

Evaluación del programa 'En familia. Educar Para la Vida (EFEPV)'. Fase 3: Mejora del Análisis de eficacia y evaluación de impacto.

ANEXO 1. INFORME FINAL.

Programa En Familia: Educar para la Vida (EFEPV). Ampliación del análisis de eficacia. Evaluación de eficacia periodo 2025-2026.

Mayo 2026

fad
Juventud

Financiado por





ANEXO 1. INFORME FINAL.
Programa En Familia: Educar para la Vida (EFEPV).
Ampliación del análisis de eficacia.
Evaluación de eficacia periodo 2025-2026.

RESUMEN EJECUTIVO

Se presentan los resultados del análisis de eficacia del programa En Familia: Educar para la Vida (EFEPV), desarrollado por Fad Juventud como intervención universal de prevención familiar orientada al fortalecimiento de competencias parentales y factores de protección familiar en el periodo de evaluación 2025-2026.

Objetivo:

- Analizar la eficacia del programa En Familia: Educar para la Vida (EFEPV).
- Determinar el cambio producido en los/las familiares tras la intervención en las dimensiones de funcionalidad familiar
- Describir posibles diferencias en función de características de los/las familiares que reciben el programa (sexo del participante y comunidad autónoma de residencia)
- Explorar el mantenimiento de los posibles cambios observados tras la intervención mediante una medida de seguimiento a tres meses en una submuestra del grupo experimental.

Metodología:

- Diseño cuasi-experimental pretest–postest con grupo control no equivalente.
- Durante 2025 se realizaron únicamente medidas pretest y postest, mientras que en 2026 se incorporó adicionalmente una medida de seguimiento a tres meses en una submuestra longitudinal.
- El análisis de eficacia se realizó mediante ANCOVA, ajustando estadísticamente las puntuaciones postest por el nivel inicial pretest. Complementariamente, análisis de cambio fiable individual (Reliable Change Index) para estimar el porcentaje de participantes que experimentan mejoras reales más allá del error de medida de la escala.
- El análisis de seguimiento a tres meses incluye índice de retención del efecto $((MSEG - MPRE) / (MPOST - MPRE)) \times 100$, cálculo de trayectorias individuales (Reliable Change Index, RCI) y contraste entre grupos mediante ANCOVA.
- La disponibilidad de datos longitudinales completos fue limitada en el grupo experimental, debido tanto a programas aún en curso como a incidencias administrativas y dificultades de emparejamiento longitudinal.

Participantes:

Eficacia pretest-postest

Muestra pareada pretest-postest formada por 415 participantes con datos completos en ambas medidas

- Grupo experimental: 129 familiares asistentes al programa EFEPV. 82.9% mujeres. Edad media: 49 años. Castilla-La Mancha (64.3%) y Comunidad de Madrid (35.7%).
- Grupo control: 286 padres y madres sin intervención. 78% mujeres. Edad media: 47 años. Castilla-La Mancha (50.3%) y Comunidad de Madrid (49.7%)

Para seguimiento a tres meses:

- Grupo experimental: 16 familiares asistentes al programa EFEPV
- Grupo control: 95 padres y madres sin intervención

Resultados:

- Una vez ajustadas las diferencias iniciales mediante ANCOVA, el grupo experimental obtiene puntuaciones significativamente superiores al grupo control en cinco de las seis dimensiones evaluadas de funcionamiento familiar. Las mayores diferencias entre grupos se observan en: Gestión de conflictos ($d = .69$; $p < .001$), Funcionamiento global ($d = .58$; $p < .001$) y Establecimiento de Normas familiares ($d = .40$; $p < .001$).
- También se observan diferencias significativas, aunque de menor magnitud, en: Consenso/Corresponsabilidad Familiar ($d = .31$; $p < .01$) y Clima familiar ($d = .29$; $p < .01$).
- La dimensión Estrés Familiar no mostró diferencias significativas entre grupos.
- El grupo experimental presenta porcentajes claramente superiores de mejora fiable respecto al grupo control, especialmente en: Gestión de conflictos (28.7% vs. 3.8%), Establecimiento de Normas Familiares (27.9% vs. 6.6%) y Funcionamiento global (24.8% vs. 7.0%).
- Los análisis moderadores no mostraron diferencias significativas en función del sexo de la persona participante ni de la comunidad autónoma de implementación, sugiriendo un patrón relativamente estable.
- En el seguimiento a tres meses, el grupo control mostró una elevada estabilidad temporal, mientras que el grupo experimental mantuvo, en la mayoría de dimensiones estabilidad o mantenimiento de las mejoras del postest.
- El patrón de trayectorias individuales resulta compatible con un mantenimiento general del efecto del programa a corto-medio plazo.
- Sin embargo, en la medida de seguimiento, la comparación entre grupo experimental y control no es significativa debido a la limitada muestra del primer grupo.

