizó un grupo control en lista de espera, con tratamiento diferido tras un periodo de espera de 4 semanas, y las evaluaciones se realizaron en línea base y a las 4 semanas, con una evaluación de seguimiento a los 3 meses postprograma.

En el estudio clínico de Fals-Stewart et al. (2009), ambas condiciones presentaron la misma duración total del tratamiento, consistente en 32 sesiones programadas de 60 minutos distribuidas a lo largo de 20 semanas. La diferencia entre condiciones residió en la estructura de las sesiones y en la participación de la pareja. En la condición de intervención, *Behavioral Couples Therapy* (BCT), el tratamiento combinó 12 sesiones de terapia de pareja, a las que asistieron conjuntamente ambos miembros de la pareja, con 20 sesiones individuales realizadas únicamente por el paciente. En contraste, en la condición control, *Individual-Based Therapy* (IBT), el tratamiento consistió exclusivamente en 32 sesiones individuales, sin participación de la pareja en ninguna fase del tratamiento.

En Gamarel et al. (2025), el programa It Takes Two se administró mediante cuatro sesiones de intervención basadas en la pareja, de aproximadamente 45 minutos cada una. El estudio incluyó un grupo control que recibió materiales educativos en formato vídeo. Debido a la pandemia de COVID-19, parte de las sesiones se realizaron en modalidad digital (telehealth). Las evaluaciones se llevaron a cabo en línea base y en los seguimientos a los 3, 6, 9 y 12 meses postintervención.

Los desenlaces evaluados abarcaron variables relacionales y psicológicas, incluyendo comunicación de pareja, ajuste diádico, afrontamiento diádico, regulación emocional, estrés percibido, distrés psicológico, apoyo social, satisfacción relacional, confianza e inestabilidad de pareja, así como variables específicas de salud, entre ellas consumo de alcohol, vulnerabilidad al VIH, conductas sexuales de riesgo, incidencia de infecciones de transmisión sexual, congruencia de metas y estigma relacional.

Para la evaluación de los desenlaces se emplearon instrumentos validados de autoinforme y medidas observacionales. En concreto, se utilizaron el *Dyadic Adjustment Scale* (DAS; Spanier, 1976), la *Revised Dyadic Adjustment Scale* (RDAS; Busby et al., 1995) y la *Relationship Satisfaction Measure* (Whitton et al., 2008); el *Dyadic Coping Inventory*

(DCI; Randall et al., 2016); la *Perceived Stress Scale* (PSS; Cohen et al., 1983) y el *Outcome Questionnaire-45* (OQ-45; Lambert et al., 2004); la *Difficulties in Emotion Regulation Scale* (DERS; Gratz y Roemer, 2004); la escala adaptada de *Support for the Relationship* (Sprecher y Felmlee, 1992); la *Goal Congruence Scale* (Gamarel y Golub, 2020); la *Confidence Scale* (Stanley et al., 1994) y el *Relationship Instability Index* (Whitton et al., 2015); el *Timeline Follow-Back* (TLFB; Sobell y Sobell, 1996); la *Relationship Stigma Scale* (Gamarel et al., 2014); la *Conflict Resolution Style Scale* (Kurdek, 1994); el *Interactional Dimensions Coding System* (IDCS; Kline et al., 2004) y las subescalas del *Communication Skills Test* (CST; Jenkins y Saiz, 1995); así como un índice compuesto de vulnerabilidad VIH/ITS basado en conducta sexual, uso de PrEP y supresión viral (Gamarel et al., 2019).

Tabla 2

*Características de los estudios incluidos*

Autor(es) y años	País	Diseño	Muestra	Tipo de pareja	Intervención	Comparador	Duración/ Formato	Outcomes evaluados	Instrumento
Barden et al. (2021)	EEUU	ECA con grupo en lista de espera	52 parejas del mismo sexo (104 participantes): 30 parejas lesbianas y 22 parejas gays	Parejas del mismo sexo (lesbianas y gays)	<b>Within Our Reach (WOR)</b> es un programa de educación relacional basado en Prevention and Relationship Education Program. (PREP) que se centra en el fortalecimiento de habilidades de pareja, incluyendo la mejora de la comunicación, la gestión del estrés y la promoción de relaciones saludables, con énfasis en la seguridad relacional física, emocional, de compromiso y comunitaria.	Lista de espera (waitlist)	Intervención grupal WOR de 12 h impartida en 4 semanas	Afrontamiento diádico Regulación emocional Distrés psicológico Ajuste diádico	Dyadic Coping Inventory (DCI) Difficulties in Emotion Regulation Scale (DERS) Outcome Questionnaire-45.2 (OQ-45.2) Dyadic adjustment: Revised Dyadic Adjustment Scale (RDAS)
Fals-Stewart, O'Farrell y Lam (2009)	EEUU	ECA	100 parejas (52 de hombres gays; 48 mujeres lesbianas) con un miembro de la pareja con trastorno por uso de alcohol	Parejas del mismo sexo con un miembro de la pareja con trastorno por uso de alcohol	Behavioral Couples Therapy (BCT) adaptada a población LGB. Incluye: (a) <i>Recovery Contract</i> : acuerdo de sobriedad supervisado por la pareja; (b) entrenamiento en comunicación positiva, expresividad emocional y reducción de interacciones negativas; (c) aumento de actividades gratificantes compartidas; (d) establecimiento de rutinas diarias que apoyan la recuperación; (e) resolución de conflictos orientada a la abstinencia; (f) apoyo mutuo en mantenimiento de cambios.	Tratamiento <b>individual</b> de facilitación de 12 pasos, sin participación de la pareja.	<b>BCT + IBT</b> : 12 sesiones de pareja (BCT) + 20 sesiones individuales (IBT), <b>32 sesiones</b> de 60 min durante <b>20 semanas</b> . <b>IBT</b> : 32 sesiones <b>individuales</b> de 60 min durante <b>20 semanas</b> (sin participación de la pareja).	Consumo de alcohol (PDHD). Ajuste diádico (DAS).	<b>Alcohol use:</b> <i>Timeline Followback Interview (TLFB)</i>  <i>Dyadic Adjustment Scale (DAS)</i>

Autor(es) y años	País	Diseño	Muestra	Tipo de pareja	Intervención	Comparador	Duración/ Formato	Outcomes evaluados	Instrumento
Gamarel et al. (2025)	EEUU	ECA	52 parejas (104 personas): mujeres trans y sus parejas	Mujeres transgénero y sus parejas	“It Takes Two”, programa de prevención del VIH basado en la pareja, centrado en mejorar la comunicación, la congruencia de metas y reducir el estigma relacional	Videos educativos	Cuatro sesiones de asesoramiento de 1 hora; modalidad presencial o digital ( <i>telehealth</i> )	Vulnerabilidad al VIH.  Conducta sexual de riesgo  Incidencia de ITS  Estigma relacional.  Congruencia de metas.  Comunicación de pareja.	Composite HIV Vulnerability Index  Self-reported condomless anal/vaginal sex (main and outside partners) Self-reported STI diagnosis (past 3 months) <i>Relationship Stigma Scale</i> <i>7-item Goal Congruence Scale</i> <i>Conflict Resolution Style Scale</i>
Whitton, Weitbrecht, Kuryluk y Hutsell (2016)	EEUU	ECA con lista de espera	20 parejas (40 participantes)	Parejas de hombres gays	<i>Strengthening Same-Sex Relationships</i> (SSSR), programa de educación relacional culturalmente sensible adaptado para parejas masculinas del mismo sexo	Lista de espera	Programa SSSR refinado de 10 horas, impartido generalmente en 3 semanas: una sesión inicial de 6 h en fin de semana y dos sesiones posteriores de 2 h entre semana; un grupo lo recibió en dos sesiones de fin de semana (6 h + 4 h) por motivos logísticos.	Comunicación de pareja  Estrés percibido  Apoyo social  Satisfacción de pareja  Inestabilidad de pareja	<i>Communication Patterns Questionnaire (CPQ)</i> (autoinforme)  <i>Interaccional Dimensions Coding System (IDCS)</i> (observacional)  <i>Perceived Stress Scale (PSS)</i>  <i>Adapted Social Support Scale</i>  <i>Relationship Assessment Scale (RAS)</i>  <i>Relationship Instability Index (RII)</i>

Autor(es) y años	País	Diseño	Muestra	Tipo de pareja	Intervención	Comparador	Duración/ Formato	Outcomes evaluados	Instrumento
Whitton (2017)	EEUU	Ensayo controlado aleatorizado o piloto con grupo en lista de espera (randomized waitlist-control trial)	37 parejas de mujeres (74 participantes)	Parejas del mismo sexo (mujeres) en relación comprometida de al menos 6 meses	<b>Strengthening Same-Sex Relationships Program</b> – Female version (SSSR-F). Programa grupal de educación relacional, culturalmente sensible, de 10 horas (11 unidades), centrado en habilidades de comunicación, afrontamiento del estrés minoritario, negociación de expectativas y fortalecimiento del vínculo.	Grupo en lista de espera (waitlist control)	<b>Programa grupal de 10 horas</b> , impartido en aproximadamente <b>2,5-3 semanas</b> (una sesión intensiva de fin de semana de 6 h y dos sesiones posteriores de 2 h entre semana)	Comunicación de pareja (positiva y negativa; observacional y autoinformada)  Estrés percibido  Apoyo social  Satisfacción relacional  Confianza en la relación  Inestabilidad relacional	Interaccional Dimensions Coding System (IDCS)  Communication Skills Test (CST)  Perceived Stress Scale (PSS)  Support for the Relationship Measure (SRM, versión adaptada)  Relationship Satisfaction Scale (adaptación de McNulty & Karney)  Confidence Scale (CS)  Relationship Instability Index (RII)

## Riesgo de sesgo de los estudios individuales

La evaluación del riesgo de sesgo se realizó siguiendo las recomendaciones del Manual Cochrane mediante la herramienta RoB 2 para ensayos aleatorizados. A continuación, se describen los principales riesgos de sesgos identificados mediante la herramienta RoB 2 y se resumen en la Tabla 3.

**Tabla 3**

*Evaluación del riesgo de sesgo (herramienta RoB 2)*

Estudio	D1. Proceso de Aleatorización	D2. Desviaciones de la Intervención	D3. Datos de Resultado Perdidos	D4. Medición del Desenlace	D5. Selección del Resultado Informado	Riesgo Global
Barden et al., 2021	Algunas preocupaciones	Algunas preocupaciones	Bajo riesgo	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones
Fals-Stewart et al., 2009	Algunas preocupaciones	Bajo riesgo	Bajo riesgo	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones
Gamarel et al., 2025	Algunas preocupaciones	Bajo riesgo	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones
Whitton et al., 2016	Algunas preocupaciones	Algunas preocupaciones	Alto riesgo	Algunas preocupaciones	Algunas preocupaciones	Alto riesgo
Whitton et al., 2017	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones	Algunas preocupaciones

En relación con el proceso de aleatorización, en cuatro de los estudios incluidos (Barden et al., 2021; Fals-Stewart et al., 2009; Whitton et al., 2016; Whitton et al., 2017) se informa la asignación aleatoria de los participantes, pero no se describe con suficiente detalle el método técnico de generación de la secuencia aleatoria ni los procedimientos de ocultación de la asignación. No obstante, en todos ellos los grupos de intervención y control fueron comparables en las características basales, sin desequilibrios clínicamente relevantes, lo que sugiere que el proceso de asignación fue adecuado, aunque insuficientemente documentado. En consecuencia, este dominio se clasificó como “algunas preocupaciones” en estos estudios.

Por su parte, el estudio de Gamarel et al. (2025) describe de forma clara el proceso de aleatorización, indicando la asignación de las parejas en una proporción 1:1 mediante un generador de asignación aleatoria con estratificación por seroestatus VIH. Sin embargo,

aunque la asignación se realizó tras la evaluación basal, no se informa de manera explícita sobre la existencia de procedimientos formales de ocultación de la asignación, por lo que este dominio también se clasificó como “algunas preocupaciones”.

En cuanto a las desviaciones de la intervención asignada, ninguno de los estudios informó desviaciones sistemáticas ni cruces entre condiciones tras la aleatorización. En Fals-Stewart et al. (2009) y Gamarel et al. (2025), los análisis se realizaron siguiendo un enfoque de intención de tratar o intención de tratar modificada mediante modelos longitudinales apropiados, lo que permitió estimar adecuadamente el efecto de la asignación a la intervención. Por este motivo, este dominio se consideró de bajo riesgo de sesgo en ambos estudios.

En Barden et al. (2021), Whitton et al. (2016) y Whitton et al. (2017), aunque no se identificaron desviaciones relevantes de la intervención tras la aleatorización, la ausencia de un enfoque explícito de intención de tratar y la información limitada sobre el posible impacto del conocimiento de la asignación en la conducta de los participantes impiden descartar completamente desviaciones diferenciales no controladas. En consecuencia, este dominio se clasificó como “algunas preocupaciones” en estos estudios.

Respecto a los datos de resultado perdidos, el manejo de la información faltante fue adecuado en Barden et al. (2021) y Fals-Stewart et al. (2009). En ambos estudios se emplearon métodos estadísticos apropiados para datos longitudinales, como la imputación múltiple multinivel y la estimación por máxima verosimilitud con información completa (FIML), sin observarse pérdidas diferenciales relevantes ni identificarse evidencia clara de que los datos faltantes dependieran del valor verdadero del desenlace. En consecuencia, este dominio se consideró de bajo riesgo de sesgo en ambos casos.

En Gamarel et al. (2025), aunque las tasas de retención fueron aceptables a lo largo del seguimiento, no se evaluó explícitamente si la falta de datos dependía del valor del desenlace, lo que llevó a clasificar este dominio como “algunas preocupaciones”. De forma similar, en Whitton et al. (2017), el manejo de los datos perdidos fue adecuado mediante modelos multinivel, pero la ausencia de análisis específicos sobre la posible dependencia del valor verdadero del desenlace justificó igualmente un juicio de “algunas preocupaciones”.

Por el contrario, en Whitton et al. (2016) se observaron pérdidas sustanciales, parcialmente diferenciales y relacionadas con el desenlace (ruptura de la relación), sin análisis de sensibilidad que permitieran descartar sesgo, por lo que este dominio se clasificó como de alto riesgo de sesgo.

En relación con la medición del desenlace, todos los estudios emplearon instrumentos validados y procedimientos de evaluación estandarizados, aplicados de forma idéntica en los grupos de intervención y control. No obstante, la mayoría de los desenlaces fueron de naturaleza subjetiva y autoinformada, y ninguno de los estudios informó cegamiento de los participantes ni de los evaluadores. La combinación de desenlaces subjetivos y ausencia de cegamiento plantea un riesgo plausible de sesgo de medición, aunque no se identificó evidencia clara de medición diferencial entre grupos. Por este motivo, este dominio se clasificó como “algunas preocupaciones” en todos los estudios.

Finalmente, en lo relativo a la selección del resultado informado, ninguno de los estudios incluidos informó un registro prospectivo del protocolo ni un plan de análisis estadístico preespecificado. Aunque en todos los casos los desenlaces descritos en la sección de métodos fueron reportados en los resultados, incluyendo resultados significativos y no significativos, la ausencia de preregistro impide descartar completamente la posibilidad de

selección entre múltiples desenlaces o análisis elegibles. En consecuencia, este dominio se clasificó como “algunas preocupaciones” en todos los estudios.

De acuerdo con la herramienta RoB 2 del Manual Cochrane, el juicio global del riesgo de sesgo se determinó a partir del dominio con el mayor nivel de riesgo identificado en cada estudio. Así, cuatro de los estudios incluidos (Barden et al., 2021; Fals-Stewart et al., 2009; Gamarel et al., 2025; Whitton et al., 2017) presentaron un riesgo global de sesgo clasificado como “algunas preocupaciones”, mientras que el estudio de Whitton et al. (2016) se clasificó como de alto riesgo de sesgo, debido al alto riesgo identificado en el dominio de datos de resultado perdidos.

De forma complementaria, se utilizaron las listas de verificación CASPe para extraer y contrastar la información metodológica reportada en los estudios, lo que permitió informar y afinar las respuestas a las *signalling questions* y los juicios emitidos en cada uno de los dominios evaluados mediante la herramienta RoB 2, sin sustituirla.

## **Resultados de los estudios individuales**

Los resultados de los estudios individuales se organizaron en función de dos grandes dominios de desenlace: desenlaces psicológicos y desenlaces relacionales, de acuerdo con los objetivos de la revisión y con la naturaleza de las variables evaluadas en los estudios incluidos. En conjunto, los estudios mostraron una notable heterogeneidad en función del tipo de intervención afirmativa aplicada, el dominio evaluado y los instrumentos empleados para la medición de los resultados (Tabla 4).

### **Desenlaces psicológicos**

En el ámbito psicológico, los resultados fueron variables y dependientes del desenlace específico evaluado. En relación con el distrés psicológico, Barden et al. (2021) identificaron

una reducción significativamente mayor en el grupo de intervención en comparación con el grupo control en el postratamiento, con un tamaño del efecto moderado.

Respecto al estrés percibido, los hallazgos fueron heterogéneos. Whitton et al. (2016) observaron reducciones significativas del estrés percibido en el grupo de intervención frente a la lista de espera, con efectos de magnitud moderada mantenidos durante el seguimiento. Sin embargo, en un estudio posterior del mismo equipo (Whitton et al., 2017), aunque se observó una reducción mayor en el grupo de intervención, las diferencias entre condiciones no alcanzaron significación estadística.

En cuanto a la regulación emocional, Barden et al. (2021) no identificaron diferencias estadísticamente significativas entre los grupos de intervención y control en dificultades de regulación emocional en el postratamiento, observándose tamaños del efecto pequeños.

Finalmente, en relación con desenlaces clínicos específicos, como el consumo elevado de alcohol, los estudios de Fals-Stewart et al. (2009) mostraron resultados comparables entre los grupos en el pretratamiento y el postratamiento. No obstante, durante el seguimiento a 12 meses, las parejas que recibieron terapia conductual de pareja presentaron porcentajes significativamente menores de días de consumo elevado en comparación con el tratamiento individual, tanto en parejas gays como lesbianas.

### **Desenlaces relacionales**

En el dominio relacional, los resultados fueron más consistentes, especialmente en variables relacionadas con la comunicación de pareja. Los programas de educación relacional evaluados por Whitton et al. (2016, 2017) mostraron reducciones significativas de la comunicación negativa, tanto observada como autoinformada, en los grupos de intervención frente a los controles, con tamaños del efecto que oscilaron entre pequeños-moderados y grandes según la medida utilizada. Asimismo, se observaron mejoras significativas en la comunicación positiva observada en los grupos de intervención en comparación con los

controles en ambos estudios, con tamaños del efecto moderados. En contraste, los resultados en comunicación positiva autoinformada fueron más heterogéneos, identificándose en un estudio ausencia de significación estadística y en otro, diferencias significativas entre condiciones.

En relación con el afrontamiento diádico y el apoyo social percibido, no se identificaron diferencias estadísticamente significativas entre los grupos de intervención y control, observándose puntuaciones comparables y tamaños del efecto pequeños en los estudios disponibles (Barden et al., 2021; Whitton et al., 2016, 2017).

En cuanto a las metas compartidas, el estigma relacional y los patrones de conflicto (implicación en el conflicto, resolución positiva de problemas y retirada comunicativa), el estudio de Gamarel et al. (2025) no identificó diferencias claras entre las condiciones a lo largo del seguimiento, manteniéndose valores comparables entre los grupos de intervención y control en todos los momentos evaluados.

Respecto a la confianza relacional, Whitton et al. (2017) observaron un incremento significativo en el grupo de intervención en comparación con el grupo control tras el tratamiento, con un tamaño del efecto pequeño a moderado. En relación con la inestabilidad relacional, no se identificaron diferencias estadísticamente significativas entre condiciones en el postratamiento en los estudios de Whitton et al. (2016, 2017), aunque se observaron reducciones intra-grupo de pequeña magnitud en el grupo de intervención a lo largo del tiempo.

Finalmente, en relación con la satisfacción relacional global y el ajuste diádico, los resultados fueron variables. Mientras que Fals-Stewart et al. (2009) y Whitton et al. (2016, 2017) observaron mejoras significativas en la satisfacción y/o el ajuste relacional en los grupos de intervención frente a los controles, Barden et al. (2021) no identificó diferencias

estadísticamente significativas entre condiciones en el ajuste diádico global tras la intervención.