Conclusiones:

- Los resultados apoyan la eficacia del programa EFEPV en la mejora de distintas dimensiones del funcionamiento familiar, especialmente en gestión de conflictos, establecimiento de normas y funcionamiento familiar global.
- Los tamaños de efectos observados presentan magnitudes pequeñas y moderadas, coherentes con la literatura sobre programas universales de prevención familiar, y resultan relevantes considerando la brevedad de la intervención (4–6 semanas).
- El análisis individual mediante RCI muestra además que una proporción relevante de familias experimenta mejoras detectables tras la participación en el programa, reforzando la relevancia práctica de los resultados observados.
- La evidencia obtenida sugiere la utilidad del programa EFEPV como herramienta de prevención familiar y fortalecimiento de competencias parentales en contextos comunitarios.

1 .EL PROGRAMA EN FAMILIA: EDUCAR PARA LA VIDA

El programa *En Familia: Educar para la Vida* (EFEPV), diseñado y desarrollado por Fad Juventud es una intervención de prevención familiar de carácter universal orientada a fortalecer las competencias parentales y la capacidad preventiva de las familias frente a distintas conductas de riesgo en la infancia y la adolescencia. El programa, diseñado originalmente en 2015 y actualizado con periodicidad, se implementa en diferentes comunidades autónomas españolas en colaboración con administraciones públicas y entidades comunitarias, dentro del marco de actuaciones financiadas, entre otros organismos, por la Delegación del Gobierno para el Plan Nacional sobre Drogas.

EFEPV parte de la consideración de la familia como principal agente de socialización y como un contexto clave para la promoción de factores de protección relacionados con la comunicación familiar, la supervisión parental, el establecimiento de normas y límites, la cohesión afectiva y la resolución de conflictos. Así, su principal finalidad es fortalecer el papel de la familia como agente preventivo, desde el convencimiento de que prevenir en la familia es algo natural, siendo otra forma más de cuidado. El programa ofrece a padres y madres herramientas prácticas y sencillas para acompañar a sus hijos e hijas en su crecimiento, fomentando hábitos saludables y evitando conductas de riesgo. Su población diana está constituida por padres y madres, preferentemente con hijos/as en edad escolar. El marco conceptual del programa integra distintos modelos teóricos ampliamente utilizados en prevención, entre ellos el modelo ecológico, el modelo biopsicosocial, el modelo de competencia social y la teoría del aprendizaje social. Desde esta perspectiva, las conductas de riesgo son entendidas como fenómenos multicausales en los que interactúan variables individuales, familiares, sociales y contextuales. De esta forma, los objetivos del programa son los siguientes:

Objetivos generales

- Empoderar a las familias en materia de prevención de conductas de riesgo y, entre ellas, la del consumo de drogas.
- Facilitar a la familia estrategias y recursos que permitan su actuación de un modo adecuado en lo que respecta a la educación y a la prevención de conductas de riesgo en las hijas y en los hijos.

Objetivos Específicos

- Adquirir habilidades para mejorar las relaciones en la familia.
- Informar sobre distintos tipos de sustancias consideradas como drogas, así como sobre los riesgos asociados a su consumo.
- Facilitar la adopción de actitudes y la adquisición de habilidades relacionadas con un adecuado clima afectivo y de comunicación en la familia.
- Favorecer el desarrollo de capacidades para mejorar la convivencia familiar a través de la negociación de normas y límites.
- Dar a conocer pautas y orientaciones básicas acerca de cómo abordar dificultades y situaciones de riesgo en la familia.
- Reflexionar y adquirir conocimientos sobre cómo trabajar los valores en la familia.
- Fomentar un uso responsable y equilibrado de las TIC en el entorno familiar, identificando tanto los beneficios como los riesgos asociados.
- Desarrollar habilidades y estrategias efectivas de regulación parental para la supervisión del uso de TIC en el hogar.