De forma adicional, Gamarel et al. (2025) incluyeron un indicador compuesto de vulnerabilidad VIH/ITS; aunque no fue un desenlace primario de esta revisión, se observó una disminución descriptiva en el grupo de intervención frente al control durante el seguimiento.

Tabla 4

## Resultados de los estudios individuales

Estudio (año), país	Desenlace	Instrumento	Pre-intervención		Post-intervención		Resultados
			Intervención	Control	Intervención	Control	
<b>DESENLACES PSICOLÓGICOS</b>							
<b>Distrés psicológico</b>							
Barden et al. (2021), EE. UU.	Psychological distress	Outcome Questionnaire-45, OQ-45 (Lambert et al., 2004)	-	-	36.74	45.38	La intervención produjo una reducción significativamente mayor del distrés psicológico en comparación con el grupo control en el postratamiento, con puntuaciones inferiores en el grupo de intervención (36.74 vs. 45.38), respaldadas por los modelos multivariados ajustados por línea base ( $b = -0.24$ , $EE = 0.12$ , $p < .05$ ; $d = -0.51$ ; IC 95 % $-0.91$ a $-0.11$ ).
<b>Estrés percibido</b>							
Whitton et al. (2016), EE. UU.	Perceived stress	Perceived Stress Scale (PSS; Cohen, Kamarck, & Mermelstein, 1983)	Preprogram (full sample): 1.53 (.12)	Preprogram (full sample): 1.53 (.12)	1.23 (SE = 0.09)	1.51 (SE = 0.10)	La intervención produjo una reducción significativa del estrés percibido en comparación con la lista de espera (postintervención: $M = 1.23$ vs. $1.51$ ; $t = -2.15$ , $p < .05$ ; $d = 0.46$ ), con efectos de magnitud moderada mantenidos en el seguimiento (pre-post $d = 0.41$ ; pre-seguimiento $d = 0.45$ ).
Whitton et al. (2017), EE. UU.			1.74 (DE = 0.67)	1.66 (DE = 0.64)	1.58 (SE = 0.17)	1.72 (SE = 0.04)	En la preintervención, los niveles de estrés percibido fueron comparables entre grupos (intervención: $M = 1.74$ ; $DE = 0.67$ ; control: $M = 1.66$ ; $DE = 0.64$ ). En el postratamiento, el grupo intervención mostró una reducción mayor ( $M = 1.58$ ) que el grupo control ( $M = 1.72$ ); no obstante, la diferencia postintervención no alcanzó significación estadística ( $z = -0.98$ , $p = .33$ , $d = -0.20$ ).
<b>Regulación emocional</b>							
Barden et al. (2021), EE. UU.	Difficulties in emotion regulation	Difficulties in Emotion Regulation Scale (DERS) (Gratz & Roemer, 2004)	-	-	71.209	72.677	No se observaron diferencias estadísticamente significativas entre la intervención y el control en dificultades de regulación emocional en el postratamiento, con medias similares entre grupos (71.209 vs. 72.677) y un tamaño del efecto pequeño ( $d = -0.12$ ; IC 95 % $-0.51$ a $0.28$ ).
<b>Consumo elevado de alcohol</b>							
Fals-Stewart et al. (2009), EE. UU. (parejas gays)	Percent Days Heavy Drinking (PDHD)	Timeline Follow-Back (TLFB) (Sobell & Sobell, 1996)	BCT: 41.94 (18.73)	IBT: 43.84 (21.63)	BCT: 5.99 (13.62)	IBT: 5.30 (14.91)	El porcentaje de días de consumo elevado fue similar entre grupos en el pretratamiento (41.94 vs. 43.84) y en el postratamiento (5.99 vs. 5.30); sin embargo, durante el seguimiento el grupo BCT presentó valores inferiores al grupo IBT (12 meses: 18.00 vs. 32.16), según comparaciones pareadas ( $ps < .05$ ).

Estudio (año),	Desenlace	Instrumento	Pre-intervención		Post-intervención		Resultados
pals-Stewart et al. (2009), EE. UU. (parejas lesbianas)			BCT: 38.62 (16.44)	IBT: 39.80 (19.74)	BCT: 5.07 (14.12)	IBT: 5.27 (14.09)	Las puntuaciones fueron comparables entre grupos en el pretratamiento (38.62 vs. 39.80) y en el postratamiento (5.07 vs. 5.27), mientras que durante el seguimiento el porcentaje de días de consumo elevado fue menor en el grupo BCT que en IBT (12 meses: 15.71 vs. 27.92), según comparaciones pareadas (ps < .05).
<b>DESENLACES RELACIONALES</b>							
<b>Comunicación negativa</b>							
Whitton et al. (2016), EE. UU.	Observed negative communication	Interactional Dimensions Coding System (Kline et al., 2004).	Preprogram (full sample): 2.68 (.20)	Preprogram (full sample): 2.68 (.20)	1.75 (SE = 0.09)	2.58 (SE = 0.10)	La intervención produjo una reducción significativa de la comunicación negativa observada en comparación con la lista de espera (postintervención: M = 1.75 vs. 2.58; t = -6.09, p < .01; d = 1.08), con mejoras mantenidas desde la preintervención al seguimiento (pre-post d = 0.71; pre-seguimiento d = 0.37).
Whitton et al. (2017), EE. UU.			2.31 (DE = 0.95)	2.50 (DE = 0.99)	2.05 (SE = 0.17)	2.52 (SE = 0.21)	En la preintervención, ambos grupos mostraron niveles similares de comunicación negativa observada (intervención: M = 2.31; DE = 0.95; control: M = 2.50; DE = 0.99). Tras la intervención, el grupo intervención presentó una reducción de la comunicación negativa (M = 2.05), mientras que el grupo control permaneció estable (M = 2.52). La diferencia postintervención fue significativa (z = -2.02, p = .04, d = -0.55).
Whitton et al. (2016), EE. UU.	Self-Reported negative communication	Subscales of Communication Skills Test (CST; Jenkins & Saiz, 1995)	Preprogram (full sample): 2.18 (.16)	Preprogram (full sample): 2.18 (.16)	1.78 (.10)	2.12 (.11)	La intervención se asoció con reducciones significativas de la comunicación negativa autoinformada en comparación con la lista de espera (postintervención: M = 1.78 vs. 2.12; t = -2.39, p < .05; d = 0.41), con efectos pequeños a moderados a lo largo del seguimiento (pre-post d = 0.48; pre-seguimiento d = 0.30).
Whitton et al. (2017), EE. UU.			2.86 (DE = 1.34)	2.92 (DE = 1.50)	2.30 (SE = 0.24)	2.92 (SE = 0.29)	En la preintervención, ambos grupos mostraron niveles similares de comunicación negativa autoinformada (intervención: M = 2.86; DE = 1.34; control: M = 2.92; DE = 1.50). Tras la intervención, el grupo intervención presentó una reducción notable (M = 2.30), mientras que el grupo control no mostró cambios (M = 2.92). La diferencia postintervención fue significativa (z = -2.95, p = .003, d = -0.49).
<b>Comunicación positiva</b>							
Whitton et al. (2016), EE. UU.	Observed positive communication	Interactional Dimensions Coding System (Kline et al., 2004).	Preprogram (full sample): 5.57 (.25)	Preprogram (full sample): 5.57 (.25)	6.60 (SE = 0.17)	5.61 (SE = 0.20)	Se observó un aumento significativo de la comunicación positiva observada en el grupo de intervención frente a la lista de espera (postintervención: M = 6.60 vs. 5.61; t = 3.76, p < .01; d = 0.83), con cambios de magnitud moderada desde la preintervención al postratamiento y al seguimiento (pre-post d = 0.67; pre-seguimiento d = 0.38).

Estudio (año),	Desenlace	Instrumento	Pre-intervención		Post-intervención		Resultados
Whitton et al. (2017), EE. UU.			5.70 (DE = 1.21)	5.24 (DE = 1.46)	6.23 (SE = 0.33)	5.22 (SE = 0.33)	En la línea base, ambos grupos presentaron niveles comparables de comunicación positiva observada (intervención: M = 5.70; DE = 1.21; control: M = 5.24; DE = 1.46). En el postratamiento, el grupo intervención mostró un aumento claro (M = 6.23), frente a la estabilidad del grupo control (M = 5.22). La diferencia entre grupos fue significativa (z = 2.26, p = .02, d = 0.78).
Whitton et al. (2016), EE. UU.	Self-Reported positive communication	Subscales of Communication Skills Test (CST; Jenkins & Saiz, 1995)	Preprogram (full sample): 5.36 (.13)	Preprogram (full sample): 5.36 (.13)	5.63 (.08)	5.42 (.08)	No se observaron diferencias estadísticamente significativas entre condiciones en comunicación positiva autoinformada (postintervención: M = 5.63 vs. 5.42; t = 1.90, p < .10; d = 0.32), aunque se identificaron cambios pequeños desde la preintervención al seguimiento (pre-post d = 0.39; pre-seguimiento d = 0.36).
Whitton et al. (2017), EE. UU.			4.65 (DE = 0.94)	4.58 (DE = 0.99)	5.14 (SE = 0.16)	4.67 (SE = 0.14)	En la preintervención, los niveles de comunicación positiva autoinformada fueron similares entre grupos (intervención: M = 4.65; DE = 0.94; control: M = 4.58; DE = 0.99). En el postratamiento, el grupo intervención mostró un incremento (M = 5.14), mientras que el grupo control apenas varió (M = 4.67). La diferencia postintervención fue significativa (z = 3.06, p = .002, d = 0.54).
<b>Afrontamiento diádico</b>							
Barden et al. (2021), EE. UU.	Dyadic Coping	Dyadic Coping Inventory (DCI) (Randall et al., 2016)	-	-	133.074	129.158	No se identificaron diferencias significativas entre condiciones en afrontamiento diádico tras la intervención, con puntuaciones similares en los grupos de intervención y control (133.074 vs. 129.158), y un tamaño del efecto pequeño y no significativo (d = 0.23; IC 95 % -0.17 a 0.63).
<b>Apoyo social percibido</b>							
Whitton et al. (2016), EE. UU.	Social support	Scale adapted from Support for Relationship Measure (SRM) (Sprecher & Felmlee, 1992)	Preprogram (full sample): 5.66 (.17)	Preprogram (full sample): 5.66 (.17)	5.66 (.13)	5.71 (.14)	No se observaron efectos significativos de la intervención sobre el apoyo social (postintervención: M = 5.66 vs. 5.71; t = 0.29; d = 0.06), con cambios longitudinales pequeños (pre-post d = 0.12; pre-seguimiento d = 0.30).
Whitton et al. (2017), EE. UU.			5.31 (DE = 0.99)	5.23 (DE = 1.01)	5.35 (SE = 0.15)	5.19 (SE = 0.16)	En la preintervención, ambos grupos presentaron niveles similares de apoyo social percibido (intervención: M = 5.31; DE = 0.99; control: M = 5.23; DE = 1.01). Tras la intervención, las puntuaciones permanecieron estables en ambos grupos (intervención: M = 5.35; control: M = 5.19), sin diferencias significativas entre condiciones (p = .17; d = 0.18).
<b>Metas compartidas</b>							
Gamarel et al. (2025), EE. UU.	Goal congruence	Goal Congruence Scale (Gamarel. & Golub, 2020)	4.07 (DE = 0.94)	4.20 (DE=0.78)	4.44 (DE= 0.70)	4.24 (DE= 0.80)	Las puntuaciones de congruencia de metas fueron similares entre los grupos en la línea base (4.20 vs. 4.07) y se mantuvieron comparables durante el seguimiento, sin diferencias claras entre condiciones (12 meses: 4.44 vs. 4.44).

Estudio (año),	Desenlace	Instrumento	Pre-intervención	Post-intervención	Resultados		
<b>Estigma relacional</b>							
Gamarel et al. (2025), EE. UU.	Relationship Stigma	Relationship Stigma Scale (Gamarel et al., 2014)	1.20 (30) 0.40	1.36 (26) 0.66	1.43 (23) 0.93	1.36 (22) 0.77	En la línea base, las puntuaciones de estigma relacional fueron similares entre el grupo de intervención y el control (1.36 vs. 1.20). Durante el seguimiento, los valores permanecieron cercanos entre condiciones, sin diferencias evidentes en las medias descriptivas (12 meses: 1.43 vs. 1.36).
<b>Implicación en el conflicto (comunicación)</b>							
Gamarel et al. (2025), EE. UU.	Conflict Engagement	Conflict Resolution Style Scale (Kurdek, 1994)	3.98 (DE = 0.92)	4.10 (DE = 0.85)	4.18 (DE = 0.77)	4.28 (DE = 0.70)	En la línea base, las puntuaciones de implicación en el conflicto fueron similares entre los grupos de intervención y control (3.98 vs. 4.10). A lo largo del seguimiento, ambos grupos mostraron valores comparables y trayectorias paralelas, sin diferencias claras entre condiciones (por ejemplo, 12 meses: 4.18 vs. 4.28).
<b>Resolución positiva de problemas (comunicación)</b>							
Gamarel et al. (2025), EE. UU.	Positive problem-solving communication	Conflict Resolution Style Scale (Kurdek, 1994)	4.00 (DE = 0.76)	3.99 (DE = 0.69)	4.17 (DE = 0.69)	4.14 (DE = 0.75)	Las puntuaciones basales de resolución positiva de problemas fueron prácticamente idénticas entre grupos (4.00 vs. 3.99). Durante el seguimiento, ambos grupos mostraron incrementos similares y valores comparables, sin observarse diferencias atribuibles a la intervención (12 meses: 4.14 vs. 4.17).
<b>Retirada comunicativa</b>							
Gamarel et al. (2025), EE. UU.	Withdrawal	Conflict Resolution Style Scale (Kurdek, 1994)	3.61 (DE = 0.83)	3.75 (DE = 0.73)	4.12 (DE = 0.76)	4.02 (DE = 0.77)	En la preintervención, las puntuaciones de retirada comunicativa fueron similares entre grupos (3.75 vs. 3.61). A lo largo del seguimiento, ambos grupos presentaron valores cercanos y sin un patrón diferencial consistente (12 meses: 4.12 vs. 4.02).
<b>Confianza relacional</b>							
Whitton et al. (2017), EE. UU.	Relationship confidence	Confidence Scale (Stanley, Hoyer, & Trathen, 1994)	6.15 (1.02)	6.15 (1.04)	6.49 (SE = 0.14)	6.13 (SE = 0.20)	En la preintervención, ambos grupos presentaron niveles idénticos de confianza relacional (intervención: M = 6.15; DE = 1.02; control: M = 6.15; DE = 1.04). En el postratamiento, el grupo intervención mostró un incremento de la confianza (M = 6.49), frente a una ligera disminución en el grupo control (M = 6.13). Esta diferencia postintervención fue significativa ( $z = 2.22$ , $p = .03$ , $d = 0.40$ ).
<b>Inestabilidad relacional</b>							
Whitton et al. (2016), EE.UU.	Relationship instability	Relationship Instability Index (Whitton, Kuryluk, & Khaddouma, 2015)	Preprogram (full sample): 2.90 (.66)	Preprogram (full sample): 2.90 (.66)	2.51 (SE = 0.28)	2.85 (SE = 0.31)	No se encontraron diferencias significativas entre condiciones en inestabilidad relacional (postintervención: M = 2.51 vs. 2.85; $t = 0.76$ ; $d = 0.10$ ), aunque se observaron reducciones dentro del grupo de intervención a lo largo del tiempo (pre-post $d = 0.14$ ; pre-seguimiento $d = 0.31$ ).

Estudio (año),	Desenlace	Instrumento	Pre-intervención		Post-intervención		Resultados
Whitton et al. (2017), EE.UU.	Relationship instability		3.61 (2.47)	3.72 (2.43)	3.09 (SE = 0.43)	3.66 (SE = 0.41)	En la preintervención, los niveles de inestabilidad relacional fueron comparables entre grupos (intervención: M = 3.61; DE = 2.47; control: M = 3.72; DE = 2.43). Tras la intervención, el grupo intervención mostró una reducción de la inestabilidad (M = 3.09), mientras que el grupo control apenas varió (M = 3.66). Sin embargo, la diferencia postintervención no alcanzó significación estadística ( $z = -1.55$ , $p = .12$ , $d = -0.23$ ).
<b>Satisfacción relacional global/ajuste diádico</b>							
Fals-Stewart et al. (2009), EE.UU. (Gay couples – Patient)	Global Relationship Satisfaction	Dyadic Adjustment Scale (DAS) (Spanier, 1976)	BCT: 88.23 (22.91)	IBT: 86.84 (23.11)	BCT: 119.37 (12.66)	IBT: 110.41 (14.25)	Las puntuaciones de ajuste relacional global fueron similares entre grupos en el pretratamiento (88.23 vs. 86.84), mientras que en el postratamiento el grupo BCT mostró valores superiores al grupo IBT (119.37 vs. 110.41), según comparaciones pareadas.
Fals-Stewart et al. (2009), EE.UU. (Gay couples – Partner)			BCT: 80.20 (24.31)	IBT: 79.36 (22.81)	BCT: 109.03 (15.91)	IBT: 101.39 (13.62)	No se observaron diferencias entre grupos en las puntuaciones basales de ajuste relacional (80.20 vs. 79.36), mientras que tras la intervención el grupo BCT presentó mayores puntuaciones que IBT (109.03 vs. 101.39), según comparaciones pareadas.
Fals-Stewart et al. (2009), EE.UU. (Lesbian couples – Patient)			BCT: 92.73 (20.41)	IBT: 93.20 (23.11)	BCT: 111.37 (12.66)	IBT: 103.22 (15.23)	Las puntuaciones de ajuste relacional global fueron comparables entre grupos en el pretratamiento (92.73 vs. 93.20), mientras que en el postratamiento se observaron mayores valores en el grupo BCT frente a IBT (111.37 vs. 103.22), según comparaciones pareadas.
Fals-Stewart et al. (2009), EE.UU. (Lesbian couples – Partner)			BCT: 86.27 (23.63)	IBT: 87.24 (22.81)	BCT: 106.03 (15.91)	IBT: 99.98 (16.52)	No se observaron diferencias basales entre grupos (86.27 vs. 87.24), mientras que tras la intervención el grupo BCT mostró mayores niveles de ajuste relacional que IBT (106.03 vs. 99.98), según comparaciones pareadas.
Whitton et al. (2016), EE.UU.	Relationship Satisfaction	Relationship Satisfaction Measure (Whitton, Stanley, Makrman & Baucom, 2008)	6.01 (SE = 0.04)	6.01 (SE = 0.04)	6.01 (SE = 0.04)	5.86 (SE = 0.05)	Se observaron mejoras significativas en la satisfacción relacional en el grupo de intervención frente a la lista de espera (postintervención: M = 6.01 vs. 5.86; $t = 2.42$ , $p < .05$ ; $d = 0.27$ ), con efectos pequeños desde la preintervención al seguimiento (pre–post $d = 0.19$ ; pre–seguimiento $d = 0.18$ ).

Estudio (año),	Desenlace	Instrumento	Pre-intervención		Post-intervención		Resultados
Winstan et al. (2017), EE. UU.			5.67 (1.02)	5.64 (0.94)	5.94 (SE = 0.12)	5.65 (SE = 0.16)	En la preintervención, ambos grupos mostraron niveles muy similares de satisfacción relacional (intervención: M = 5.67; DE = 1.02; control: M = 5.64; DE = 0.94). Tras la intervención, el grupo intervención presentó un aumento de la satisfacción relacional (M = 5.94), mientras que el grupo control permaneció prácticamente estable (M = 5.65). La diferencia postintervención entre grupos fue estadísticamente significativa ( $z = 3.10$ , $p = .002$ , $d = 0.36$ ), lo que indica un efecto positivo atribuible a la intervención.
Barden et al. (2021), EE. UU.	Dyadic adjustment	Revised dyadic adjustment scale (RDAS) (Busby et al., 1995)	-	-	49.406	49.232	No se observaron diferencias estadísticamente significativas entre la condición de intervención y el grupo control en el ajuste relacional global en el postratamiento, con valores prácticamente equivalentes entre grupos (49.406 vs. 49.232), de acuerdo con los modelos multivariados ajustados por línea base ( $d = 0.13$ ; IC 95 % -0.30 a 0.56).
<b>Vulnerabilidad VIH/ITS</b>							
Gamarel et al. (2025), EE. UU.	HIV vulnerability / STI outcomes	Composite HIV/STI vulnerability index based on self-reported sexual behavior, PrEP use, and viral suppression (Gamarel et al., 2019)	59.5 (25/42)	56.8 (25/44)	43.6 (17/39)	58.3 (28/48)	En la línea base, la proporción de participantes con vulnerabilidad al VIH fue similar entre el grupo de intervención y el control (59.5 % vs. 56.8 %). Durante el seguimiento, la proporción disminuyó en el grupo de intervención mientras se mantuvo más elevada en el grupo control (12 meses: 43.6 % vs. 58.3 %).

## Resultados de la síntesis

Debido a la heterogeneidad clínica y metodológica entre los cinco estudios incluidos, en términos de tipo de intervención, poblaciones, desenlaces evaluados e instrumentos de medida, no fue posible realizar un metaanálisis. En su lugar, se llevó a cabo una síntesis narrativa de los resultados, organizada por dominios de desenlace psicológicos y relacionales.

Las síntesis incluyeron un número reducido de estudios por dominio, predominantemente ensayos controlados aleatorizados, realizados en contextos comunitarios y clínicos en población LGBTIQ+. Los estudios presentaron tamaños muestrales variables y emplearon instrumentos heterogéneos para evaluar desenlaces conceptualmente similares, especialmente en los dominios de comunicación, ajuste relacional y bienestar psicológico, lo que limitó la comparabilidad directa entre estudios.

De acuerdo con la herramienta RoB 2 del Manual Cochrane, el juicio global del riesgo de sesgo se determinó a partir del dominio con el mayor nivel de riesgo identificado en cada estudio. En este sentido, cuatro de los estudios incluidos presentaron un riesgo global de sesgo clasificado como “algunas preocupaciones”, mientras que el estudio de Whitton et al. (2016) se clasificó como de alto riesgo de sesgo, debido al elevado riesgo identificado en el dominio de datos de resultado perdidos (pérdidas sustanciales relacionadas con el desenlace principal y sin análisis de sensibilidad).

Las principales limitaciones metodológicas estuvieron relacionadas con la imposibilidad de cegamiento de participantes y personal, así como con el uso frecuente de medidas autoinformadas, lo que introduce un riesgo plausible de sesgo de medición. No obstante, la mayoría de los estudios empleó procedimientos adecuados de aleatorización y utilizó instrumentos validados para la evaluación de los desenlaces.

En relación con los patrones generales de los resultados, las síntesis mostraron efectos consistentes a favor de la intervención en los dominios de comunicación negativa, comunicación positiva observada, confianza relacional, distrés psicológico y satisfacción o ajuste relacional en algunos estudios. Por el contrario, no se identificaron efectos claros en afrontamiento diádico, apoyo social percibido, metas compartidas, patrones de conflicto, estigma relacional autoinformado por las mujeres trans ni inestabilidad relacional, observándose resultados comparables entre los grupos de intervención y control. En el caso del consumo elevado de alcohol, los efectos favorables a la intervención se observaron principalmente durante el seguimiento, pero no de forma inmediata tras el tratamiento. En conjunto, los efectos favorables se concentraron en desenlaces relacionales y psicológicos directamente vinculados a la interacción de pareja y a procesos comunicativos observables, mientras que los desenlaces más distales o estructurales mostraron menor sensibilidad al cambio a corto plazo.

La heterogeneidad observada entre los resultados puede explicarse por diferencias en el contenido y la intensidad de las intervenciones, la duración del seguimiento, los instrumentos utilizados y el informante de los desenlaces (medidas observacionales frente a autoinformadas). Estas diferencias justificaron la elección de una síntesis narrativa frente a una síntesis cuantitativa.

No se realizaron análisis de sensibilidad, dado que no se llevaron a cabo síntesis estadísticas cuantitativas ni metaanálisis.

### **Sesgos en la publicación**

El riesgo de sesgo debido a resultados faltantes derivados de los sesgos en la publicación se evaluó de forma cualitativa, dado el número reducido de estudios incluidos. Con un total de cinco estudios, no fue posible aplicar métodos gráficos ni pruebas estadísticas para la detección del sesgo de publicación, de acuerdo con las recomendaciones del Manual

Cochrane. El proceso de búsqueda y selección reflejado en el diagrama PRISMA muestra que todos los informes identificados y potencialmente elegibles fueron recuperados y evaluados a texto completo, sin pérdidas de estudios por falta de acceso. No obstante, no puede descartarse completamente la presencia de sesgo de publicación, ya que no se identificaron estudios a través de registros de ensayos ni literatura gris, y es posible que estudios con resultados nulos no hayan sido publicados. En conjunto, el riesgo de sesgo debido a resultados faltantes se considera incierto y debe tenerse en cuenta al interpretar los resultados de las síntesis, aunque no se identificaron indicios claros de sesgo de publicación a partir del proceso de selección de estudios.

### **Certeza de la evidencia**

En conjunto, la evidencia disponible sugiere una tendencia general favorable a las intervenciones culturalmente adaptadas dirigidas a parejas LGBTQ+ en determinados desenlaces relacionales y psicológicos próximos, especialmente en variables de comunicación de pareja, confianza relacional, distrés psicológico y satisfacción o ajuste relacional en algunos estudios. No obstante, estos efectos no se observaron de forma consistente en todos los desenlaces evaluados y, para variables más distales —como afrontamiento diádico, apoyo social percibido, estigma relacional autoinformado o inestabilidad de la relación— los resultados fueron heterogéneos o no mostraron diferencias claras entre grupos.

La interpretación de estos hallazgos debe realizarse con cautela a la luz del riesgo de sesgo. Cuatro de los estudios presentaron un riesgo global de sesgo clasificado como algunas preocupaciones y, de acuerdo con la herramienta RoB 2, la dirección prevista del sesgo a nivel global fue impredecible, al no identificarse mecanismos que favorecieran sistemáticamente a uno de los grupos. En contraste, uno de los estudios mostró un riesgo global de sesgo alto, debido a pérdidas sustanciales de datos relacionadas con el desenlace

principal y a la ausencia de análisis de sensibilidad, lo que podría haber favorecido la estimación de los efectos de la intervención. Este aspecto limita la confianza en los resultados de dicho estudio y refuerza la necesidad de una interpretación prudente.

Además, la evidencia se ve condicionada por el reducido número de estudios, tamaños muestrales moderados y diversidad en los tipos de intervención, comparadores y tiempos de seguimiento, así como por el uso predominante de medidas autoinformadas en contextos donde el cegamiento no fue posible. Estos factores contribuyen a una valoración global limitada de la solidez de la evidencia, especialmente para desenlaces secundarios o más distales.

En síntesis, los estudios incluidos apuntan en su mayoría hacia una dirección favorable de los efectos de las intervenciones adaptadas, particularmente en procesos relacionales cercanos, aunque la evidencia disponible es heterogénea y metodológicamente limitada. Por ello, los resultados apoyan la plausibilidad y el potencial beneficio de este tipo de intervenciones, pero no permiten establecer conclusiones firmes ni generalizables, siendo necesarios ensayos adicionales de mayor tamaño y rigor metodológico para confirmar la dirección y magnitud de los efectos observados.

## **Discusión**

La presente revisión sistemática tuvo como objetivo evaluar la efectividad de la terapia afirmativa, en comparación con otras intervenciones psicológicas, en la reducción de la homofobia interiorizada y la mejora del bienestar psicológico y relacional en personas adultas LGBTQ+ en contextos de pareja. En conjunto, los resultados obtenidos muestran que las intervenciones psicológicas culturalmente adaptadas y sensibles a las experiencias específicas de las minorías sexuales y de género se asocian con mejoras observadas

principalmente en determinadas variables relacionales y psicológicas relevantes, como la comunicación de pareja, la satisfacción relacional, el afrontamiento del estrés y la reducción del distrés psicológico; no obstante, la evidencia disponible presenta una notable heterogeneidad clínica y metodológica, lo que limita la posibilidad de establecer conclusiones firmes sobre la superioridad de un modelo terapéutico específico (Barden et al., 2021; Fals-Stewart et al., 2009; Gamarel et al., 2025; Whitton et al., 2016, 2017).

Los estudios incluidos indican que los programas de educación relacional adaptados a parejas del mismo sexo, como *Strengthening Same-Sex Relationships* (Whitton et al., 2016, 2017) y *Within Our Reach* (Barden et al., 2021), produjeron mejoras significativas en la comunicación de pareja y reducciones del estrés percibido, con tamaños del efecto pequeños a moderados y, en algunos casos, mantenidos en el seguimiento. Asimismo, la intervención *Behavioral Couples Therapy* adaptada a población LGB mostró beneficios tanto en el funcionamiento relacional como en variables clínicas específicas, como la reducción del consumo de alcohol, en comparación con un tratamiento individual sin participación de la pareja (Fals-Stewart et al., 2009). En el ámbito de la salud sexual y relacional, el programa *It Takes Two* evidenció tendencias a la reducción del estigma relacional durante el seguimiento, aunque no se observaron efectos significativos en otros desenlaces primarios (Gamarel et al., 2025).

Aunque ninguno de los estudios incluidos evaluó de forma directa la homofobia interiorizada como desenlace principal, los resultados observados pueden interpretarse a la luz del modelo de estrés de minorías (Meyer, 2003), que conceptualiza la homofobia interiorizada como un estresor proximal con impacto tanto individual como diádico. Desde esta perspectiva, las mejoras en comunicación, satisfacción relacional, afrontamiento diádico y reducción del estrés psicológico pueden considerarse indicadores indirectos de una

disminución del impacto del estigma interiorizado en la dinámica de pareja, en línea con estudios empíricos y aportaciones teórico-clínicas previas que han vinculado la homofobia interiorizada con mayor conflicto, menor intimidad y menor percepción de apoyo mutuo en la relación de pareja (Pepping y Halford, 2014; Hambrook, 2025).

La integración explícita de componentes afirmativos y culturalmente sensibles en las intervenciones revisadas resulta coherente con la literatura teórica que subraya la importancia de abordar el estrés minoritario en el trabajo clínico con parejas LGBTQ+. Estudios previos señalan que el afrontamiento diádico del estrés constituye un mecanismo clave para el bienestar psicológico y relacional, independientemente del diagnóstico clínico (Landolt et al., 2023). En este sentido, las intervenciones afirmativas no solo actúan sobre habilidades relacionales generales, sino que también validan la identidad sexual y de género, contextualizan el malestar en experiencias de estigmatización y promueven estrategias compartidas para manejar el estrés minoritario, lo que podría contribuir a una mayor adecuación clínica para parejas LGBTQ+ en comparación con modelos que no integran explícitamente estos factores.

Los hallazgos de esta revisión son coherentes con revisiones sistemáticas previas identificadas mediante búsqueda manual. Kousteni y Anagnostopoulos (2019) concluyeron que existe evidencia preliminar de beneficios de intervenciones psicológicas dirigidas a parejas del mismo sexo, especialmente en variables como la satisfacción con la intervención, el funcionamiento relacional y conductas relacionadas con la salud, si bien subrayaron la limitada cantidad de estudios disponibles y que el riesgo de sesgo global de los ensayos incluidos fue moderado. De forma similar, la revisión de Brito Silva et al. (2022) señaló que las intervenciones afirmativas y culturalmente adaptadas muestran resultados prometedores en el bienestar psicológico y relacional de personas LGBTQ+, pero destacó la limitada

calidad metodológica y la heterogeneidad de los estudios disponibles. La presente revisión amplía estas conclusiones al centrarse específicamente en contextos de pareja y al integrar los resultados desde el marco del estrés minoritario y la homofobia interiorizada, aportando una lectura clínica más articulada de los efectos observados.

No obstante, la evidencia incluida presenta importantes limitaciones. En primer lugar, el número de estudios es reducido y las muestras son, en general, pequeñas y poco diversas, con predominio de participantes blancos y de nivel socioeconómico medio-alto, lo que limita la generalización de los resultados. En segundo lugar, existe una considerable heterogeneidad en cuanto a los tipos de intervención, los desenlaces evaluados y los instrumentos utilizados, lo que impidió la realización de un metanálisis y obligó a optar por una síntesis narrativa. Además, la ausencia de evaluaciones directas de homofobia interiorizada constituye una limitación relevante, dado que este es un constructo central en la pregunta de investigación planteada.

En cuanto a las limitaciones del proceso de revisión, aunque se siguieron las directrices PRISMA 2020 y las recomendaciones metodológicas del Manual Cochrane, la búsqueda se restringió a artículos publicados en determinados idiomas y bases de datos, lo que podría haber excluido estudios relevantes no indexados o publicados en otros contextos. Asimismo, no se realizó una búsqueda exhaustiva de literatura gris, lo que puede haber contribuido a un sesgo de publicación. A pesar de estas limitaciones, la utilización de criterios de elegibilidad explícitos, el cribado y la extracción de datos por duplicado y la evaluación sistemática del riesgo de sesgo fortalecen la validez interna de la revisión.

Desde una perspectiva clínica, los resultados sugieren que las intervenciones de pareja dirigidas a población LGBTQ+ deberían incorporar explícitamente principios de práctica afirmativa —tal como enmarcan las *Guidelines for Psychological Practice with Sexual*

*Minority Persons* de la APA (2021)— y atender al impacto del estigma y el estrés asociado a la condición de minoría, factores que estas guías describen como clínicamente relevantes. La validación identitaria, la contextualización del malestar psicológico y relacional y el fortalecimiento de redes de apoyo emergen como componentes clave para promover el bienestar en la pareja. En términos de políticas y servicios de salud mental, estos hallazgos apoyan la necesidad de formación específica de los profesionales en competencias culturales y afirmativas, así como el desarrollo de programas adaptados a las realidades de las parejas LGBTQ+.

En relación con futuras líneas de investigación, resulta prioritario el diseño de ensayos controlados aleatorizados con muestras más amplias y diversas, que incluyan medidas explícitas de homofobia interiorizada y estrés minoritario como desenlaces primarios, así como seguimientos a largo plazo que permitan evaluar la estabilidad de los efectos. Asimismo, sería relevante comparar de forma directa intervenciones afirmativas con modelos tradicionales no adaptados, con el fin de esclarecer en qué medida la inclusión de componentes afirmativos aporta beneficios adicionales al bienestar relacional.

En respuesta a la pregunta de investigación planteada, la evidencia disponible no permite afirmar de manera concluyente que la terapia afirmativa sea más efectiva que otras intervenciones psicológicas en la reducción directa de la homofobia interiorizada, dado que este desenlace no fue evaluado de forma específica en los estudios incluidos. Sin embargo, los resultados sugieren que las intervenciones afirmativas o culturalmente adaptadas muestran una mayor adecuación clínica y resultados más favorables en determinados desenlaces del bienestar psicológico y relacional de las parejas LGBTQ+, actuando probablemente sobre mecanismos indirectos vinculados al estrés minoritario y al estigma interiorizado. En este sentido, la terapia afirmativa se perfila como un enfoque especialmente relevante y

prometedor para el trabajo con parejas LGBTQ+, aunque se requiere investigación adicional de mayor calidad metodológica para confirmar su superioridad frente a otros modelos terapéuticos.

### Referencias bibliográficas

- American Psychological Association. (2021). Guidelines for psychological practice with sexual minority persons.  
<https://www.apa.org/about/policy/psychological-sexual-minority-persons.pdf>
- Belous, C. K. (15 de mayo de 2015). Couple therapy with lesbian partners using an affirmative-contextual approach. *American Journal of Family Therapy*, 43(3), 269–281. <https://doi.org/10.1080/01926187.2015.1012234>
- Biaggio, M., Coan, S., y Adams, W. (2002). Couples therapy for lesbians. *Journal of Lesbian Studies*, 6(1), 129–138. [https://doi.org/10.1300/J155v06n01\\_12](https://doi.org/10.1300/J155v06n01_12)
- Boe, J. L., Bermúdez, J. M., Sharstrom, K. A., y Baldwin, D. R. (junio de 2019). Easing the transition: A critical narrative therapy approach to working with committed couples navigating gender transition. *Journal of Systemic Therapies*, 38(1), 1–16.  
<https://doi.org/10.1521/jsyt.2019.38.1.1>
- Cabaj, R. P., y Klinger, R. L. (1996). Psychotherapeutic interventions with lesbian and gay couples. In *Textbook of homosexuality and mental health* (pp. 485–501).  
<https://psycnet.apa.org/record/1996-97801-028>
- CASPe. (2018). Plantilla para la lectura crítica de ensayos clínicos aleatorizados. Programa de Habilidades en Lectura Crítica. <https://www.redcaspe.org>
- Cabello López, J. B., López Briz, E., y Pijoan Zubizarreta, J. I. (27 de mayo de 2022). Lectura crítica de estudios de tratamiento: Ensayos clínicos aleatorizados. En J. B. Cabello López (Ed.), *Lectura crítica de la evidencia clínica* (pp. 36–56). Elsevier España.
- Centro Cochrane Iberoamericano. (2012). *Manual Cochrane de revisiones sistemáticas de intervenciones* (versión 5.1.0 [actualizado en marzo de 2011]; J. P. T. Higgins y S.

Green, Eds.). Centro Cochrane Iberoamericano.

<https://www.cochrane.es/?q=es/node/269>

Crenshaw, Kimberle (1989) "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics," *University of Chicago Legal Forum*: Vol. 1989, Article 8.

<https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>

de Brito Silva, B., Soares de Almeida-Segundo, D., de Miranda Ramos, M., Bredemeier, J., y Cerqueira-Santos, E. (21 de septiembre de 2021). Couple and Family Therapies and Interventions with Lesbian, Gay and Bisexual Individuals: A Systematic Review. *Journal of Couple & Relationship Therapy*, 21(1), 52–79.

<https://doi.org/10.1080/15332691.2021.1978360>

Fals-Stewart, W., O'Farrell, T. J., y Lam, W. K. (diciembre de 2009). Behavioral couple therapy for gay and lesbian couples with alcohol use disorders. *Journal of substance abuse treatment*, 37(4), 379–387. <https://doi.org/10.1016/j.jsat.2009.05.001>

Gamarel, K. E., Operario, D., Stein, E. S., Pollack, L. M., Neilands, T. B., McCree, B., Johnson, J., Correll-King, W. M., Olem, D., Johnson, M. O., y Sevelius, J. M. (noviembre de 2025). Randomized Controlled Trial of a Couples-Based HIV Intervention for Transgender Women and Their Partners: Findings from the "It Takes Two" Project. *AIDS and behavior*, 29(11), 3544–3556.

<https://doi.org/10.1007/s10461-025-04798-5>

Gray, T. D., Dovala, T. K., y Cordova, J. V. (23 de noviembre de 2022). Affirmative Adaptations of the Relationship Checkup to Meet the Needs of LGBTQ Couples. *Journal of Couple & Relationship Therapy*, 22(4), 299–323.