La intervención evaluada se compone de cuatro módulos centrales cuyos componentes y sesiones centrales son: a) comunicación, b) clima y afectividad familiar, c) organización familiar (normas y límites) y d) gestión de tensiones y conflictos. Las sesiones, habitualmente de carácter grupal y participativo, son dinamizadas por profesionales con formación en psicología, educación, intervención social o prevención de adicciones. El programa combina exposiciones breves, dinámicas grupales, análisis de casos, tareas para casa y actividades experienciales orientadas a facilitar la reflexión y el aprendizaje práctico de madres y padres. Los materiales didácticos del programa incluyen guías específicas tanto para familias como para profesionales dinamizadores. Estas guías desarrollan contenidos relacionados con factores de riesgo y protección, parentalidad positiva, comunicación, regulación emocional, convivencia familiar y prevención de conductas problemáticas, incorporando actividades prácticas y recursos aplicados para el trabajo con familias.

1.1. Evaluaciones previas del programa

Las evaluaciones previas del programa *En Familia: Educar para la Vida* (EFEPV) se han desarrollado de forma progresiva, pasando de una evaluación centrada en el proceso y satisfacción de las familias participantes hacia análisis específicos de eficacia sobre dimensiones del funcionamiento familiar. En una primera fase (2023-2024), se realizó una evaluación de proceso sobre 1.466 usuarios/as de distintas comunidades autónomas, especialmente Comunidad de Madrid y Castilla-La Mancha. Los resultados mostraron niveles muy elevados de satisfacción general con el programa (i.e. puntuaciones medias superiores a 4.5 sobre 5) y valoraciones especialmente positivas de las personas formadoras, la utilidad práctica de los contenidos y los módulos relacionados con comunicación familiar, gestión de conflictos y establecimiento de normas. Además, las familias informaron de mejoras percibidas en la convivencia, la comunicación con sus hijos/as y el afrontamiento de situaciones educativas complejas, lo que apoyó la continuidad y consolidación del programa en distintos contextos comunitarios.

Posteriormente, durante 2025, se llevó a cabo una primera evaluación de eficacia mediante un diseño cuasi-experimental PRE-POST con grupo control equiparado. Los resultados mostraron mejoras estadísticamente significativas en establecimiento de normas y gestión de conflictos familiares, tanto a nivel grupal como intraindividual, con tamaños de efecto pequeños a moderados. La dimensión de gestión de conflictos fue el componente con una evolución más consistente tras la intervención, seguida del sistema de normas y límites familiares. En contraste, el grupo control no presentó cambios relevantes entre medidas. Asimismo, los análisis sugirieron que el programa podría mostrar un mayor impacto en familias con hijos/as menores de 12 años y especialmente en madres, aunque las submuestras reducidas limitaban la interpretación de estos análisis estratificados. La evaluación permitió además desarrollar y analizar preliminarmente la escala EF-FAD como instrumento específico para la evaluación del funcionamiento familiar en el contexto del programa, observándose adecuados niveles de consistencia y estabilidad.

Estas evaluaciones previas proporcionaron evidencia favorable sobre la aceptación, viabilidad y potencial eficacia del programa EFEPV, así como el entrenamiento en una metodología de evaluación en las personas implementadoras, base indispensable para el desarrollo de análisis posteriores y estudios orientados a consolidar la evidencia empírica del programa.

2.OBJETIVOS DE LA EVALUACIÓN

El objetivo general del proyecto es analizar la eficacia del programa En Familia: Educar para la Vida (EFEPV). Se desglosa en los siguientes objetivos específicos:

- Determinar el cambio producido en los/las familiares tras la intervención en las dimensiones evaluadas: (nombre y definición de los factores EF-FAD)
 - Factor 01: Clima Familiar
 - Factor 02: Estrés Familiar (invertido)
 - Factor 03: Establecimiento de Normas
 - Factor 04: Consenso/Corresponsabilidad
 - Factor 05: Gestión de Conflictos
- Describir posibles diferencias en función de características de los/las familiares que reciben el programa.
- Explorar el mantenimiento de los posibles cambios observados tras la intervención mediante una medida de seguimiento a tres meses en una submuestra del grupo experimental.

Las hipótesis planteadas en este análisis de eficacia son las siguientes:

- Se observarán cambios significativos entre las medidas pretest y postest del grupo experimental. Se espera que este grupo aumentará significativamente sus puntuaciones en las dimensiones de funcionalidad familiar en la medida postest respecto a la medida pretest (Hipótesis 1)
- No se observarán cambios significativos entre las medidas pretest y postest del grupo control (Hipótesis 2).
- Así, se espera, que en la medida postest, las puntuaciones entre los grupos experimental y control sean diferentes (Hipótesis 3).
- Se espera que las mejoras observadas en el grupo experimental se mantengan parcialmente en la medida de seguimiento a tres meses (Hipótesis 4).
- Se espera que las diferencias entre el grupo experimental y el grupo control observadas en post-test se mantengan, al menos parcialmente, en la medida de seguimiento a tres meses (Hipótesis 5).

3.METODOLOGÍA

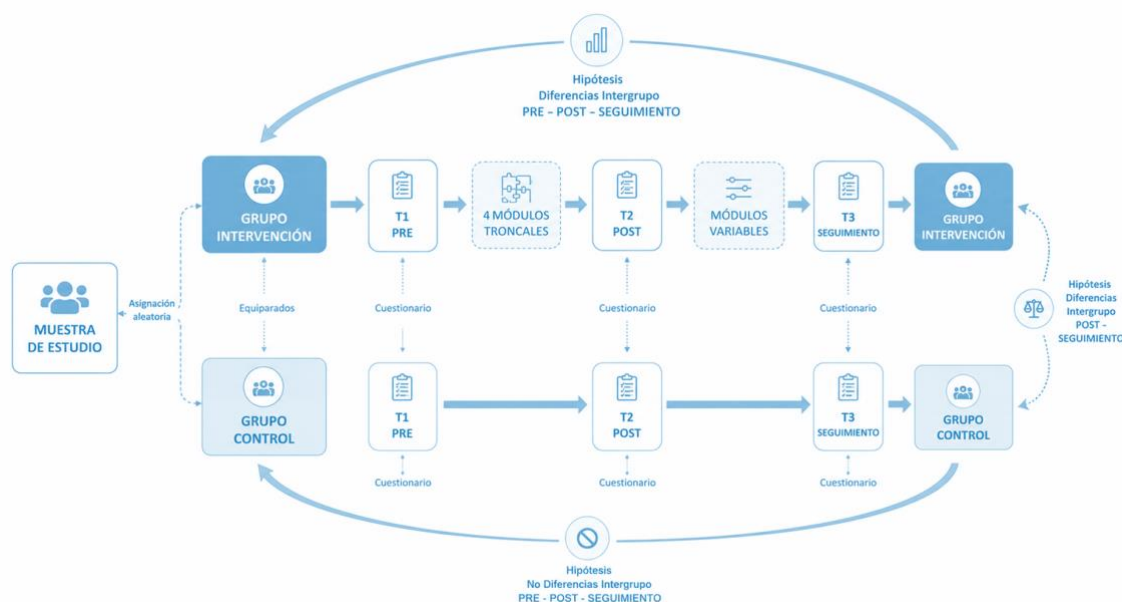
3.1.Diseño

La evaluación de eficacia del programa EFEPV utilizó un diseño cuasi-experimental pretest-postest, con grupo control no equiparado. En la evaluación correspondiente al periodo 2025-2026 se realizaron dos medidas principales: antes (pretest) y después (postest) de la intervención. En ellas se evaluó repetidamente a dos grupos: a) familiares participantes en el programa EFEPV (grupo experimental) y b) un conjunto de familias que no recibieron intervención alguna (grupo control), lo que permite comparar el efecto del programa, diferenciándolo de la no intervención. La elección del diseño responde a las condiciones reales de implementación del programa, en las que la asignación aleatoria no era viable.