<https://doi.org/10.1080/15332691.2022.2149652>

- Green, R.J., y Mitchell, V. (2015). Gay, lesbian, and bisexual issues in couple therapy. *Clinical Handbook of Couple Therapy* (5th ed., pp. 489–511).  
<https://psycnet.apa.org/record/2015-25194-018>
- Greenan, D. (2019). Resiliency-focused couple and family therapy with gay men. In *Integrative couple and family therapies* (pp. 199–217).  
<https://doi.org/10.1037/0000151-010>
- Hambrook, D. (2025). Working with sexual minority couples. In *Case studies in cognitive behavioural couple therapy* (pp. 106–124). <https://doi.org/10.4324/9781003024439>
- Hatch, S. G., Rothman, K., Roddy, M. K., Dominguez, R. M., Le, Y., y Doss, B. D. (2021). Heteronormative relationship education for same-gender couples. *Family Process*, 60(1), 119–133. <https://doi.org/10.1111/famp.12548>
- Landolt, S. A., Weitkamp, K., Roth, M., Sisson, N. M., y Bodenmann, G. (diciembre de 2023). Dyadic coping and mental health in couples: A systematic review. *Clinical Psychology Review*, 106, 102344. <https://doi.org/10.1016/j.cpr.2023.102344>
- Lebow, J., y Snyder, D. K. (29 de septiembre de 2022). Couple therapy in the 2020s: Current status and emerging developments. *Family process*, 61(4), 1359–1385.  
<https://doi.org/10.1111/famp.12824>
- Levy, B., y Middleton, A. (7 de enero de 2025). We carry on living in this othered world: Interpretative phenomenological analysis of lesbian/queer co-mothers' relationships with professional support. *Journal of Family Therapy*, 47(1), 1–16.  
<https://doi.org/10.1111/1467-6427.12482>
- Lord, S. A. (22 de marzo de 2017). Mindfulness and spirituality in couple therapy: The use of meditative dialogue to help couples develop compassion and empathy for themselves and each other. *Australian and New Zealand Journal of Family Therapy*, 38(1), 98–114. <https://doi.org/10.1002/anzf.1201>

- Malley, M. (2 de mayo de 2002). Systemic therapy with lesbian and gay clients: A truly social approach to psychological practice. *Journal of Community & Applied Social Psychology*, 12(3), 237–241. <https://doi.org/10.1002/casp.676>
- Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: Conceptual issues and research evidence. *Psychological Bulletin*, 129(5), 674–697. <https://doi.org/10.1037/0033-2909.129.5.674>
- Niki, D. (29 de diciembre de 2017). Now you see me, now you don't: addressing bisexual invisibility in relationship therapy. *Sexual and Relationship Therapy*, 33(1–2), 45–57. <https://doi.org/10.1080/14681994.2017.1419563>
- Ouzzani, M., Hammady, H., Fedorowicz, Z., y Elmagarmid, A. (5 de diciembre de 2016). Rayyan—A web and mobile app for systematic reviews. *Systematic Reviews*, 5(1), Article 210. <https://doi.org/10.1186/s13643-016-0384-4>
- 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., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., y Moher, D. (4 de enero de 2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pepping, C. A., y Halford, W. K. (17 de diciembre de 2014). Relationship education and therapy for same-sex couples. *Australian and New Zealand Journal of Family Therapy*, 35(4), 431–444. <https://doi.org/10.1002/anzf.1075>
- Randall, A. K., Totenhagen, C. J., Calveri, E. G., Bodenmann, G., Greving, K., y Rostosky, S. S. (2025). The Couples Coping Enhancement Training–Sexual Minority Stress (CCET-SMS) Relationship Education Program: Feasibility, Acceptability, and

- Preliminary Evidence of Effectiveness. *LGBTQ+ Family: An Interdisciplinary Journal*, 21(3), 238-253. <https://doi.org/10.1080/27703371.2024.2405148>
- Samons, S. L. (19 de octubre de 2009). Can this marriage be saved? Addressing male-to-female transgender issues in couples therapy. *Sexual and Relationship Therapy*, 24(2), 152–162. <https://doi.org/10.1080/14681990903002748>
- Sejal M. Barden, Ryan G. Carlson, Thomas Toomey, Lauren Locklear, Dalena Dillman Taylor y Nicole Silverio (2021): Examining the Effects of Relationship Education for Same-Sex Couples: A Randomized Control Trial, *Marriage & Family Review*, <https://doi.org/10.1080/01494929.2021.1921896>
- Souza, C., y Santos, M. A. dos. (2025). Couple therapy for lesbian women: Current evidence-base and recommendations for therapists. *Psicologia: Teoria e Prática*, 27(1), ePTPCP15547. <https://doi.org/10.5935/1980-6906/ePTPCP15547.en>
- Su, T., y Parker, M. L. (6 de junio de 2022). Narrative couple therapy with sexual minority couples: Exploring sexual intimacy. *Contemporary Family Therapy: An International Journal*, 44(4), 333–343. <https://doi.org/10.1007/s10591-022-09644-7>
- Tudor-Sfetea, C., y Topciu, R. (30 de septiembre de 2024). A Systematic Review of Evidence-Based Cognitive and/or Behavioural Interventions Targeting Mental Health in LGBTQ+ Populations. *Clinical psychology in Europe*, 6(3), e11323. <https://doi.org/10.32872/cpe.11323>
- Whitton, S. W., Weitbrecht, E. M., Kuryluk, A. D., y Hutsell, D. W. (2 de marzo de 2016). A randomized waitlist-controlled trial of culturally sensitive relationship education for male same-sex couples. *Journal of Family Psychology*, 30(6), 763–768. <https://doi.org/10.1037/fam0000199>
- Whitton, S. W., Scott, S. B., Dyar, C., Weitbrecht, E. M., Hutsell, D. W., y Kuryluk, A. D. (2017). Piloting relationship education for female same-sex couples: Results of a

small randomized waitlist-control trial. *Journal of Family Psychology*, 31(7), 878–888. <https://doi.org/10.1037/fam0000337>

## Apéndices

### Apéndice A. Estrategia de búsqueda

Bases de datos consultadas	Estrategia de búsqueda
APA PsycInfo, CINAHL with Full Text, MEDLINE Complete y Psychology and Behavioral Sciences Collection (vía EBSCO-host)	(((bisexual*[Title/Abstract] OR gay[Title/Abstract] OR lesbian*[Title/Abstract] OR homosexual*[Title/Abstract] OR "men who have sex with men"[Title/Abstract] OR queer[Title/Abstract] OR "sexual minorities"[Title/Abstract] OR transgender[Title/Abstract] OR transsexual*[Title/Abstract] OR "women who have sex with women"[Title/Abstract]) OR ("Sexual and Gender Minorities" OR "Homosexuality" OR "Bisexuality"[MeSH Terms])) AND (couple*[Title/Abstract] OR partner*[Title/Abstract])) AND (("couple therap*[Title/Abstract] OR "couples therap*[Title/Abstract] OR "marital therap*[Title/Abstract] OR "relationship therap*[Title/Abstract] OR "couple counsel*"OR "marital counsel*[Title/Abstract] OR "relationship counsel*[Title/Abstract] OR "couple intervention*[Title/Abstract] OR "relationship intervention*[Title/Abstract] OR "dyadic intervention*[Title/Abstract] OR "couple-based intervention*[Title/Abstract] OR "partner-based intervention*"OR "partner intervention*[Title/Abstract]) OR ("Couples Therapy" OR "Marital Therapy" OR "Family Therapy"[MeSH Terms]))

**Apéndice B. Evaluación de la calidad metodológica del estudio de Barden et al. (2021) según CASPe**

<b>Ítem CASPe</b>	<b>Respuesta</b>	<b>Frase literal del artículo</b>	<b>Justificación</b>
<p><b>1. ¿Se formuló una pregunta claramente definida?</b>  <i>Una pregunta debe definirse en términos de la población de estudio, la intervención realizada, la comparación y los desenlaces considerados (positivos y negativos).</i></p>	Sí	<p><i>“We propose the following research question: What are the effects of relationship education for same-sex couples’ dyadic coping, emotion regulation, psychological distress, and dyadic adjustment in the treatment group compared to control group at post-assessment (one month after random assignment)?”</i>            P → Parejas del mismo sexo adultas (same-sex couples).            I → Educación en relaciones (Relationship Education – Within Our Reach (WOR)).            C → Grupo control en lista de espera (wait-list control).            O → malestar psicológico, ajuste diádico, afrontamiento diádico y regulación emocional</p>	<p>El estudio formula una pregunta claramente definida en términos de población, intervención, comparación y desenlaces. Aunque los autores no diferencian explícitamente un desenlace principal y secundarios, el malestar psicológico se considera el desenlace principal por ser el eje central de los resultados y la discusión, mientras que el ajuste diádico, el afrontamiento diádico y la regulación emocional se consideran desenlaces secundarios.</p>
<p><b>2. ¿Los participantes fueron asignados aleatoriamente?</b>            ¿Se generó adecuadamente la secuencia?            ¿Se mantuvo oculta la secuencia de aleatorización?            ¿Son iguales en línea basal?</p>	No sé	<p>“Random assignment in the larger study occurred at a 1:1 ratio after completion of assessments at intake,”            “Individuals in the treatment and control condition did not differ on baseline characteristics, as indicated by Cohen’s d values less than .05. Specifically, no differences existed between groups on demographics including gender, age, ethnicity, employment status, partner status, or number of children.”</p>	<p>Aunque el estudio informa de asignación aleatoria en una proporción 1:1 y de comparabilidad basal entre los grupos, no se describen el método de generación ni la ocultación de la secuencia de aleatorización, por lo que el riesgo de sesgo de selección no puede ser evaluado con certeza.</p>
<p><b>3. ¿Se mantuvo la comparabilidad de los grupos a través del estudio?</b></p>	No sé	<p><b>Comparabilidad basal:</b>  <i>“The individuals in the treatment and control condition did not differ on baseline characteristics, as indicated by Cohen’s d values less than .05. pecifically, no differences existed between</i></p>	<p>Aunque el artículo indica comparabilidad basal entre grupos, no aporta información suficiente sobre posibles desviaciones, contaminación/cocrintervenciones o</p>

**Desviaciones de la intervención por problemas en la asignación.**

**Desviaciones de la intervención por problemas en la adhesión al tratamiento.**

**¿Cómo se analizó el estudio: ITT mITT APP, AT?**

**4. ¿Son importantes las pérdidas ocurridas durante el estudio?** No

**¿Difieren según el grupo?**

**¿Las pérdidas podrían depender de su valor o resultado?**

**¿Se hace análisis de sensibilidad?**

*groups on demographics including gender, age, ethnicity, employment status, partner status, or number of children.*

**Desviaciones por asignación:** no se describen.

**Desviaciones por adhesión / pérdidas:**

*“Data were missing at post-treatment (14.6%)./ “As with most community-based trials, this study experienced attrition over the 6-month study period.”*

**Tipo de análisis:**

*“data were imputed using multiple imputation processes in Mplus 8.0.” / “Twenty imputed datasets were created, and all reported results are pooled across those datasets.”*

**Cuántas pérdidas hubo:**

*“Data were missing at post-treatment (14.6%).”*

**Difieren según grupo:** No se informa explícitamente si las pérdidas difieren entre el grupo de tratamiento y el grupo control.

**Pérdidas-valor-resultado:**

*“To determine the pattern of missingness, we examined whether missing data were substantively related to important study variables such as demographic variables or outcome variables. No such relationships were found.” // We also examined Little’s MACR test to examine pattern of missingness, which was not statistically significant,  $\chi^2 = 67.087$ ,  $df=112$ ,  $p=1.00$ . “As such, we fail to reject the null hypothesis, providing support that our data are missing completely at random.”*

**Datos perdidos:**

*“Therefore, data were imputed using multiple imputation processes in Mplus 8.0...”*

*“Twenty imputed datasets were created, and all reported results are pooled across those datasets.”*

**Análisis de sensibilidad:** no consta.

adherencia diferencial que permita asegurar que la comparabilidad se mantuvo a lo largo del estudio.

Además, no se declara explícitamente un análisis por intención de tratar; se manejaron los datos faltantes mediante imputación múltiple (20 conjuntos), lo que aproxima un enfoque ITT pero no permite etiquetarlo como mITT con certeza.

Hubo datos faltantes a post-tratamiento (14.6%). Los autores evaluaron el patrón de missingness y no hallaron relación con variables demográficas ni con los desenlaces; además, el test de Little apoyó que los datos eran MCAR. Los datos faltantes se manejaron mediante imputación múltiple (20 conjuntos). No se informa explícitamente si las pérdidas difieren entre grupos y no se describe la realización de análisis de sensibilidad. Por tanto, las pérdidas no parecen importantes en términos de sesgo.

<p><b>5. ¿Fue adecuada la medición de los desenlaces?</b>  <b>Tipo de desenlace medido y método usado.</b>  <b>Cegamiento del paciente, clínico, evaluador, estadístico.</b>  <b>Si hay problema, ¿es diferencial entre los grupos?</b></p>	Sí	<p><b>Tipo de desenlace medido y metodo usado:</b>  “included in this study are responses to a general demographics form, the Dyadic Coping Inventory, the Difficulties in Emotion Regulation Scale, the Outcomes Questionnaire 45.2, and Revised Dyadic Adjustment Scale.” /  “<i>participants completed the study measures via an iPad,</i>”</p> <p><b>Cegamiento del paciente:</b>  “...receive the treatment immediately, or to the wait-list control”</p> <p><b>Cegamiento del clínico:</b>  No se describe cegamiento de los facilitadores en el artículo.</p> <p><b>Cegamiento del evaluador:</b>  “<i>participants completed the study measures via an iPad,</i>”</p> <p><b>Cegamiento estadístico:</b> no consta.</p> <p><b>Problema diferencial:</b>  “<i>participants completed the study measures via an iPad,</i>”</p>	<p>Los desenlaces se midieron mediante instrumentos estandarizados y validados (DCI, DERS, OQ-45.2 y RDAS), aplicados de forma uniforme mediante autoinformes vía iPad en ambos grupos. No se describe cegamiento de participantes, facilitadores, evaluadores ni del analista, lo que supone una posible limitación por sesgo de medición. Sin embargo, el procedimiento de evaluación fue idéntico en ambos grupos, por lo que el posible sesgo por falta de cegamiento no parece diferencial entre los grupos.</p>
<p><b>6. ¿Se evitó la comunicación selectiva de resultados?</b></p> <p><b>(Mirar el registro de ensayos)</b>  <b>¿Hay reporte selectivo de desenlaces o reporte selectivo de análisis?</b></p>	No sé	<p><b>Registro del ensayo:</b> no consta registro.</p> <p><b>Selección de desenlaces en este subestudio:</b>  “Participants in the current investigation completed a larger battery of assessments...”  “However, due to the small sub-sample and limitations with power, we selected outcomes...”</p> <p><b>Desenlaces declarados y reportados:</b>  “the current paper... examines the impact of RE on four outcomes: dyadic coping, psychological distress, emotion regulation, and dyadic adjustment.”  (Estos cuatro desenlaces se reportan en Results, incluidos los no significativos.)</p>	<p>No consta registro prospectivo ni protocolo que permita verificar los desenlaces y análisis preespecificados. Además, los autores indican que, aunque se administró una batería amplia de instrumentos, seleccionaron desenlaces para este subanálisis por limitaciones de potencia, lo que impide descartar completamente reporte selectivo de resultados. Aun así, los cuatro desenlaces que declaran evaluar se informan en resultados, incluidos los no significativos.</p>
<p><b>7. ¿Cuál fue el efecto del tratamiento?</b></p>	Sí	<p><b>Desenlace principal:</b>  There was a significant difference between the treatment and waitlist control conditions on the measure of individual psychological distress (OQ), as indicated by the slopes of the</p>	<p>Aunque los coeficientes son negativos (<math>b = -0.24</math>; <math>d = -0.51</math>), esto indica una reducción del estrés psicológico en el</p>

**¿Qué desenlaces se han medido? Detalla los positivos y los negativos.**

outcome questionnaire ( $b = 0.24$ ,  $SE = .12$ ,  $p = .04$ ,  $R^2 = .06$ , Cohen's  $d = -0.51$ ,  $IC95\% = -.91$  to  $-.11$ ).

**Desenlaces secundarios:**

In contrast to our hypotheses, there was no significant difference between the treatment and waitlist control conditions on the measures of dyadic coping or difficulties in emotion regulation, as indicated by the slopes of the dyadic coping inventory ( $b = .11$ ,  $SE = .11$ ,  $p = .31$ ,  $R^2 = .01$ , Cohen's  $d = .23$ ,  $IC95\% = -.17$  to  $.63$ ), the Difficulties in Emotion Regulation Scale ( $b = -.06$ ,  $SE = 0.16$ ,  $p = .73$ ,  $R^2 = .01$ , Cohen's  $d = -.12$ ,  $IC95\% = -.51$  to  $.28$ ), or the RDAS scale ( $b = .16$ ,  $SE = 0.17$ ,  $p = .33$ ,  $R^2 = .03$ , Cohen's  $d = .13$ ,  $IC95\% = -.30$  to  $.56$ )

grupo tratado, ya que puntuaciones más bajas en el OQ representan menor malestar. El desenlace de distrés psicológico individual mostró una reducción significativa y de magnitud moderada (desenlace con efecto significativo), mientras que los desenlaces relacionales (afrentamiento diádico, regulación emocional y ajuste diádico) no mostraron efectos significativos (desenlaces sin efecto significativo). Por tanto, los resultados son mixtos y el beneficio se limita al distrés psicológico individual, sin evidencia de efecto en los desenlaces relacionales.

**8. ¿Cuál es la precisión de los estimadores del efecto?**

*¿Cuáles son sus intervalos de confianza?*

Sí

**Desenlace principal:** “(OQ) ( $b = -0.24$ ,  $SE = .12$ ,  $p = .04$ ,  $R^2 = .06$ , Cohen's  $d = -0.51$ ,  $IC95\% = -.91$  to  $-.11$ ).”

**Desenlaces secundarios:** “... dyadic coping inventory ( $b = .11$ ,  $SE = .11$ ,  $p = .31$ ,  $R^2 = .01$ , Cohen's  $d = .23$ ,  $IC95\% = -.17$  to  $.63$ ), the Difficulties in Emotion Regulation Scale ( $b = -.06$ ,  $SE = 0.16$ ,  $p = .73$ ,  $R^2 = .01$ , Cohen's  $d = -.12$ ,  $IC95\% = -.51$  to  $.28$ ), or the RDAS scale ( $b = .16$ ,  $SE = 0.17$ ,  $p = .33$ ,  $R^2 = .03$ , Cohen's  $d = .13$ ,  $IC95\% = -.30$  to  $.56$ ).”

El intervalo de confianza del desenlace de distrés psicológico no incluye el 0 ( $IC95\% = -0.91$  a  $-0.11$ ), lo que indica un efecto estadísticamente significativo con precisión moderada (desenlace con mejor precisión). En cambio, los intervalos de confianza de los desenlaces relacionales son amplios e incluyen el 0, lo que refleja baja precisión y gran incertidumbre sobre la magnitud real del efecto (desenlaces con baja precisión). Por tanto, la precisión es aceptable solo para el desenlace principal y limitada para los desenlaces secundarios.

<p><b>9. ¿Pueden aplicarse estos resultados en tu medio o población local?</b></p>	Sí	<p>Few studies have empirically examined the impact of RE programs with same-sex couples (Buzzella et al., 2012; Newcomb et al., 2017; Whitton et al., 2016; Whitton et al., 2017) and the generalizability of these studies is limited. In this paper, we present the findings of a randomized controlled trial of a relationship education program with a sub-sample of same-sex couples (N=52). [...] Recently, Hatch et al. (2021) studied the effectiveness of a web-based, heteronormative RE program (OurRelationship) with a subsample of same-sex couples (n=49) from a larger randomized controlled trial.</p>	<p>La población del estudio cumple los criterios de inclusión de esta revisión: adultos <math>\geq 18</math> años, parejas del mismo sexo y participantes LGBTQ+ en contexto de pareja. Además, en la introducción los autores mencionan varios estudios similares en esta población (ref. 8, ref. 40, ref. 63, ref. 64 y ref. 24 del artículo), lo que apoya la consistencia externa de los resultados. Por tanto, los hallazgos son aplicables a una población local de parejas LGBTQ+ en intervención psicológica o psicoeducativa. No obstante, la generalización de los resultados es limitada a parejas del mismo sexo reclutadas en contextos comunitarios similares, tal como señalan los propios autores. Aunque el estudio evalúa utilidades en términos de reducción del estrés psicológico individual y otros desenlaces relacionales, no se analizan disutilidades ni posibles efectos adversos asociados a la intervención. Tampoco se consideran las preferencias de los pacientes, ni se discuten costes o carga del tratamiento, ni se realiza un balance explícito entre utilidades y disutilidades. Por tanto, no se han tenido en cuenta todos los</p>
<p><b>10. ¿Se han tenido en cuenta todos los resultados y su importancia clínica?</b></p> <p>Utilidades y disutilidades de cada desenlace. Balance de efectos positivos/ negativos. Preferencias del paciente, costes, etc.</p>	No	<p>“...included in this study are responses to a general demographics form, the Dyadic Coping Inventory, the Difficulties in Emotion Regulation Scale, the Outcomes Questionnaire 45.2, and Revised Dyadic Adjustment Scale.”</p>	

**11. ¿Los beneficios que se espera obtener justifican los riesgos y los costes?**

No

*These findings suggest that WOR was effective in decreasing symptoms of psychological distress in same-sex couples immediately following the completion of the intervention.*

Es improbable que pueda deducirse solo de un ensayo, pero ¿qué piensas tú al respecto?

resultados relevantes para establecer la importancia clínica global.

El beneficio observado se limita a una reducción del estrés psicológico individual inmediatamente tras la intervención, sin efectos significativos en los desenlaces relacionales.

Además, el estudio no evalúa costes, preferencias de los pacientes ni posibles riesgos o efectos adversos asociados al programa. Por tanto, con la información disponible no puede concluirse que los beneficios obtenidos justifiquen claramente los riesgos y los costes asociados a la intervención.

---

Apéndice C. Evaluación de la calidad metodológica del estudio de Fals-Stewart et al. (2009) según CASPe.

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>1. ¿Se orienta el ensayo a una pregunta claramente definida?</b> Una pregunta debe definirse en términos de la población de estudio, la intervención realizada, la comparación y los desenlaces considerados (positivos y negativos).</p>	Sí	<p>“Gay (n = 52) and lesbian (n = 48) patients with alcohol use disorder (AUD) and their non-substance-abusing same-sex relationship partners were randomly assigned to equally intensive interventions consisting of (a) behavioral couples therapy (BCT) plus individual-based treatment (IBT) or (b) IBT only.”</p>	<p>El estudio define claramente los cuatro componentes PICO: población (pacientes gays y lesbianas con AUD y sus parejas), intervención (BCT + IBT), comparación (IBT solo) y desenlaces, cumpliendo los criterios CASPe de pregunta claramente formulada. Los desenlaces están claramente definidos y medidos con instrumentos validados (TLFB para consumo y DAS para ajuste relacional).</p>
<p><b>2. ¿Fue aleatoria la asignación a los tratamientos?</b> ¿Se generó adecuadamente la secuencia? ¿Se mantuvo oculta la secuencia de aleatorización? ¿Son iguales en línea basal?</p>	Sí	<p><b>Aleatorización:</b> “Participants were randomly assigned to BCT or IBT; they were not informed of their assignment until they arrived at the first treatment session.”</p> <p><b>Cómo se generó la secuencia:</b> no consta.</p> <p><b>Ocultación de la secuencia:</b> “..they were not informed of their assignment until they arrived at the first treatment session.”</p> <p><b>Línea basal:</b> “Random assignment was effective; comparisons of background characteristics of participants assigned to BCT versus IBT for both gay and lesbian couples revealed no significant differences (i.e., all ps &gt; .25).”</p>	<p>La asignación a los tratamientos se describe como aleatoria. Sin embargo, no se especifica el método de generación de la secuencia de aleatorización, por lo que no es posible confirmar si esta se generó adecuadamente.</p> <p>La ocultación de la asignación fue parcial: los participantes desconocían su grupo hasta el inicio de la primera sesión. No obstante, no se describe ningún sistema formal de ocultación de la secuencia, por lo que este aspecto se considera incierto.</p> <p>La comparabilidad en línea basal está adecuadamente demostrada, ya que no se encontraron diferencias significativas entre los grupos BCT e IBT en las características iniciales (<i>todas p &gt; .25</i>).</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p data-bbox="197 240 734 459"><b>3. ¿Se mantuvo la comparabilidad de los grupos a través del estudio? Desviaciones de la intervención por problemas en la asignación.</b></p> <p data-bbox="197 496 734 603"><b>Desviaciones de la intervención por problemas en la adhesión al tratamiento.</b></p> <p data-bbox="197 639 734 715"><b>¿Cómo se analizó el estudio: ITT mITT APP,AT?</b></p>	<p data-bbox="745 240 936 268">Sí</p>	<p data-bbox="965 240 1464 304"><b>Comparabilidad inicial (línea basal):</b></p> <p data-bbox="965 309 1464 528">“Random assignment was effective; comparisons of background characteristics of participants assigned to BCT versus IBT for both gay and lesbian couples revealed no significant differences (i.e., all ps N .25).”</p> <p data-bbox="965 533 1464 564"><b>Desviaciones de la asignación:</b></p> <p data-bbox="965 569 1464 639">No se describen desviaciones de la asignación tras la aleatorización.</p> <p data-bbox="965 644 1464 715"><b>2. Adherencia al tratamiento (comparabilidad de exposición):</b></p> <p data-bbox="965 719 1464 938">“For gay couples, the number of sessions attended by the partners with AUD randomly assigned to BCT (M = 24.29, SD = 3.62) or IBT (M = 23.00, SD = 3.04) was not significantly different.”</p> <p data-bbox="965 943 1464 1161">“For lesbian couples, the number of sessions attended by the partners with AUD randomly assigned to BCT (M = 22.66, SD = 4.89) or IBT (M = 24.93, SD = 5.11) was not significantly different.”</p> <p data-bbox="965 1166 1464 1198"><b>3. Tipo de análisis:</b></p> <p data-bbox="965 1203 1464 1305">“Data from all participants randomly assigned to conditions were used (i.e., “intention-to-treat” analyses).”</p> <p data-bbox="965 1310 1464 1374">“Full-information maximum likelihood was used to address data</p>	<p data-bbox="1494 240 2033 343"><b>Justificación</b></p> <p data-bbox="1494 347 2033 603">No se describen desviaciones de la intervención debidas a errores en la asignación tras la aleatorización. La adherencia al tratamiento fue comparable entre grupos: <b>el número de sesiones recibidas no difirió significativamente entre BCT e IBT</b>, tanto en parejas gays como lesbianas, indicando exposición similar a la intervención.</p> <p data-bbox="1494 608 2033 790">Las pérdidas durante el seguimiento fueron moderadas y similares entre grupos (aproximadamente 77% con datos completos; pérdidas del 12% en gays y 8% en lesbianas).</p> <p data-bbox="1494 794 2033 1007">El análisis se realizó por <b>intención de tratar (ITT)</b>, utilizando datos de todos los participantes aleatorizados y métodos adecuados para manejar los datos perdidos, preservando la comparabilidad inicial de los grupos.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>4. ¿Son importantes las pérdidas ocurridas durante el estudio?</b></p> <p>¿Difieren según el grupo?            ¿Las pérdidas podrían depender de su valor o resultado?            ¿Se hace análisis de sensibilidad?</p>	No sé	<p>missingness in the multilevel growth models”</p> <p><b>1. Intensidad y duración del tratamiento:</b>            “Each treatment condition consisted of 32 scheduled 60- minute sessions conducted over a 20-week period”</p> <p><b>2. Cointervenciones diferentes:</b>            No se describen cointervenciones diferentes entre los grupos durante el estudio.</p> <p><b>3. Pérdidas- valor y resultado:</b>No se describe en el artículo ningún análisis sobre la relación entre las pérdidas y el resultado clínico.</p> <p><b>Análisis de sensibilidad:</b> no consta.</p>	<p>Ambos grupos recibieron la misma duración y frecuencia de tratamiento (32 sesiones de 60 minutos en 20 semanas). No se describen cointervenciones diferenciales ni tratamientos adicionales distintos entre grupos.</p> <p>La adherencia fue comparable entre BCT e IBT, ya que el número de sesiones recibidas no difirió significativamente en parejas gays ni lesbianas, indicando que los grupos fueron tratados de forma similar durante el estudio.</p> <p>No se aporta información que permita valorar si las pérdidas dependieron de la evolución clínica o del resultado, por lo que este aspecto se considera incierto.</p> <p>No se realizó ni se describe ningún análisis de sensibilidad para evaluar el impacto potencial de diferentes supuestos de imputación o escenarios extremos sobre los resultados, aunque el manejo de datos perdidos se realizó mediante métodos adecuados.</p>
<p><b>5. ¿Fue adecuada la medición de los desenlaces?</b></p> <p>Tipo de desenlace medido y método usado.            Cegamiento del paciente,</p>	No	<p><b>1. Tipo de desenlace y método:</b>            “The substance use index derived from the TLFB used in the present investigation was percent days of heavy drinking (PDHD).”</p>	<p>Los desenlaces principales se midieron mediante instrumentos validados (TLFB para consumo y DAS para ajuste relacional), lo que apoya una medición técnicamente adecuada desde el punto de vista psicométrico.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p>clínico, evaluador, estadístico. Si hay problema, ¿es diferencial entre los grupos?</p>		<p>“The Dyadic Adjustment Scale (DAS; Spanier, 1976) is the most widely used measure of global relationship satisfaction...” <b>2. Cegamiento:</b> No se describe en el artículo ningún procedimiento de cegamiento de pacientes, terapeutas, evaluadores ni estadísticos. <b>3. Problema diferencial entre grupos:</b> No se analiza ni se describe si la ausencia de cegamiento pudo afectar de forma diferente a BCT e IBT.</p>	<p>Sin embargo, no se describe ningún procedimiento de cegamiento de pacientes, terapeutas ni evaluadores, y la intervención (BCT vs IBT) es claramente distinguible. El desenlace principal (PDHD) se obtuvo mediante entrevista estructurada (Timeline Followback Interview), basada en el recuerdo del propio paciente, por lo que la ausencia de cegamiento puede introducir sesgo de información. No se aporta información que permita descartar que este sesgo pudiera ser diferencial entre los grupos. No se identifica en el artículo ningún registro previo del ensayo ni un protocolo publicado que permita comparar los desenlaces preespecificados con los finalmente reportados. Aunque los desenlaces principales (PDHD y DAS) se presentan de forma coherente con los objetivos del estudio, no es posible descartar la existencia de reporte selectivo de desenlaces o de análisis.</p>
<p><b>6. ¿Se evitó la comunicación selectiva de resultados??</b> (Mirar el registro de ensayos) ¿Hay reporte selectivo de desenlaces o reporte selectivo de análisis?</p>	No sé	<p>En ningún sitio aparece:</p> <ul style="list-style-type: none"> <li>● número de registro</li> <li>● ClinicalTrials.gov</li> <li>● ISRCTN</li> <li>● protocolo previo</li> </ul>	<p>No se identifica en el artículo ningún registro previo del ensayo ni un protocolo publicado que permita comparar los desenlaces preespecificados con los finalmente reportados. Aunque los desenlaces principales (PDHD y DAS) se presentan de forma coherente con los objetivos del estudio, no es posible descartar la existencia de reporte selectivo de desenlaces o de análisis.</p>
<p><b>7. ¿Cuál es el efecto del tratamiento para cada desenlace?</b> ¿Qué desenlaces se han medido? Detalla los positivos y los negativos.</p>	Sí	<p><b>1. ¿Qué desenlaces se han medido?</b> “The substance use index derived from the TLFB used in the present investigation was percent days of heavy drinking (PDHD).” “The Dyadic Adjustment Scale (DAS; Spanier, 1976) is the most widely used</p>	<p>El desenlace principal fue el porcentaje de días de consumo pesado (PDHD), claramente definido y medido mediante TLFB. BCT mostró un efecto beneficioso consistente frente a IBT, con menos días de consumo pesado en el postratamiento</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>8. ¿Cuál es la precisión de los estimadores del efecto?</b> ¿Cuáles son sus intervalos de confianza?</p>	No	<p>measure of global relationship satisfaction in couples therapy research.”</p> <p><b>2. ¿Cuál fue el efecto del tratamiento sobre el desenlace principal?</b> “BCT had a significantly lower percentage of days of heavy drinking during the year after treatment than patients who received IBT only.”</p> <p><b>3. ¿Qué efecto sobre los desenlaces secundarios?</b> “In addition, both gay and lesbian couples who received BCT reported higher levels of relationship adjustment at the end of treatment and in the year after treatment than those who received IBT only. Thus, the response of gay and lesbian couples with an alcoholic member to BCT was consistent with what has been observed with heterosexual couples.”</p> <p><b>4. ¿Hubo desenlaces negativos o favorables al control?</b> No se describen desenlaces en los que IBT sea superior a BCT. No se presentan intervalos de confianza para las estimaciones del efecto en los resultados.</p>	<p>y en todos los seguimientos hasta 12 meses, con diferencias estadísticamente significativas. Como desenlace secundario, el ajuste relacional (DAS) mejoró significativamente más en BCT que en IBT durante la intervención y el seguimiento. No se identifican desenlaces en los que IBT sea superior a BCT, por lo que el efecto global del tratamiento es favorable a BCT. Los resultados se presentan de forma clara, con identificación del desenlace principal, desenlaces secundarios y dirección consistente del efecto a favor de BCT.</p> <p>Aunque los resultados se presentan con valores de p y diferencias estadísticamente significativas, no se informan intervalos de confianza para los</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>9. ¿Pueden aplicarse estos resultados en tu medio o población local?</b>            ¿Crees que los pacientes incluidos en el ensayo son demasiado distintos a tus pacientes?            ¿Hay algún otro ensayo parecido a este?            ¿Es consistente con este?</p>	Sí	<p>Las diferencias entre grupos se expresan mediante valores de p y comparaciones por pares, por ejemplo: “BCT had significantly different scores than those in IBT at the assessment period based on pairwise contrasts (all ps &lt; .05).” (Tabla 2, nota al pie)</p> <p><b>1. Comparabilidad de pacientes:</b>            “Gay (n = 52) and lesbian (n = 48) patients with alcohol use disorder (AUD) and their non-substance-abusing same-sex relationship partners...”</p> <p><b>Existencia de estudios similares:</b>            “Despite the substantial evidence base for BCT, controlled studies of BCT with gay and lesbian patients with AUD have not appeared in the literature.”</p> <p><b>Consistencia con estudios previos:</b>            “As predicted, the response of gay and lesbian couples to BCT was consistent with what has been observed with heterosexual couples.”</p> <p><b>Potencial de aplicabilidad clínica:</b>            The present results indicate great promise for the use of BCT with gay and lesbian couples in which one</p>	<p>estimadores del efecto (ni para PDHD ni para DAS).            La ausencia de intervalos de confianza impide valorar directamente la precisión de las estimaciones y el rango de valores plausibles del efecto, por lo que la precisión de los resultados no puede considerarse adecuadamente evaluada.</p> <p>La población estudiada está claramente definida y coincide con la diana de esta revisión (adultos LGBTQ+ en contexto de pareja con AUD). No existían previamente ensayos controlados de BCT en esta población, lo que sitúa este estudio como el primero de su tipo. Sin embargo, los autores indican que los resultados son consistentes con los observados previamente en parejas heterosexuales y en meta-análisis previos, y consideran que la intervención tiene un potencial clínico relevante. Todo ello apoya que los resultados pueden considerarse aplicables, con cautela, a contextos clínicos similares y a la población objetivo de esta revisión.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>10. ¿Se han tenido en cuenta todos los resultados y su importancia clínica?</b></p> <p>Utilidades y disutilidades de cada desenlace. Balance de efectos positivos/ negativos. Preferencias del paciente, costes, etc.</p>	No	<p>member of the couple enters treatment for an AUD.</p> <p><b>Utilidades clínicas (consumo):</b> “...those who received BCT had a significantly lower percentage of days of heavy drinking during the year after treatment than patients who received IBT only. In addition, both gay and lesbian couples who received BCT reported higher levels of relationship adjustment at the end of treatment and in the year after treatment than those who received IBT only.”</p> <p><b>Utilidadess clínicas (relacional):</b> “Both gay and lesbian couples who received BCT reported higher levels of relationship adjustment at the end of treatment and in the year after treatment than those who received IBT only.”</p> <p><b>Limitaciones relevantes:</b> “Several important limitations also should be noted: (a) modest sample sizes used in each trial; (b) lack of secondary outcome measures (e.g., substance-related problems, cost-benefit analysis, partner violence) often included in studies of BCT with heterosexual patients; and (c) exclusive use of primary alcohol problem patients and their</p>	<p>El estudio evalúa adecuadamente los beneficios clínicos principales (reducción del consumo de alcohol y mejora del ajuste relacional) y describe limitaciones metodológicas relevantes. Sin embargo, no se realiza un análisis explícito del balance entre beneficios y posibles perjuicios, ni se consideran aspectos clave para la toma de decisiones clínicas como costes, carga del tratamiento, preferencias de los pacientes o impacto económico. Por ello, no puede considerarse que se hayan tenido en cuenta todos los resultados clínicamente relevantes y su importancia para la decisión clínica.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>11. ¿Los beneficios que se espera obtener justifican los riesgos y los costes?</b> Es improbable que pueda deducirse solo de un ensayo, pero ¿qué piensas tú al respecto?</p>	No	<p>non-substance-abusing partners (thus, it is unclear if results observed here would generalize to patients with primary drug abuse problems or couples in which both members have a current substance problem).”</p> <p><b>Preferencias del paciente, etc. :</b> No consta.</p> <p><b>Beneficios clínicos relevantes:</b> “those who received BCT reported significantly lower proportions of days of heavy drinking in the year after treatment than couples in which the patient with AUD received IBT only.” “couples who received BCT reported higher levels of relationship adjustment at the end of treatment and in the year after treatment than those who received IBT only.”</p> <p><b>Beneficio mantenido:</b> “That posttreatment follow-up effects were maintained through 1 year offers new evidence for long-term sustained effects of treatment in an RCT for both gay and lesbian substance abusers.”</p> <p><b>Carga del tratamiento:</b> “Each treatment condition consisted of 32 scheduled 60- minute sessions conducted over a 20-week period.”</p>	<p>Aunque el estudio muestra <b>beneficios clínicos claros y sostenidos</b> en reducción del consumo y mejora del ajuste relacional, se trata de una <b>intervención larga e intensiva</b> (32 sesiones en 20 semanas) y los propios autores reconocen la <b>ausencia de análisis coste-beneficio y de desenlaces negativos relevantes</b>. En ausencia de información sobre <b>costes, riesgos clínicos y carga económica</b>, no puede concluirse que los beneficios justifiquen de forma clara los riesgos y los costes del tratamiento.</p>

---

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
		<b>Ausencia de costes y balance:</b> “lack of secondary outcome measures (e.g., substance-related problems, cost-benefit analysis, partner violence)”	

---

Apéndice D. Evaluación de la calidad metodológica del estudio de Gamarel et al. (2025) (“It Takes Two”) según CASPe.

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>1. ¿Se orienta el ensayo a una pregunta claramente definida?</b></p> <p>Una pregunta debe definirse en términos de la población de estudio, la intervención realizada, la comparación y los desenlaces considerados (positivos y negativos).</p> <p>Escribe los desenlaces.</p>	Sí	<p><i>“The aim of this project is to alleviate HIV vulnerability by enhancing communication, improving shared goals, and decreasing relationship stigma.” [...] “We conducted a randomized controlled trial with 52 couples (104 individuals).” [...] “the intervention group received four 1-hour counseling sessions [...] “the control group received educational videos” ; “Specifically, intervention participants would have a lower prevalence of composite vulnerability for HIV, any condomless sex, any condomless sex with an outside partner, a lower incidence of STIs, a higher group mean for goal congruence and conflict engagement communication and positive communication, and a lower group mean for withdrawal communication and relationship stigma relative to control group participants”</i></p>	<p>El estudio define claramente los cuatro componentes PICO: población (parejas con al menos una mujer trans), intervención (4 sesiones de consejería basada en parejas), comparación (videos educativos sobre prevención del VIH) y desenlaces (vulnerabilidad al VIH, conductas sexuales, ITS, estigma relacional, metas compartidas y comunicación).</p>
<p><b>2. ¿Fue aleatoria la asignación a los tratamientos?</b></p> <p>¿Se generó adecuadamente la secuencia?</p> <p>¿Se mantuvo oculta la secuencia de aleatorización?</p>	Sí	<p><b>Asignación aleatoria:</b>  <i>“We implemented a parallel randomized controlled trial design, randomly assigning couples in a 1:1 ratio to one of two groups.”</i></p> <p><b>Generación de la secuencia:</b>  <i>“We used a simple random allocation generator for the randomization process stratified by couple-level HIV serostatus.”</i></p> <p><b>Ocultación de la secuencia:</b></p>	<p>La asignación fue claramente aleatoria, realizada mediante un generador de asignación aleatoria por ordenador en proporción 1:1 y estratificada por seroestatus VIH de la pareja, lo que asegura una secuencia adecuada. La aleatorización se realizó tras la evaluación basal, reduciendo el riesgo de manipulación. No se describe un mecanismo formal de ocultación de la secuencia y el estudio no fue enmascarado, lo que introduce riesgo de sesgo de selección y de ejecución. No obstante, los grupos estaban globalmente bien</p>