Adicionalmente, en el periodo 2026, se incorporó una medida de seguimiento a tres meses (SEG) con finalidad descriptiva y exploratoria.

El grupo experimental recibió el programa EFEPV, en su formato de cuatro módulos o sesiones troncales, con una periodicidad semanal/quincenal, por lo que la duración total de intervención es aproximadamente de 4 a 6 semanas. La intervención se complementa tras la evaluación con módulos variables que pueden oscilar entre 1 y 10 sesiones adicionales. El tipo de muestreo fue intencional. Se intentó obtener tamaños muestrales similares en cada grupo.

Figura 1. Diseño experimental



3.2.Participantes

Para los análisis de eficacia con muestras relacionadas (análisis intraindividual), una vez descartados aquellos registros que no tenían su correspondiente cuestionario inicial, intermedio y/o final, no pudieron ser emparejados por clave alfanúmerica, o fueron excluidos por patrones azarosos o discapacidad social, la muestra pareada pretest-postest de estudio para los análisis de eficacia quedó conformada por 415 participantes, con datos completos en el pretest y postest.

Grupo experimental. Los participantes fueron 129 familiares asistentes al programa EFEPV durante el periodo evaluado, con hijos/as entre 6 y 17 años y edad media 49 años (DT = 6.9). De ellos, el 82.9% eran mujeres (n = 107) y el 17.1% hombres (n = 22). Las personas participantes residían en las comunidades autónomas de Madrid (35.7%) y Castilla La Mancha (64.3%). Un 92.2% compartía crianza y tareas educativas, frente a un 7.8% que lo hacía en solitario. Otras características familiares pueden ser observadas en la tabla 1.

Grupo control. Respecto al grupo control, las personas participantes fueron 286 padres/madres, con hijos/as entre 6 y 17 años. De ellos, el 78% eran mujeres (n = 223) y el 22% hombres (n = 63). Las personas participantes residían en las comunidades autónomas de Madrid (49.7%) y Castilla La Mancha (50.3%). Un 86.7% compartía crianza y tareas educativas, frente a un 13.3% que lo hacía en solitario. Otras características familiares pueden ser observadas en la tabla 1.

Tabla 1. Características de los grupos. Linea base variables sociodemográficas	GRUPO EXPERIMENTAL (n=129)	GRUPO CONTROL (n=286)	DIFERENCIA (chi-cuadrado)
Comunidad autónoma Castilla La mancha	64.3%	50.3%	**
Comunidad autónoma Madrid	35.7%	49.7%	**
Edad media	49	47	*
Hombres	17.1%	22%	*
Mujeres	82.9%	78%	*
Comparte tareas educativas con pareja con la que convive	92.2%	86.7%	*

*p<.05; ** p<.01; ***: p<.001; (n.s.): no significativo

3.2.1. Atrición (pérdida muestral)

Como se observa en la tabla 2, la disponibilidad de datos longitudinales completos fue limitada al cierre de la recogida de datos en el grupo experimental.

La disponibilidad de datos longitudinales completos respecto a la muestra inicial (n = 805) fue del 16.5% en el grupo experimental y del 75.6% en el grupo control (n = 403).

	GRUPO EXPERIMENTAL			GRUPO CONTROL		
	Castilla La Mancha	Comunidad de Madrid	Total	Castilla La Mancha	Comunidad de Madrid	Total
Solo PRE	405 (68.6%)	130 (60.2%)	535 (66.1%)	33 (16.8%)	52 (25.3%)	85 (21.1%)
Solo POST	53 (9.0%)	25 (11.6%)	78 (9.6%)	-	-	-
Solo SEGUIMIENTO	24 (4.1%)	6 (2.8%)	30 (3.7%)	-	-	-
PRE- SEGUIMIENTO	19 (3.2%)	7 (3.2%)	26 (3.2%)	5 (2.5%)	9 (4.4%)	14 (3.5%)
PRE-POST	82 (13.9%)	35 (16.2%)	117 (14.4%)	104 (52.8%)	105 (51.0%)	209 (51.9%)
PRE-POST- SEGUIMIENTO	6 (1.0%)	11 (5.1%)	17 (2.1%)	55 (27.9%)	40 (19.4%)	95 (23.6%)
Total	589 (100%)	216 (100%)	805 (100%)	197 (100%)	206 (100%)	403 (100%)

Nota: En sombreado, la muestra válida para ser pareada y realizar análisis sobre ella

Respecto al grupo experimental, esta cifra requiere una aclaración metodológica. Se refiere a la pérdida de participantes durante el estudio, no necesariamente al abandono real del programa, ya que la estructura de la base de datos sugiere la coexistencia de distintos factores metodológicos y contextuales.

En este sentido, la explicación a esta pérdida muestral es multifactorial:

- a) En primer lugar, una proporción significativa de los casos clasificados inicialmente como “solo pre” corresponde a programas todavía en curso de implementación en el momento de extracción de la base de datos. Por tanto, parte de la ausencia de medidas postest o seguimiento no reflejaría pérdida efectiva de participantes, sino falta de disponibilidad temporal de dichas evaluaciones (porcentaje desconocido)
- b) En segundo lugar, otro porcentaje de pérdida emparejada puede deberse a abandono del programa, lo que imposibilita el cumplimentado de medidas postest y seguimiento (porcentaje desconocido)
- c) En tercer lugar, los patrones de casos sugieren también dificultades o incidencias logístico-administrativas en la recogida de medidas pretest, postest o seguimiento. Es decir, casos que no cumplimentaron la primera, segunda o tercera medida. Esta interpretación se ve reforzada por la presencia de participantes con medidas pretest y seguimiento disponibles sin correspondencia emparejada en el postest (porcentaje aportado por medidas postest 16.5%, porcentaje aportado por medidas pretest desconocido). Comparando los datos del grupo experimental y control, mientras que el grupo control presentó una estructura longitudinal mucho más estable y cerrada administrativamente, los patrones apuntan a que el grupo experimental estuvo más expuesto a calendarios heterogéneos, incorporación progresiva de grupos, distintos equipos de personas facilitadoras en los programas de intervención y variabilidad logística en la administración de los cuestionarios.
- d) Por último, en un porcentaje menor, errores de consigna de clave alfanumérica de emparejamiento de casos no recuperable o asignable (2%)

Respecto a las explicaciones b) y c), al no existir un registro sistemático de asistencia a las sesiones, no puede determinarse cuántas familias dejaron realmente de acudir y cuántas simplemente no respondieron al cuestionario de seguimiento.

En esta línea, para descartar autoselección por nivel basal, se realizó análisis de diferencias entre aquellos participantes que solo completaron los cuestionarios pretest de aquellos que completaron cuestionarios pretest y postest. Este análisis permitió descartar que existiera una autoselección asociada al nivel inicial de funcionamiento familiar, ya que quienes completaron el postest no se diferenciaron de quienes solo participaron en el pretest en ninguna de las dimensiones de la EF-FAD (todos los valores $p \geq .169$; $|d| \leq 0.14$), ni por sexo ($\chi^2 = 0.08$; $p = .784$). Sí se observó una mayor participación en el post-test en Castilla-La Mancha ($\chi^2 = 6.41$; $p = .011$).

3.2.2. Equivalencia inicial

Aunque ambos grupos comparten variables sociodemográficas similares (Ej. comunidad autónoma de residencia, sexo, edad media, número de hijos, etc.), existen diferencias en la equivalencia inicial que conviene subrayar. Así, en las variables evaluadas, y aunque con perfiles similares, el grupo experimental parte con niveles de funcionamiento familiar que permiten mayor margen de mejora que el grupo control. Todas las diferencias son estadísticamente significativas ($p \leq .033$ en todas las dimensiones), con magnitudes pequeñas pero sistemáticas ($|d|$ entre .13 y .24). Esta asimetría podría explicarse por el hecho de que las familias que se acercan al programa EFEPV no constituyen una muestra aleatoria de la población, sino que representarían, en la medida inicial, a familias que ya presentan mayor dificultad o problemáticas en las dimensiones que el programa pretende trabajar.