<b>Ítem CASPe</b>	<b>Respuesta</b>	<b>Frase literal del artículo</b>	<b>Justificación</b>
¿Son iguales en línea basal?		<p>“This study was conducted without any masking.” ; “Randomization of couples occurred after each member of the dyad provided a rapid point-of-care HIV-antibody test result and completed a baseline survey.”</p> <p><b>Iguals en línea basal:</b>            “There was only one significant difference between the intervention and control groups.” ; “The groups did not differ on study modality, age, race, ethnicity, income, employment, housing stability, incarceration history, length of relationship, couple-level HIV serostatus, gender of the trans woman’s partner, or type of sexual agreement at baseline.”</p>	<p>balanceados en la línea basal, con una única diferencia significativa en nivel educativo, lo que apoya una aleatorización adecuada, aunque con ocultación incompleta.</p>
<p><b>3. ¿Se mantuvo la comparabilidad de los grupos a través del estudio?</b>            Desviaciones de la intervención por problemas en la asignación.            Desviaciones de la intervención por problemas en la adhesión al tratamiento.            ¿Cómo se analizó el</p>	<p><b>Sí</b></p>	<p><b>Desviaciones de la intervención por problemas en la asignación:</b>            “Five of those dyads declined to participate or were lost to follow-up during the enrollment process, resulting in 52 dyads (n=104) being randomized, with 26 dyads (n=72) in each condition.”</p> <p><b>Desviaciones de la intervención por problemas en la adherencia / seguimiento:</b>            “Follow-up retention rates were computed separately for each partner, which ranged from 77% to 96%.”</p> <p><b>¿Cómo se analizó el estudio?:</b></p>	<p>No se describen desviaciones relevantes en la asignación ni cruces entre grupos tras la aleatorización. El seguimiento fue adecuado, con tasas de retención entre 77% y 96%, lo que limita el sesgo por pérdidas. El análisis incluyó a todos los participantes aleatorizados mediante modelos mixtos longitudinales que incorporan todas las mediciones repetidas, lo que corresponde a un análisis por intención de tratar modificado (mITT). No se realizó un análisis por protocolo ni por tratamiento recibido, y no se reportan exclusiones diferenciales entre grupos, por lo que la comparabilidad de los grupos se mantuvo razonablemente a lo largo del estudio.</p>

Ítem CASPe estudio: ITT mITT APP,AT?	Respuesta	Frase literal del artículo	Justificación
<p><b>4. ¿Son importantes las pérdidas ocurridas durante el estudio?</b> ¿Difieren según el grupo? ¿Las pérdidas podrían depender de su valor o resultado? ¿Se hace análisis de sensibilidad?</p>	No	<p>“These hypotheses were tested using multilevel random coefficient models [...] incorporated a random intercept for each couple to adjust for the dyadic nature of the data and a random intercept for each participant to adjust for multiple measurements over time..” ; “Planned post hoc comparisons testing differences between conditions at 3-, 6-, 9-, and 12-months ultimately served as the direct test for an intervention effect.”</p> <p><b>¿Son importantes las pérdidas?:</b> “Follow-up retention rates were computed separately for each partner, which ranged from 77% to 96%.”</p> <p><b>¿Difieren según el grupo?:</b> No se reporta.</p> <p><b>¿Las pérdidas podrían depender de su valor o resultado?:</b> No se reporta.</p> <p><b>¿Se hace análisis de sensibilidad?:</b> No se reporta ningún análisis de sensibilidad (worst-case/best-case) en el artículo.</p>	<p>Las pérdidas fueron moderadas y aceptables para un ensayo longitudinal, con tasas de retención entre 77% y 96%. No se reportan diferencias explícitas entre grupos en las tasas de seguimiento, lo que reduce el riesgo de sesgo diferencial. No se realizó análisis de sensibilidad, pero dado el bajo y equilibrado porcentaje de pérdidas, es poco probable que estas modifiquen de forma relevante los resultados. Por ello, se considera que las pérdidas no son importantes para la validez del estudio.</p>
<p><b>5. ¿Fue adecuada la medición de los desenlaces?</b> Tipo de desenlace medido y método usado. Cegamiento del paciente, clínico, evaluador, estadístico.</p>	No	<p><b>¿Se midieron los resultados de la misma manera en ambos grupos?:</b> “Assessments were conducted at baseline and at 3, 6, 9, and 12 months.”</p> <p><b>Tipo de desenlace medido y método usado:</b> “We created a composite index of HIV vulnerability using participant self-reported data on their sexual behaviors, PrEP usage, and viral suppression.” ; “Participants were</p>	<p>Los desenlaces fueron medidos con instrumentos validados y utilizados de forma estandarizada en ambos grupos. Sin embargo, el estudio se realizó sin enmascaramiento y la mayoría de los desenlaces fueron autorreportados, lo que introduce un riesgo relevante de sesgo de información. No se describen procedimientos de cegamiento del evaluador ni verificación objetiva sistemática para los desenlaces conductuales. Por tanto, aunque los instrumentos son adecuados, la medición de</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p>Si hay problema, ¿es diferencial entre los grupos?</p>		<p>considered to have had an incident STI if they reported a positive result ... in the past 3 months.” ; “Participants completed the relationship stigma scale ... showed good internal reliability (<math>\alpha = 0.90</math> for trans women; <math>\alpha = 0.85</math> for partners of trans women).” ; “Participants completed a 7-item scale ... (<math>\alpha = 0.86</math>) goal congruence .” ; “Participants completed the conflict resolution style scale ... (<math>\alpha = 0.83</math>; <math>\alpha = 0.89</math>; <math>\alpha = 0.78</math>).”</p>	<p>los desenlaces no puede considerarse plenamente adecuada desde el punto de vista metodológico.</p>
<p><b>6. ¿Se evitó la comunicación selectiva de resultados?</b> (Mirar el registro de ensayos) ¿Hay reporte selectivo de desenlaces o reporte selectivo de análisis?</p>	<p>No sé</p>	<p><b>¿Cegamiento del evaluador/paciente?:</b> “This study was conducted without any masking.”</p> <p><b>¿Hay medición diferencial entre grupos?:</b> No se describe medición diferencial entre grupos en el artículo.</p> <p><b>¿Está registrado el ensayo?:</b> El artículo no informa explícitamente sobre un número de registro de ensayo clínico.</p> <p><b>¿Se reportan todos los desenlaces preespecificados?:</b> “It Takes Two did not have an impact on HIV vulnerability, communication, or goal congruence; however, there were reductions in relationship stigma reported by partners of trans women.” [...] “Although our intervention aimed to reduce HIV vulnerability, our analysis did not show significant differences between the intervention and control groups.”</p>	<p>El artículo no informa explícitamente sobre un registro prospectivo del ensayo ni sobre desenlaces preespecificados, lo que limita la capacidad de descartar comunicación selectiva de resultados. Aunque los autores reportan tanto resultados negativos (sin efecto en vulnerabilidad VIH, comunicación y metas) como positivos (reducción de estigma relacional) y discuten las limitaciones del estudio, la ausencia de un registro previo impide confirmar que no haya existido reporte selectivo de desenlaces o de análisis. Por tanto, no puede asegurarse que se haya evitado completamente este sesgo.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>7. ¿Cuál es el efecto del tratamiento para cada desenlace?</b>  ¿Qué desenlaces se han medido?  Detalla los positivos y los negativos.</p>	Sí	<p><b>¿Hay reporte selectivo de análisis?:</b>  “Although our findings regarding HIV vulnerability and mechanisms of change were not statistically significant, the ‘It Takes Two’ intervention resulted in a significant reduction in relationship stigma reported by partners of trans women.”</p> <p><b>¿Cuál fue el tamaño del efecto del tratamiento?:</b>  “Although our intervention aimed to reduce HIV vulnerability, our analysis did not show significant differences between the intervention and control groups.”</p> <p><b>¿Qué desenlaces se han medido?:</b>  “Specifically, intervention participants would have a lower prevalence of composite vulnerability for HIV, any condomless sex, any condomless sex with an outside partner, a lower incidence of STIs, a higher group mean for goal congruence and conflict engagement communication and positive communication, and a lower group mean for withdrawal communication and relationship stigma relative to control group participants...”</p> <p><b>Desenlaces con efecto nulo:</b>  “‘It Takes Two’ did not have an impact on HIV vulnerability, communication, or goal congruence.”</p> <p><b>Desenlace con efecto positivo:</b></p>	<p>El tamaño del efecto del tratamiento fue nulo o no significativo para el desenlace principal y para la mayoría de los desenlaces secundarios, sin diferencias significativas entre grupos. Los desenlaces evaluados están claramente definidos y reportados. El único efecto positivo observado fue una reducción significativa del estigma relacional en las parejas de mujeres trans. En conjunto, el efecto del tratamiento está claramente descrito para cada desenlace, cumpliendo el criterio CASPe de este ítem.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>8. ¿Cuál es la precisión de los estimadores del efecto?</b></p> <p>¿Cuáles son sus intervalos de confianza?</p>	No	<p>“The ‘It Takes Two’ intervention resulted in a significant reduction in relationship stigma reported by partners of trans women.”</p> <p><b>¿Cuáles son sus intervalos de confianza?:</b> “OR=26.18 (90% CI: 2.33-294.60), p=0.027.”</p> <p><b>¿Cuál es la precisión / potencia?:</b> “We were permitted to reduce our enrollment goals due to the pandemic, which limited our statistical power.”</p> <p><b>Limitación de tamaño muestral:</b> “Our study lacked sufficient power to rigorously evaluate mediation effects.”</p> <p><b>Dificultad para medir con precisión:</b> “Collecting biomedical data, such as viral load and PrEP adherence, was difficult because participants were unable or unwilling to provide hair samples or visit labs.”</p>	<p>El estudio sí reporta intervalos de confianza, por ejemplo, OR = 26.18 (90% CI: 2.33–294.60), pero estos son muy amplios, lo que indica baja precisión de la estimación del efecto. Además, los autores reconocen una limitación importante de potencia estadística por la reducción del tamaño muestral durante la pandemia. En conjunto, aunque se informan intervalos de confianza, la precisión de los estimadores es limitada y no cumple adecuadamente los criterios CASPe.</p>
<p><b>9. ¿Pueden aplicarse estos resultados en tu medio o población local?</b></p> <p>¿Crees que los pacientes incluidos en el ensayo son demasiado distintos a tus pacientes?</p> <p>¿Hay algún otro ensayo parecido a este?</p>	No	<p><b>¿Pueden aplicarse estos resultados en tu medio o población local?:</b> “our findings may not be generalizable to trans women and their partners in other geographic regions and settings.”</p> <p><b>Diferencia de población:</b> “our findings may not be generalizable to trans women and their partners in other geographic regions and settings.”</p>	<p>La población estudiada es muy específica (parejas con al menos una mujer trans, principalmente en California), y los propios autores reconocen que los resultados pueden no ser generalizables a otros contextos geográficos o poblaciones. Existe un ensayo piloto previo similar (CHIP), que constituye el antecedente directo de <i>It Takes Two</i> (referencia 21), pero con resultados distintos, ya que el estudio piloto mostró reducciones en conductas de riesgo mientras que este ensayo no mostró efecto en vulnerabilidad al VIH. Por tanto, la aplicabilidad de los resultados a otros medios o poblaciones es limitada según los criterios CASPe.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
¿Es consistente con este?		<p><b>Otros ensayos similares:</b>  <i>“A pilot study found that CHIP was feasible, acceptable, and significantly reduced condomless sex acts...”</i></p>	
<p><b>10. ¿Se han tenido en cuenta todos los resultados y su importancia clínica?</b></p>	Sí	<p><b>Consistencia de resultados:</b>  <i>“It Takes Two did not have an impact on HIV vulnerability, communication, or goal congruence; however, there were reductions in relationship stigma...”</i></p> <p><b>Resultados e importancia clínica:</b>  <i>“While the intervention did not significantly impact HIV vulnerability, communication, or shared goals, there were reductions in relationship stigma among partners of transgender women...”</i></p>	<p>Los autores informan de forma completa tanto de los desenlaces principales sin efecto (vulnerabilidad al VIH, comunicación y congruencia de metas) como del único desenlace con efecto positivo (reducción del estigma relacional). Además, discuten de forma explícita las principales limitaciones del estudio, incluyendo la falta de potencia estadística y las dificultades para medir algunos desenlaces con precisión. Sin embargo, no se analizan de forma explícita costes, preferencias de los pacientes ni el impacto económico del tratamiento. En conjunto, aunque se presentan de manera transparente los principales beneficios y limitaciones, no se han tenido plenamente en cuenta todos los aspectos necesarios para valorar de forma completa la importancia clínica global de los resultados, según los criterios CASPe.</p>
Utilidades y disutilidades de cada desenlace.		<p><b>Utilidades y disutilidades de cada desenlace:</b></p>	
Balance de efectos positivos/ negativos.		<p><i>“The ‘It Takes Two’ intervention resulted in a significant reduction in relationship stigma reported by partners of trans women.”</i></p>	
Preferencias del paciente, costes, etc.		<p><i>“Although our intervention aimed to reduce HIV vulnerability, our analysis did not show significant differences between the intervention and control groups.”</i></p> <p><b>Balance de efectos positivos/negativos:</b>  <i>“It Takes Two did not have an impact on HIV vulnerability, communication, or goal congruence; however, there were reductions in relationship stigma...”</i></p>	

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>11. ¿Los beneficios que se espera obtener justifican los riesgos y los costes?</b></p> <p>Es improbable que pueda deducirse solo de un ensayo, pero ¿qué piensas tú al respecto?</p>	No	<p><b>Preferencias del paciente, etc.:</b> no consta.</p> <p><b>Beneficio:</b>            “The ‘It Takes Two’ intervention resulted in a significant reduction in relationship stigma reported by partners of trans women.”</p> <p><b>Riesgos / ausencia de daño clínico:</b>            “It Takes Two” did not have an impact on HIV vulnerability, communication, or goal congruence; however, there were reductions in relationship stigma reported by partners of trans women”</p> <p><b>Costes:</b> no consta.</p> <p><b>Balance global:</b>            “Although our intervention aimed to reduce HIV vulnerability, our analysis did not show significant differences between the intervention and control groups; however, there were reductions in relationship stigma...”</p>	<p>El estudio muestra un beneficio clínicamente relevante limitado a un único desenlace (reducción significativa del estigma relacional), sin efecto sobre el desenlace principal ni sobre la mayoría de los desenlaces secundarios. No se describen efectos adversos relevantes, pero no se realiza ningún análisis formal de riesgos, ni se evalúan costes, carga del tratamiento ni preferencias de los pacientes. Por tanto, aunque existe un beneficio puntual, no hay información suficiente para concluir que los beneficios esperados justifican los riesgos y los costes, según los criterios CASPe.</p>

Apéndice E. Evaluación de la calidad metodológica del estudio de Whitton et al. (2016) según CASPe

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>1. ¿Se orienta el ensayo a una pregunta claramente definida?</b></p> <p>Una pregunta debe definirse en términos de la población de estudio, la intervención realizada, la comparación y los desenlaces considerados (positivos y negativos). Escribe los desenlaces.</p>	Sí	<p><i>“Participants were 20 male same-sex couples in cohabiting, committed relationships of at least 6 months.”</i></p> <p><i>“Groups were then randomized to either an immediate intervention condition (two groups with a total of 11 couples) or a waitlist condition with delayed treatment after a 4-week wait period (three groups with a total of 12 couples).”</i></p> <p><i>“At each assessment, we measured three risk factors targeted by the program (couple communication, perceived stress, social support), as well as two indices of global relationship quality (satisfaction and instability). To obtain observational measures of couple communication, the videotaped couple problem discussions were coded using the Interactional Dimensions Coding System..”</i></p>	<p>El estudio define claramente los cuatro componentes PICO: población (20 parejas masculinas del mismo sexo en relaciones estables <math>\geq 6</math> meses), intervención (programa educativo grupal SSSR), comparación (grupo control en lista de espera) y desenlaces, cumpliendo los criterios CASPe de pregunta claramente formulada.</p> <p>Los desenlaces están explícitamente enumerados y medidos con instrumentos validados (observación estructurada y escalas estandarizadas).</p>
<p><b>2. ¿Fue aleatoria la asignación a los tratamientos?</b></p> <p>¿Se generó adecuadamente la secuencia? ¿Se mantuvo oculta la secuencia de aleatorización?</p>	Sí	<p><b>Asignación aleatoria:</b> “Groups were then randomized to either an immediate intervention condition ... or a waitlist condition with delayed treatment after a 4-week wait period.”</p> <p><b>Generación de la secuencia:</b> No descrita en el artículo.</p> <p><b>Ocultación de la secuencia:</b> No descrita en el artículo.</p>	<p>La asignación a los tratamientos fue aleatoria, ya que los autores indican explícitamente que los grupos fueron “randomized” a intervención inmediata o lista de espera. No se describe el método de generación de la secuencia ni la ocultación de la misma, por lo que la generación y la ocultación de la secuencia presentan incertidumbre y potencial riesgo de sesgo de selección. Los grupos fueron comparables en línea basal, sin diferencias significativas en variables demográficas ni</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
¿Son iguales en línea basal?		<b>Línea basal:</b> “Conditions did not differ on any demographic (age, income, education, or race), or relationship variable (length, marital status, baseline satisfaction, or couple communication; all <i>ps</i> > .05).	relacionales relevantes, lo que sugiere que la aleatorización fue eficaz para equilibrar los grupos.
<p><b>3. ¿Se mantuvo la comparabilidad de los grupos a través del estudio?</b></p> <p>Desviaciones de la intervención por problemas en la asignación.</p> <p>Desviaciones de la intervención por problemas en la adhesión al tratamiento.</p> <p>¿Cómo se analizó el estudio: ITT mITT APP,AT?</p>	No	<p><b>Comparabilidad a lo largo del estudio:</b> “Three couples dropped out during the wait period; therefore, final group sizes ranged from three to six couples, with 11 couples in the immediate treatment condition and 9 couples in the waitlist condition.”</p> <p><b>Desviaciones por problemas en la asignación / pérdidas:</b> “Two couples in the immediate treatment condition were unable to be scheduled for the 3-month follow-up assessment; one broke up and one stopped responding to our contact.”</p> <p><b>Desviaciones por problemas en la adhesión al tratamiento:</b> “All 20 couples attended the full 10-hr refined SSSR program for male couples.”</p> <p><b>Análisis y manejo de datos faltantes (tipo de análisis):</b> “We used multiple imputation to handle missing data in analyses involving follow-up data., one strategy recommended for clinical trials with missing values”</p>	<p>La comparabilidad entre los grupos no se mantuvo completamente a lo largo del estudio.</p> <p>Se produjeron pérdidas diferenciales durante el periodo de espera y en el seguimiento, con tamaños finales diferentes entre condiciones, lo que introduce riesgo de sesgo de desgaste y compromete la comparabilidad inicial.</p> <p>No hubo desviaciones por falta de adhesión, ya que todos los participantes asignados a la intervención completaron el programa.</p> <p>No se declara un análisis por intención de tratar (ITT). El uso de imputación múltiple sugiere un enfoque mITT, pero no permite confirmar un ITT estricto.</p> <p>Por estos motivos, según los criterios de Cabello, no puede considerarse que la comparabilidad se haya mantenido a lo largo del estudio.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>4. ¿Son importantes las pérdidas ocurridas durante el estudio?</b>            ¿Difieren según el grupo?            ¿Las pérdidas podrían depender de su valor o resultado?            ¿Se hace análisis de sensibilidad?</p>	Sí	<p><b>Magnitud de las pérdidas:</b>            “Three couples dropped out during the wait period; therefore, final group sizes ranged from three to six couples, with 11 couples in the immediate treatment condition and 9 couples in the waitlist condition.”</p> <p><b>Diferencias según grupo:</b>            “...with 11 couples in the immediate treatment condition and 9 couples in the waitlist condition“.</p> <p><b>Pérdidas en seguimiento:</b>            “Two couples in the immediate treatment condition were unable to be scheduled for the 3-month follow-up assessment; one broke up and one stopped responding to our contact.”</p> <p><b>Análisis de sensibilidad:</b> no consta en el artículo.</p> <p><b>Manejo de datos faltantes:</b>            “We used multiple imputation to handle missing data in analyses involving follow-up data, one strategy recommended for clinical trials with missing values ”</p>	<p>Se produjeron pérdidas relevantes durante el periodo de espera y en el seguimiento, con una proporción no despreciable respecto al tamaño muestral inicial (5 de 20 parejas).</p> <p>Las pérdidas fueron diferenciales según el grupo, con tamaños finales distintos entre condiciones (11 parejas en el grupo de intervención y 9 en el grupo control), concentrándose inicialmente en el grupo control (waitlist) durante el periodo de espera.</p> <p>Una de las pérdidas se debió a la ruptura de la pareja, lo que sugiere que las pérdidas podrían estar relacionadas con el desenlace principal (inestabilidad relacional), introduciendo riesgo de sesgo de desgaste (attrition bias) potencialmente relacionado con el desenlace.</p> <p>No se declara un análisis de sensibilidad formal para evaluar el impacto de las pérdidas. Se utilizó imputación múltiple para el manejo de datos faltantes, pero no se presentan análisis alternativos que permitan confirmar la robustez de los resultados.</p> <p>Por estos motivos, según los criterios de Cabello, las pérdidas deben considerarse importantes.</p>
<p><b>5. ¿Fue adecuada la medición de los desenlaces?</b>            Tipo de desenlace medido y método usado.             Cegamiento del paciente,</p>	No	<p><b>Adecuación de la medición de desenlaces:</b>            “Intercoder reliability, assessed on 27% of the videos, was high (intraclass correlation coefficients = .85 and .87 for positive and negative communication, respectively).”</p> <p><b>Tipo de desenlaces y métodos de medición:</b></p>	<p>Los desenlaces son relevantes y están medidos con instrumentos validados (escalas estandarizadas y codificación observacional con alta fiabilidad), lo que apoya una medición técnicamente adecuada.</p> <p>Sin embargo, no se declara ningún procedimiento de cegamiento: los participantes conocen su asignación, los terapeutas conocen la intervención y no se especifica</p>

<b>Ítem CASPe</b>	<b>Respuesta</b>	<b>Frase literal del artículo</b>	<b>Justificación</b>
<p>clínico, evaluador, estadístico. Si hay problema, ¿es diferencial entre los grupos?</p>		<p>“At each assessment, we measured three risk factors targeted by the program (couple communication, perceived stress, social support), as well as two indices of global relationship quality (satisfaction and instability).”</p>	<p>cegamiento de los evaluadores de los vídeos ni del analista estadístico.</p>
		<p><b>Medición observacional de comunicación:</b> “The videotaped couple problem discussions were coded using the Interactional Dimensions Coding System...”</p>	<p>Dado que varios desenlaces son subjetivos y autorreportados (estrés, apoyo social, satisfacción, inestabilidad), la ausencia de cegamiento introduce riesgo de sesgo de detección y de respuesta.</p>
		<p><b>Fiabilidad de los evaluadores:</b> “Intercoder reliability, assessed on 27% of the videos, was high (intraclass correlation coefficients = .85 and .87 for positive and negative communication, respectively)”</p>	<p>No hay indicios de que este problema sea diferencial entre grupos, ya que ambos grupos fueron evaluados con los mismos instrumentos y procedimientos.</p>
		<p><b>Desenlaces autorreportados:</b> “Participants self-reported on couple communication ... perceived stress ... social support ... relationship satisfaction ... relationship instability.”</p>	<p>Por estos motivos, según los criterios de Cabello, la medición de los desenlaces no puede considerarse plenamente adecuada.</p>
		<p><b>Cegamiento (paciente/clínico – intervención grupal):</b> “Each group session was led by two facilitators”</p>	
		<p><b>Cegamiento (evaluador/estadístico):</b></p>	

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>6. ¿Se evitó la comunicación selectiva de resultados?</b> (Mirar el registro de ensayos) ¿Hay reporte selectivo de desenlaces o reporte selectivo de análisis?</p>	No sé	<p>No se declara cegamiento de evaluadores ni del analista estadístico en el artículo.</p> <p><b>Comunicación selectiva / registro de ensayo:</b> No se menciona en el artículo ningún número de registro de ensayo ni la existencia de un protocolo previo publicado.</p> <p><b>Desenlaces preespecificados en métodos:</b> “At each assessment, we measured three risk factors targeted by the program (couple communication, perceived stress, social support), as well as two indices of global relationship quality (satisfaction and instability).”</p> <p><b>Reporte de resultados principales:</b> “Results, displayed in Table 1, indicated within-person changes in three of four targeted risk factors over the course of the inter- vention (i.e., from before to immediately after the program).”</p> <p><b>Comparación intervención vs. control:</b> “Compared to waitlist controls, participants had significantly less negative communication and more positive communication than waitlist controls, as rated by observers (large effects) and by self-report (small to medium sized effects). Perceived stress was significantly lower and</p>	<p>No se dispone de un registro de ensayo ni de un protocolo previo que permita comprobar si los desenlaces y análisis publicados coinciden con los inicialmente planificados.</p> <p>El artículo informa resultados para todos los desenlaces descritos en el apartado de métodos, y no se identifican desenlaces claramente anunciados que no hayan sido posteriormente reportados.</p> <p>Sin embargo, en ausencia de un registro o protocolo accesible, no es posible confirmar ni descartar la existencia de comunicación selectiva de resultados o de análisis.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>7. ¿Cuál es el efecto del tratamiento para cada desenlace?</b> ¿Qué desenlaces se han medido?</p> <p>Detalla los positivos y los negativos.</p>	Sí	<p>relationship satisfaction higher in the immediate treatment than waitlist condition (medium and small effect sizes, respectively). Social support and relationship instability did not differ by condition.”</p> <p><b>Efecto del tratamiento:</b> “Compared to waitlist controls, participants had significantly less negative communication and more positive communication than waitlist controls, as rated by observers (large effects) and by self-report (small to medium sized effects). Perceived stress was significantly lower and relationship satisfaction higher in the immediate treatment than waitlist condition (medium and small effect sizes, respectively). Social support and relationship instability did not differ by condition.”</p> <p><b>Desenlaces medidos:</b> “At each assessment, we measured three risk factors targeted by the program (couple communication, perceived stress, social support), as well as two indices of global relationship quality (satisfaction and instability).”</p> <p><b>Efectos positivos (pre–post):</b> “Participants engaged in less destructive and more positive communication immediately</p>	<p>Se evaluaron cinco desenlaces principales: comunicación de pareja, estrés percibido, apoyo social, satisfacción relacional e inestabilidad relacional. En la comparación con el grupo control (waitlist), la intervención mostró efectos positivos claros sobre la comunicación (reducción de comunicación negativa e incremento de comunicación positiva), el estrés percibido y la satisfacción relacional. En contraste, no se observaron diferencias entre grupos en apoyo social ni en inestabilidad relacional. En los análisis pre–post (within-person), la comunicación y el estrés mejoraron de forma inmediata tras la intervención y estos efectos se mantuvieron en el seguimiento a los 3 meses. El apoyo social, que no mostró mejoría inmediata, presentó una mejoría tardía significativa en el seguimiento. De forma similar, la satisfacción relacional y la inestabilidad relacional no mejoraron de forma inmediata al finalizar el programa, pero ambas mostraron una mejoría significativa a los 3 meses en el análisis longitudinal. No obstante, es importante distinguir que, aunque la inestabilidad relacional mostró una mejoría tardía en el análisis within-person, no se confirmaron diferencias significativas frente al grupo control, por lo que este desenlace debe considerarse como de efecto nulo en la comparación principal entre grupos.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
8. ¿Cuál es la precisión de los estimadores del efecto?	No	<p>postintervention than prior to the intervention.”</p> <p>“Perceived stress also decreased significantly from pre- to postprogram (<math>p &lt; .01</math>; small effect size).”</p> <p><b>Efectos positivos mantenidos / tardíos:</b></p> <p>“These gains were maintained at follow-up: all measures of couple communication and perceived stress remained significantly better than at preprogram (<math>ps &lt; .05</math>).”</p> <p>“Social support, which did not improve from pre- to immediate postprogram, did improve by 3-month follow-up (<math>p = .01</math>).”</p> <p>“...by the 3-month follow-up, both [satisfaction and instability] were significantly improved (<math>ps = .04</math> and <math>.04</math>, respectively; small effect sizes).”</p> <p><b>Efectos negativos o nulos (vs control):</b></p> <p>“Social support and relationship instability did not differ by condition.”</p> <p><b>Sin efecto inmediato (pre–post):</b></p> <p>“Neither index of relationship quality (satisfaction or instability) improved significantly by the immediate postprogram assessment...”</p> <p><b>Precisión / tamaño del efecto:</b></p> <p>“Cohen’s d effect sizes, which describe the difference between means in standard</p>	<p>En conjunto, el tratamiento presenta efectos positivos consistentes sobre procesos intermedios clave (comunicación y estrés), efectos tardíos sobre algunos desenlaces de calidad relacional y efectos nulos sobre la inestabilidad relacional en la comparación principal entre grupos.</p> <p>El estudio informa tamaños del efecto (Cohen’s d) para los principales desenlaces, tanto en los análisis pre–post como en la comparación entre grupos, lo que permite</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
¿Cuáles son sus intervalos de confianza?		<p>deviation units, were used to interpret the magnitude of differences.”</p> <p>“Effect sizes were larger for observational methods (ds = .67–.71; medium to large effects), than self-reports (ds = .39–.48; small to medium effects).”</p> <p><b>Intervalos de confianza:</b> No se reportan intervalos de confianza para ninguno de los desenlaces ni para los tamaños del efecto en el artículo.</p>	<p>valorar la magnitud clínica de los efectos observados. Sin embargo, no se presentan intervalos de confianza para ninguno de los desenlaces ni para los tamaños del efecto, ni en las Tablas 1 y 2 ni en el texto de resultados. En ausencia de intervalos de confianza, no es posible valorar la precisión de los estimadores ni la incertidumbre asociada a las estimaciones. Por otro lado, no se reportan intervalos de confianza, la precisión de los estimadores del efecto no puede considerarse adecuadamente evaluada, por lo que la respuesta debe ser No.</p>
<p><b>9. ¿Pueden aplicarse estos resultados en tu medio o población local?</b></p> <p>¿Crees que los pacientes incluidos en el ensayo son demasiado distintos a tus pacientes?</p> <p>¿Hay algún otro ensayo parecido a este?</p> <p>¿Es consistente con este?</p>	No sé	<p><b>Aplicabilidad / pertinencia clínica:</b></p> <p>“Consequently, practitioners are increasingly likely to encounter same-sex couples seeking relationship-focused services.”</p> <p><b>Población incluida / similitud con tu población:</b></p> <p>“Participants were 20 male same-sex couples in cohabiting, committed relationships of at least 6 months.”</p> <p>“The sample was largely White (83%; 10% Black, 2.5% Hispanic, 2.5% multiracial).”</p> <p>“Mean years of education was 16 (SD = 2.19); 93% had completed college.”</p> <p><b>Contexto del estudio:</b></p> <p>“...a sample of 20 committed male same-sex couples living in the midwestern United States...”</p>	<p>El estudio aborda una intervención relevante para la práctica con parejas del mismo sexo y los autores argumentan que la demanda de este tipo de servicios está aumentando, lo que apoya la pertinencia del tema. Sin embargo, la aplicabilidad directa depende mucho de si tu población se parece a la muestra: parejas masculinas, mayoritariamente blancas, con nivel educativo alto, y además con un perfil probablemente poco representativo de parejas más diversas o con mayor nivel de malestar. El contexto sociocultural (EE. UU. y estados con menor aceptación/legalidad) también puede no equivaler al entorno local. Existe evidencia previa tipo piloto(ref. 5 del artículo) en una muestra similar que sugiere coherencia y apoyo externo inicial, pero el propio artículo reconoce limitaciones (tamaño muestral pequeño y necesidad de muestras más representativas). Por ello, la transferibilidad a otros contextos/poblaciones no puede asegurarse con certeza y la respuesta más ajustada es “No sé”.</p>

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
10. ¿Se han tenido en cuenta todos los	No	<p>“...expanding the population of interest to couples living in states with less social acceptance of and legal protections for sexual minorities.”</p>	
		<p><b>Otros estudios parecidos / consistencia externa:</b>            “An initial pilot study of SSSR with 12 married or engaged male same-sex couples in Boston, Massachusetts (Buzzella, Whitton, &amp; Tompson, 2012), provided preliminary data that supported program feasibility and acceptability, and informed program refinement. ”</p>	
		<p><b>Limitaciones de generalización reconocidas por los autores:</b>            “First, the small sample size may have limited power to detect program effects...”            “Third, participants were mostly White, of moderate to high socioeconomic status, and satisfied with their relationships.”            “Future studies using larger, more representative samples are needed to powerfully evaluate the program’s effectiveness with couples who are more diverse and who are more severely distressed”</p>	
<p><b>Resultados positivos y utilidades:</b>            “Results indicated significant program effects in comparison to waitlist controls on</p>	<p>El estudio informa de forma detallada sobre los beneficios clínicos de la intervención en múltiples desenlaces relevantes (comunicación de pareja, estrés</p>		

<b>Ítem CASPe</b>	<b>Respuesta</b>	<b>Frase literal del artículo</b>	<b>Justificación</b>
<b>resultados y su importancia clínica?</b> Utilidades y disutilidades de cada desenlace. Balance de efectos positivos/ negativos. Preferencias del paciente, costes, etc.		<p>couple constructive and destructive communication, perceived stress, and relationship satisfaction.”</p> <p>“Participating couples reported high satisfaction with the program, which showed immediate, sustained positive effects on two key risk factors—couple communication behaviors and perceived stress—as well as effects on social support, relationship satisfaction, and relationship instability that were evident by 3-month follow-up.”</p>	<p>percibido, satisfacción y estabilidad de la relación) y sobre la alta aceptabilidad del programa por parte de los participantes, lo que permite valorar adecuadamente las utilidades de la intervención y las preferencias de los pacientes.</p> <p>Sin embargo, no se analizan ni se reportan disutilidades, efectos negativos, posibles riesgos, ni cargas para los participantes, ni se presentan datos sobre costes económicos, costes de implementación o análisis de coste-efectividad. Tampoco se realiza un balance explícito entre beneficios y posibles daños de la intervención.</p>
		<p><b>Aceptabilidad / preferencias del paciente:</b>            “...men rated the overall program as very helpful in maintaining relationship satisfaction (<math>M = 6.20</math> <math>SD = 1.64</math>)”</p> <p>“Participating couples reported high satisfaction with the program...”</p> <p><b>Ausencia de análisis de disutilidades, costes o efectos negativos:</b>            No se describen efectos adversos, cargas, costes económicos ni análisis de balance beneficio–daño.</p>	<p>Al no considerar de forma sistemática todos los componentes relevantes para valorar la importancia clínica según los criterios de este ítem —especialmente las disutilidades, los riesgos y los costes—, no puede afirmarse que se hayan tenido en cuenta todos los resultados y su importancia clínica, por lo que la respuesta más adecuada es <b>No</b>.</p> <p>Aunque se describen algunos aspectos logísticos del programa (duración, número de sesiones, participación de dos terapeutas y compensación económica por evaluación), estos datos no se integran en un análisis clínico de balance beneficio–daño ni en una valoración de la carga o el coste para los pacientes o el sistema.</p>
<b>11. ¿Los beneficios que se espera obtener justifican los riesgos y los costes?</b>	No	<p>Beneficios / eficacia y aceptabilidad:            “Overall, findings from this small randomized waitlist controlled trial were supportive of the feasibility, efficacy, and</p>	<p>El estudio documenta beneficios clínicos claros, sostenidos y superiores al grupo control en múltiples desenlaces relevantes (comunicación, estrés, satisfacción y estabilidad), así como una alta aceptabilidad de la intervención, lo que apoya la</p>

<b>Ítem CASPe</b>	<b>Respuesta</b>	<b>Frase literal del artículo</b>	<b>Justificación</b>
Es improbable que pueda deducirse solo de un ensayo, pero ¿qué piensas tú al respecto?		<p data-bbox="775 236 1346 308">acceptability of the SSSR Program for Male couples.”</p> <p data-bbox="775 347 1346 671">Magnitud y relevancia de los beneficios:  “Participating couples reported high satisfaction with the program, which showed immediate, sustained positive effects on two key risk factors —couple communication behaviors and perceived stress— as well as effects on social support, relationship satisfaction, and relationship instability that were evident by 3-month follow-up.”</p>	<p data-bbox="1384 236 2107 308">existencia de un beneficio clínico potencial del programa.</p> <p data-bbox="1384 311 2107 528">Sin embargo, el artículo no informa sobre la existencia o ausencia de riesgos, efectos adversos, daños ni cargas para los participantes, ni presenta ningún análisis de costes, coste-efectividad o recursos necesarios para su implementación. Tampoco se realiza un balance explícito entre beneficios, riesgos y costes.</p> <p data-bbox="1384 531 2107 671">Por tanto, aunque los beneficios parecen clínicamente relevantes, no es posible concluir que justifiquen los riesgos y los costes, ya que estos no se evalúan de forma sistemática en el estudio.</p>

Apéndice F. Evaluación de la calidad metodológica del estudio de Whitton et al. (2017) según CASPe

Ítem CASPe	Respuesta	Frase literal del artículo	Justificación
<p><b>1. ¿Se orienta el ensayo a una pregunta claramente definida?</b></p> <p><i>Una pregunta debe definirse en términos de la población de estudio, la intervención realizada, la comparación y los desenlaces considerados (positivos y negativos). Escribe los desenlaces.</i></p>	Sí	<p><b>Pregunta / diseño del estudio:</b> “A new culturally sensitive adaptation of relationship education was developed specifically for female same-sex couples (The Strengthening Same-Sex Relationships Pro- gram, Female version; SSSR-F).” “A small randomized waitlist-control trial (N = 37 couples) was conducted to evaluate program feasibility, acceptability, and efficacy.”</p> <p><b>Desenlaces evaluados:</b> “Three proximal outcomes targeted by SSSR-F (communication, perceived stress, social support) and three distal outcomes (global relationship satisfaction, instability, and confidence) were assessed.”</p>	<p><b>P</b> = Parejas mujer–mujer en relación comprometida <math>\geq</math> 6 meses. <b>I</b> = Programa SSSR-F (educación relacional adaptada). <b>C</b> = Grupo control en lista de espera (waitlist). <b>O</b> = Comunicación, satisfacción relacional, confianza relacional, inestabilidad relacional, estrés percibido, apoyo social. El estudio evalúa seis desenlaces: comunicación, satisfacción relacional, confianza, inestabilidad, estrés y apoyo social. La comunicación y la satisfacción relacional La pregunta está claramente definida en términos de población, intervención, comparación y desenlaces relevantes.</p>
<p><b>2. ¿Fue aleatoria la asignación a los tratamientos?</b></p> <p>¿Se generó adecuadamente la secuencia? ¿Se mantuvo oculta la secuencia de aleatorización?</p>	Sí	<p><b>Pregunta general (asignación aleatoria):</b> “Groups were then randomized to either an immediate intervention condition or a waitlist condition with delayed treatment after a 4-week wait period.”</p> <p><b>¿Se genera adecuadamente la secuencia?:</b> no consta. <b>¿Se mantuvo oculta la secuencia?:</b> No consta. <b>¿Son iguales en línea basal?:</b> “Conditions did not differ on any demographic (age, income, education, or race), or relationship variable</p>	<p>La asignación a los tratamientos fue aleatoria, ya que los autores indican explícitamente que los grupos fueron “randomized” a intervención inmediata o a lista de espera. Sin embargo, no se describe el método de generación de la secuencia de aleatorización ni el procedimiento de ocultación de la asignación, por lo que estos aspectos constan como no descritos y existe riesgo potencial de sesgo de selección.</p>

¿Son iguales en línea basal?

(length, marital status, or baseline levels of proximal or distal couple outcomes; all p values > .05).”

No obstante, los grupos fueron comparables en línea basal, sin diferencias significativas en variables demográficas ni en los desenlaces iniciales. En conjunto, puede considerarse que la asignación fue aleatoria según los criterios CASPe, aunque con limitaciones metodológicas por falta de información sobre la secuencia y su ocultación.

**3. ¿Se mantuvo la comparabilidad de los grupos a través del estudio? Desviaciones de la intervención por problemas en la asignación.**

No

**Comparabilidad a lo largo del estudio :**

“One waitlist couple dropped out during the waiting period, and another dropped out during the intervention and did not respond to further contact attempts; these two couples were not included in analyses. One couple did not complete the immediate posttreatment assessment but did return for the 3-month follow-up assessment.”

**Desviaciones por problemas en la asignación:** No consta.

**Desviaciones por problemas en la adhesión al tratamiento:** “Received allocated intervention (n = 17)... Discontinued intervention (no reason given) (n = 1).”

**Tipo de análisis (ITT / mITT / PP):** “These two couples were not included in analyses.”

La comparabilidad entre los grupos no se mantuvo adecuadamente a lo largo del estudio. Durante el seguimiento se produjeron pérdidas diferenciales, ya que una pareja abandonó durante el periodo de espera y otra durante la intervención, y ambas fueron excluidas del análisis final. Además, se describe al menos una discontinuación del tratamiento sin motivo especificado, lo que introduce una desviación por problemas de adhesión. No se informan desviaciones relacionadas con errores en la asignación inicial. En cuanto al análisis, no se realizó un análisis por intención de tratar (ITT), ya que las parejas que abandonaron no fueron incluidas en los análisis, adoptándose un enfoque por protocolo. Por estos motivos, según los criterios CASPe, no puede considerarse que la comparabilidad de los grupos se haya mantenido adecuadamente a lo largo del estudio, principalmente por las pérdidas diferenciales y el análisis no basado en ITT.