Así, se sugeriría la explicación de la autoselección muestral en el grupo experimental por motivación o implicación en la crianza.

En otras palabras, las familias que acuden voluntariamente a recursos como los ofrecidos por Fad Juventud podrían estar empezando a mostrar signos de dificultad, a diferencia de la población general, que no presenta necesariamente estas problemáticas. Por otro lado, al tratarse de medidas auto-informadas, es posible que las personas que se inscriben voluntariamente en el Programa EFEPV lo hagan porque ya tienen una mayor sensibilidad hacia la crianza y mayor capacidad de autoobservación sobre su dinámica familiar. En ese caso, las diferencias observadas no reflejarían necesariamente una mayor problemática, sino una percepción más consciente al responder la autoevaluación. Esta observación de asimetría exige el uso de análisis ANCOVA para que las comparaciones posttest resulten interpretables.

3.3.Procedimiento

De forma previa a la implementación del estudio, se obtuvieron las autorizaciones correspondientes por parte del equipo de Fundación FAD Juventud y la aprobación por parte de la Oficina de Investigación Responsable de la Universidad Miguel Hernández de Elche con código AUT.DPS.VMCP.241211.

El trabajo de campo se prolongó de enero 2025 a mayo 2026. El proceso de recogida de cuestionarios, en el marco de las distintas ediciones del programa EFEPV desarrolladas en ese periodo en las comunidades autónomas de Castilla La Mancha y Comunidad Autónoma de Madrid.

- Primera tanda de evaluación: el trabajo de campo se prolongó durante los meses de enero a abril de 2025, en dos medidas (pretest-postest)
- Segunda tanda de evaluación: Los datos se recogieron progresivamente entre enero y mayo de 2026, en tres medidas (pretest-postest-seguimiento)

Cada participante utilizó una clave de identificación encriptada compuesta por (Inicial del primer nombre + tres últimos dígitos del DNI + código de dos dígitos del curso proporcionado por la persona dinamizadora del curso), garantizando la confidencialidad del proceso. Finalizada la intervención, se volvían a cumplimentar los cuestionarios posttest empleando la misma metodología previamente descrita, o bien mediante contacto a través de correo electrónico con la persona asistente.

Los participantes cumplimentaban los cuestionarios de forma individual, autoinformada y voluntaria. En el grupo experimental, estos fueron cumplimentados bajo supervisión del equipo técnico del programa. Dicho personal proporcionó previamente las instrucciones necesarias, explicó los objetivos del estudio y resolvió cualquier duda que pudiera surgir.

La recogida de datos fue realizada por personal responsable de la implementación del programa EFEPV, siguiendo el procedimiento habitual de aplicación del programa en los distintos contextos comunitarios y educativos participantes. El equipo de investigación de la Universidad Miguel Hernández no participó en la implementación de la intervención ni en la recogida de datos, limitándose sus funciones al asesoramiento metodológico, análisis estadístico de los datos y elaboración de informes técnicos y científicos.

Posteriormente, durante el mes de mayo de 2026, se procedió al análisis de eficacia.

3.4. Instrumentos

Como medida principal del funcionamiento familiar se utilizó la Escala Familiar de Funcionamiento, Ajuste y Dinámica (EF-FAD), desarrollada en el marco del propio programa EFEPV y validada psicométricamente en un estudio paralelo (Cabrera-Perona et al., en revisión, ver anexo 1).

La escala consta de 24 ítems con respuesta en formato Likert de 1 a 7 (1 = Nunca, 7 = Siempre), que se estructura en cinco dimensiones: Clima familiar, Estrés familiar, Establecimiento de Normas y Límites, Consenso familiar/corresponsabilidad y Gestión de conflictos familiares. Adicionalmente, se calcula un índice global de funcionamiento familiar como promedio de las cuatro dimensiones positivas, siguiendo la estructura bifactor identificada en el estudio de validación. Los ítems de Estrés familiar son invertidos durante la corrección, de modo que puntuaciones más altas indican menor estrés percibido. Esta inversión es coherente con la lógica del índice global: en él, una puntuación más alta refleja, en todas las dimensiones, mejor funcionamiento familiar.