**Desviaciones de la intervención por problemas en la adhesión al tratamiento.**

**¿Cómo se analizó el estudio: ITT mITT APP,AT?**

**4. ¿Son importantes las pérdidas**

Sí

**Magnitud de las pérdidas:**

Las pérdidas fueron cuantitativamente relevantes en relación con el tamaño muestral

ocurridas durante el estudio?

¿Difieren según el grupo?

¿Las pérdidas podrían depender de su valor

o resultado?

¿Se hace análisis de sensibilidad?

“One waitlist couple dropped out during the waiting period, and another dropped out during the intervention and did not respond to further contact attempts; these two couples were not included in analyses.”

¿Difieren según el grupo?:

“One waitlist couple dropped out during the waiting period, and another dropped out during the intervention...”

¿Pueden depender del resultado?:

“One couple did not complete the immediate posttreatment assessment but did return for the 3-month follow-up assessment.”

¿Análisis de sensibilidad?:

No consta.

(2 de 39 parejas, ≈5%) y ambas parejas fueron excluidas del análisis final. Las pérdidas fueron diferenciales según el grupo, ya que una ocurrió durante el periodo de espera y otra durante la intervención. Además, una de las pérdidas se debió a abandono durante la intervención y otra a falta de seguimiento, lo que sugiere que las pérdidas podrían estar relacionadas con el curso del tratamiento o con los desenlaces. Aunque los autores realizaron análisis con la muestra completa combinando ambos grupos (“full sample”), no se realizó ningún análisis de sensibilidad específico para evaluar el impacto de las pérdidas sobre los resultados. Por estos motivos, según los criterios CASPe, las pérdidas deben considerarse importantes.

La medición de los desenlaces puede considerarse técnicamente adecuada. El estudio utiliza instrumentos estandarizados y validados para todos los desenlaces, combinando medidas observacionales de la comunicación con alta fiabilidad interevaluador y escalas autoinformadas con buenas propiedades psicométricas para estrés percibido, apoyo social y calidad relacional. Sin embargo, el estudio no describe procedimientos de cegamiento. En una intervención grupal como SSSR-F no es posible cegar a las participantes ni a las facilitadoras, y no consta que los codificadores de los vídeos ni el analista estadístico

5. ¿Fue adecuada la medición de los desenlaces? Sí

Tipo de desenlace medido y método usado.

Cegamiento del paciente, clínico, evaluador, estadístico.

Si hay problema, ¿es diferencial entre los grupos?

**Tipo de desenlace y método usado, y problema diferencial:**

“At each assessment, we measured three proximal outcomes targeted by the program (couple communication, perceived stress, and social support), as well as three distal outcomes (global relationship satisfaction, confidence, and instability).”

**-Proximal outcomes:**

**1. Couple communication (observational):**

*Interactional Dimensions Coding System* (Kline et al., 2004).

**2. Couple communication (self-reported):**

*Communication Skills Test* (CST; Jenkins & Saiz, 1995).

**3. Perceived stress: *Perceived Stress Scale* (PSS; Cohen, Kamarck, & Mermelstein, 1983).**

<p><b>6. ¿Se evitó la comunicación selectiva de resultados? (Mirar el registro de ensayos)</b></p> <p><b>¿Hay reporte selectivo de desenlaces o reporte selectivo de análisis?</b></p>	<p>No sé</p>	<p><b>4. Social support:</b> <i>Support for the Relationship Measure</i> (SRM; Sprecher &amp; Felmlee, 1992), adapted 7-item scale.</p> <p><b>- Distal Outcomes:</b></p> <p><b>1. Relationship satisfaction:</b> 12-item measure of relationship satisfaction (McNulty &amp; Karney, 2001; Whitton et al., 2008).</p> <p><b>2. Relationship confidence:</b> <i>Confidence Scale</i> (Stanley, Hoyer, &amp; Trathen, 1994).</p> <p><b>3. Relationship instability:</b> <i>Relationship Instability Index</i> (Whitton, Kuryluk, &amp; Khaddouma, 2015), adapted from the <i>Marital Instability Index</i> (Booth, Johnson, &amp; Edwards, 1983).</p> <p><b>Cegamiento del paciente:</b> No consta.</p> <p><b>Cegamiento del clínico/facilitador:</b> No consta.</p> <p><b>Cegamiento del evaluador:</b> No consta.</p> <p><b>Cegamiento del estadístico:</b> No consta.</p> <p><b>Registro de ensayo / protocolo previo:</b> No consta.</p> <p><b>Desenlaces prespecificados en métodos:</b>      “At each assessment, we measured three proximal outcomes targeted by the program (couple communication, perceived stress, and social support), as well as three distal outcomes (global relationship satisfaction, confidence, and instability).”</p> <p><b>Reporte de resultados principales:</b>      “Results are shown in Table 2.”</p> <p><b>Plan de análisis descrito:</b> “Analyses to assess program effects on targeted couple outcomes were conducted using Mplus Version 7 with robust maximum likelihood estimation.”</p>	<p>estuvieran cegados a la condición de asignación.</p> <p>Dado que varios desenlaces son subjetivos y autoinformados (estrés, apoyo social, satisfacción, confianza e inestabilidad), la ausencia de cegamiento introduce un riesgo de sesgo de detección.</p> <p>No obstante, no hay evidencia de que este sesgo sea diferencial entre grupos, ya que los mismos instrumentos y procedimientos de evaluación se aplicaron en ambos brazos del estudio y en todos los momentos de medida. Por tanto, según los criterios de CASPe/Cabello, la medición de los desenlaces puede considerarse adecuada, aunque con un riesgo de sesgo de detección por ausencia de cegamiento, especialmente en los desenlaces subjetivos.</p> <p>No se dispone de un número de registro de ensayo ni de un protocolo previo accesible que permita comprobar si los desenlaces y análisis publicados coinciden con los inicialmente planificados. El artículo describe de forma clara en el apartado de métodos cuáles eran los desenlaces proximales y distales que se iban a evaluar y posteriormente informa resultados para todos ellos, sin que se identifiquen desenlaces claramente anunciados que no hayan sido reportados ni análisis principales omitidos. Asimismo, se describe un plan de análisis estadístico detallado y coherente con los resultados presentados.</p>
--	--------------	--	---

**7. ¿Cuál es el efecto del tratamiento para cada desenlace?**

Sí

**¿Qué desenlaces se han medido?**

**Detalla los positivos y los negativos.**

**Efecto del tratamiento:**

“Results of multilevel models accounting for nonindependence in dyadic data indicated statistically significant program effects on positive and negative couple communication, relationship satisfaction, and relationship confidence and small, nonsignificant program effects on stress, social support, and relationship instability.”

**Desenlaces medidos (proximales y distales):** “Three proximal outcomes targeted by SSSR-F (communication, perceived stress, social support) and 3 distal outcomes (global relationship satisfaction, instability, and confidence) were assessed at pre- and posttreatment and 3-month follow-up.”

**Efectos positivos (significativos):** “Results... indicated statistically significant program effects on positive and negative couple communication, relationship satisfaction, and relationship confidence...”

**Efectos negativos / no significativos:** “...and small, nonsignificant program effects on stress, social support, and relationship instability.”

Sin embargo, en ausencia de un registro prospectivo o de un protocolo público, no es posible descartar con certeza la existencia de comunicación selectiva de resultados o de análisis no preespecificados. Por este motivo, según los criterios CASPe/Cabello, no puede confirmarse que se haya evitado completamente la comunicación selectiva de resultados, y la respuesta más adecuada es “No sé”.

El estudio define claramente seis desenlaces: tres proximales (comunicación —observada y autorreportada—, estrés percibido y apoyo social) y tres distales (satisfacción de la relación, inestabilidad y confianza). Los resultados muestran que el tratamiento produce efectos positivos significativos en cuatro desenlaces: disminuye la comunicación negativa y aumenta la comunicación positiva (tamaños de efecto medios-altos,  $d \approx -0.49$  a  $-0.78$ ), y mejora de forma significativa la satisfacción de la relación ( $d \approx 0.36$ ) y la confianza ( $d \approx 0.40$ ). En cambio, se observan efectos pequeños y no significativos en estrés percibido, apoyo social e inestabilidad de la relación ( $d \approx 0.18-0.23$ ), por lo que el programa no demuestra beneficio claro en estos desenlaces. En el seguimiento a 3 meses, los efectos se mantienen en los desenlaces proximales, especialmente en comunicación, pero los desenlaces distales empeoran tras el postratamiento: la satisfacción y la confianza

**(Efecto por desenlace: 4 semanas; Tabla 2)** “Observed negative (  $z = -2.02$ ,  $p = .04$ ,  $d = -.55$ ); Observed positive (  $z = 2.26$ ,  $p = .02$ ,  $d = .78$ ); Self-reported negative; (  $z = -2.95$ ,  $p = .003$ ,  $d = -.49$ ; Self-reported positive (  $z = 3.06$ ,  $p = .002$ ,  $d = .54$ ); Perceived stress(  $z = -.98$ ,  $p = .33$ ,  $d = -.20$ ); Social support (  $z = 1.36$ ,  $p = .17$ ,  $d = .18$ ); Relationship satisfaction...  $z = 3.10$ ,  $p = .002$ ,  $d = .36$ ; Relationship instability (  $z = -1.55$ ,  $p = .12$ ,  $d = -.23$ ); and Relationship confidence (  $z = 2.22$ ,  $p = .03$ ,  $d = .40$ )”

**Mantenimiento del efecto/ deterioro:**

“Analyses of follow-up data suggest maintenance of effects on the proximal but not the distal outcomes.”

“Relationship satisfaction and confidence decreased significantly, and relationship instability increased marginally, from immediately post-program to the 3-month follow-up. Comparison of follow-up scores to preprogram scores suggest that all three distal couple outcome variables returned to baseline levels by 3 months after the program.”

disminuyen significativamente y la inestabilidad aumenta marginalmente, retornando aproximadamente a los niveles basales. Por tanto, el estudio informa de forma completa el efecto del tratamiento para cada desenlace, distinguiendo beneficios inmediatos y limitaciones en la eficacia a medio plazo.

**8. ¿Cuál es la precisión de los estimadores del efecto?**

**Sí**

**Intervalos de confianza 95% de los tamaños de efecto (Tabla 2):**

“Observed negative (communication):  $d = -.55$  [  $-.82$ ,  $-.27$  ]”

“Observed positive (communication):  $d = .78$  [  $.36$ ,  $1.20$  ]”

“Self-reported negative (communication):  $d = -.49$  [  $-.90$ ,  $-.09$  ]”

“Self-reported positive (communication):  $d = .54$  [  $.25$ ,  $.82$  ]”

¿Cuáles son sus intervalos de confianza?

El estudio informa sistemáticamente los tamaños de efecto acompañados de sus intervalos de confianza al 95% para todos los desenlaces, tanto en el efecto inmediato (Tabla 2) como en el seguimiento a 3 meses (Tabla 3). En los desenlaces con efecto significativo inmediato —comunicación observada y autorreportada, satisfacción y confianza— los IC 95% no incluyen el valor 0 y son relativamente estrechos (por ejemplo,

“Perceived stress:  $d = -.20 [-.42, .02]$ ”  
 “Social support:  $d = .18 [-.11, .47]$ ”  
 “Relationship satisfaction:  $d = .36 [.10, .62]$ ”  
 “Relationship instability:  $d = -.23 [-1.03, .58]$ ”  
 “Relationship confidence:  $d = .40 [.11, .69]$ ”

**Intervalos de confianza 95% en el seguimiento (Tabla 3, Post-Follow-up):**

“Observed negative:  $d = .18 [-.06, .42]$ ”  
 “Observed positive:  $d = -.12 [-.44, .20]$ ”  
 “Self-report negative:  $d = .20 [-.06, .47]$ ”  
 “Self-report positive:  $d = -.09 [-.28, .10]$ ”  
 “Perceived stress:  $d = .05 [-.12, .23]$ ”  
 “Social support:  $d = -.08 [-.31, .14]$ ”  
 “Satisfaction:  $d = -.34 [-.56, -.13]$ ”  
 “Instability:  $d = .21 [-.47, .89]$ ”  
 “Confidence:  $d = -.37 [-.62, -.12]$ ”

**Población y criterios de selección:**

“We recruited female same-sex couples who had been together in a committed relationship for at least 6 months.”  
 “Exclusion criteria included conditions that might impair participation (e.g., psychotic thought processes, severe substance abuse, or cognitive impairment; active suicidal or homicidal ideation; and severe physical violence between partners).”  
 “In the final sample ( $N = 74$  individuals; 37 couples)...”  
 “Conditions did not differ on any demographic (age, income, education, or race), or relationship variable (length, marital status, or baseline levels of proximal or distal couple outcomes; all  $p$  values  $> .05$ ).”

**Tipo de intervención afirmativa:**

comunicación observada positiva  $d = .78$ ; IC [.36, 1.20]), lo que indica buena precisión de los estimadores. En cambio, en estrés, apoyo social e inestabilidad, los intervalos de confianza son amplios e incluyen el 0 (por ejemplo, inestabilidad  $d = -.23$ ; IC [-1.03, .58]), lo que refleja baja precisión y elevada incertidumbre, probablemente relacionada con el tamaño muestral reducido y la baja potencia estadística del estudio. En el seguimiento, los IC muestran estabilidad en los desenlaces proximales (todos incluyen 0) y un deterioro significativo en satisfacción y confianza (IC que no incluyen 0), confirmando que la precisión es adecuada en los desenlaces principales y limitada en los secundarios.

Según los criterios de inclusión de esta revisión, el estudio es claramente aplicable, ya que incluye exclusivamente adultos  $\geq 18$  años, parejas del mismo sexo en contexto de relación de pareja y una intervención afirmativa específicamente diseñada para población LGBTQ+, con medición de desenlaces psicológicos y relacionales, cumpliendo plenamente los criterios de inclusión y no cumpliendo ninguno de los criterios de exclusión de esta revisión. La población del ensayo es conceptualmente similar a la población diana (parejas LGBTQ+ adultas en intervención de pareja), por lo que los resultados pueden aplicarse a contextos

**9. ¿Pueden aplicarse estos resultados en tu medio o población local?**  
**¿Crees que los pacientes incluidos en el ensayo son demasiado distintos a tus pacientes?**  
**¿Hay algún otro ensayo parecido a este?**  
**¿Es consistente con este?**

**Sí**

“A new culturally sensitive adaptation of relationship education was developed specifically for female same-sex couples (The Strengthening Same-Sex Relationships Program, Female version; SSSR-F). ”  
 “Toward this aim, our research team developed the Strengthening Same-Sex Relationships (SSSR) programs (Buzzella & Whitton, 2009; Whitton, Scott, & Buzzella, 2013). Consistent with guidelines for culturally sensitive program adaptation (American Psychological Association, 2003; Kumpfer, Alvarado, Smith, & Bellamy, 2002)...”

**Consistencia externa:**

“The male version of the program, SSSR-M (Buzzella & Whitton, 2009), has shown evidence of feasibility, acceptability, and efficacy in improving communication and relationship satisfaction among male same-sex couples (Buzzella, Whitton, & Tompson, 2012; Whitton, Weitbrecht, Kuryluk, & Hutsell, 2016).”

**Resultados y balance de efectos:**

“Results of multilevel models accounting for nonindependence in dyadic data indicated statistically significant program effects on positive and negative couple communication, relationship satisfaction, and relationship confidence and small, nonsignificant program effects on stress, social support, and relationship instability. ”

clínicos o comunitarios similares. No obstante, la aplicabilidad es limitada porque la muestra incluye únicamente parejas con al menos 6 meses de relación, excluye situaciones de alta gravedad (violencia severa, psicosis, abuso grave de sustancias) y está compuesta mayoritariamente por mujeres blancas, con alto nivel educativo y relaciones relativamente estables, lo que puede infrarepresentar a determinados subgrupos de parejas LGBTQ+, como parejas más jóvenes, con mayor diversidad étnica, menor nivel socioeconómico o mayor nivel de conflicto. Además, el estudio se realizó en dos ciudades de Estados Unidos, por lo que la generalización a otros contextos culturales y sanitarios debe hacerse con cautela. Existen ensayos similares consistentes, especialmente el programa paralelo para parejas masculinas (SSSR-M), que mostró mejoras comparables en comunicación y satisfacción, lo que apoya la coherencia externa de los hallazgos dentro de intervenciones afirmativas dirigidas a población LGBTQ+.

El estudio considera de forma explícita todos los desenlaces definidos y discute su importancia clínica, distinguiendo entre resultados proximales y distales. Se informan claramente las utilidades: mejoras significativas y clínicamente relevantes en comunicación, satisfacción y confianza a corto plazo, con mantenimiento de los efectos en

**10. ¿Se han tenido en cuenta todos los resultados y su importancia clínica?**  
 Sí  
 Utilidades y disutilidades de cada desenlace.

Balance de efectos positivos/ negativos. Preferencias del paciente, costes, etc.

**Importancia clínica y mantenimiento:**

“Analyses of follow-up data suggest maintenance of effects on the proximal but not the distal outcomes.”

**Aceptabilidad y preferencias del paciente:**

“Regarding program acceptability, participants’ satisfaction ratings of the overall program ( $M = 7.11$ ,  $SD = 1.00$ ) and confidence in recommending the program to a friend ( $M = 7.23$ ,  $SD = 0.91$ ) were very high.”

**Limitaciones y disutilidades:**

“Relationship satisfaction and confidence decreased significantly, and relationship instability increased marginally, from immediately post-program to the 3-month follow-up.”

**Costes:** no consta.

comunicación a 3 meses. También se describen las disutilidades o limitaciones: ausencia de beneficio claro en estrés, apoyo social e inestabilidad, y pérdida de los efectos en satisfacción y confianza en el seguimiento, con retorno a niveles basales. El balance de efectos positivos y negativos se analiza de forma detallada en la discusión. Además, se tienen en cuenta las preferencias del paciente mediante altas puntuaciones de satisfacción y recomendación del programa, y se evalúa la aceptabilidad y factibilidad de la intervención. No se analizan costes económicos de forma directa, pero sí se valora la carga, duración y viabilidad clínica, por lo que el estudio tiene en cuenta la importancia clínica global de los resultados.

A partir de este único ensayo piloto no puede establecerse con certeza definitiva el balance beneficio-riesgo-coste, pero los datos disponibles sugieren un perfil beneficio-riesgo favorable. Los beneficios esperados incluyen mejoras significativas en comunicación, satisfacción y confianza a corto plazo, con alta aceptabilidad y satisfacción de las participantes. Los riesgos son mínimos, al tratarse de una intervención psicoeducativa grupal sin efectos adversos reportados y con alta fidelidad de implementación. Los costes económicos no se analizan, pero la intervención es breve (10 horas), grupal y factible, lo que sugiere una carga moderada.

**11. ¿Los beneficios que se espera obtener justifican los riesgos y los costes?**

Sí

Es improbable que pueda deducirse solo de un ensayo, pero ¿qué piensas tú al respecto?

**Beneficios y eficacia:**

“Overall, findings from the small randomized waitlist controlled trial were supportive of the program’s feasibility, acceptability, and efficacy. Results provided initial evidence of immediate and sustained program effects on proximal couple outcomes targeted by the program, particularly communication.”

**Limitaciones de los beneficios:**

“Immediate effects on distal couple outcomes (global relationship satisfaction and confidence) were also observed, but not maintained over the follow-up period.”

**Aceptabilidad y bajo riesgo:**

“Regarding program acceptability, participants’ satisfaction ratings of the overall program ( $M = 7.11$ ,  $SD = 1.00$ ) and confidence in recommending the

program to a friend ( $M = 7.23$ ,  $SD = 0.91$ ) were very high.”

**Bajo poder y ausencia de daño:**

“The small sample size of this study makes it difficult to draw definitive conclusions regarding program effects on these variables.”

“The small sample size limited power to detect program effects, particularly on stress, social support, and relationship instability.”

Aunque los efectos distales no se mantienen a 3 meses, el balance global indica que los beneficios potenciales probablemente justifican los riesgos y los costes, especialmente como intervención preventiva y de bajo riesgo en población LGBTQ+, reconociendo que se requieren ensayos mayores para confirmarlo.

---