Así, las dimensiones evaluadas por la escala EF-FAD son:

Clima familiar	Dimensión relacional entre los/las familiares, incluye la percepción del a) grado de convivencia y realización de actividades conjuntas entre los/las familiares, la b) percepción del familiar de que la comunicación en la unidad familiar es satisfactoria y adecuada y c) el grado de expresión afectiva entre los/las familiares. Mayor puntuación indica mejor clima familiar.
Estrés familiar (puntuación inversa)	Evalúa el grado de afectación emocional percibida por el familiar a causa de la gestión educativa o las problemáticas referidas a los menores. Mayor puntuación indica un menor estrés familiar.
Establecimiento de Normas y límites	Evalúa el nivel de establecimiento de normas de convivencia claras, supervisión de las actividades y adopción de sanciones o consecuencias en caso de incumplimiento. Mayor puntuación indica un mayor establecimiento de normas y límites familiares.
Consenso familiar/corresponsabilidad	Evalúa el grado de acuerdo de tutores/as o familiares que comparten la responsabilidad educativa sobre la importancia de los problemas o riesgos sobre los menores, y la forma de afrontarlos. Mayor puntuación indica un mayor grado de consenso.
Gestión de conflictos familiares	Evalúa la percepción acerca de la gestión y adecuada resolución de conflictos familiares. Mayor puntuación indica una mejor gestión y resolución de conflictos.

3.5. Análisis de datos

La estrategia analítica se organiza en distintos bloques.

3.5.1. Patrón de atrición y equivalencia inicial de los grupos.

Las comparaciones se realizaron con la prueba t de Welch cuando las distribuciones son compatibles con la normalidad, y con la prueba U de Mann–Whitney en caso contrario, acompañadas del correspondiente tamaño del efecto (d de Cohen o equivalente rango-biserial transformado a la métrica de d). La atrición se analizó contrastando a quienes completan el postest con quienes solo aportan pretest, tanto en las puntuaciones EF-FAD como en variables sociodemográficas básicas.

3.5.2. Efecto del programa

El efecto del programa se estima mediante análisis de covarianza (ANCOVA), con la puntuación postest como variable dependiente, el grupo (experimental o control) como factor entre sujetos y la puntuación pretest como covariable ($POST \sim PRE + Grupo$). Esta especificación, preferible a la basada en puntuaciones de cambio en este contexto, ofrece dos ventajas relevantes para un diseño cuasi-experimental: corrige las diferencias iniciales sistemáticas entre grupos y estima el efecto del programa ajustado al punto de partida individual (Maxwell et al., 2017; Van Breukelen, 2006). Del modelo se obtienen las medias marginales estimadas (EMM) con su intervalo de confianza del 95%, la diferencia ajustada entre grupos, el estadístico η^2p y el tamaño del efecto d entre grupos (utilizando como estandarizador la desviación típica residual del ANCOVA). Como triangulación frente a las cautelas asociadas al uso del ANCOVA en diseños no aleatorizados (Lord, 1967), se calcula además la media estandarizada de respuesta (SRM) intragrupo, que expresa el cambio pre-post como múltiplo de la desviación típica del propio cambio.

Para conocer el porcentaje de grupo experimental que mostraba una mejora fiable, junto con el análisis de las medias grupales (efectos medios estimados), resultaba importante examinar también el patrón de respuesta a nivel individual. Este análisis se realizó mediante análisis Reliable Change Index (RCI; Jacobson y Truax, 1991), que considera mejora fiable el cambio positivo entre el pretest y el postest cuando éste supera el margen esperado de error estadístico de la escala ($RCI > 1.96$), y empeoramiento fiable un cambio negativo de igual magnitud ($RCI < -1.96$).

3.5.3. Efecto en función de variables moderadoras

El tercer bloque atiende a tres preguntas complementarias. La primera examina si el efecto del programa varía en función del sexo o de la comunidad autónoma, introduciendo el término de interacción Grupo \times Moderador en el modelo ANCOVA. La segunda estima, con propósito exploratorio y descriptivo, el tamaño del efecto del programa en cada comunidad autónoma por separado, como verificación de la estabilidad territorial del patrón.

3.5.4. Análisis de seguimiento a tres meses

Para el análisis de seguimiento a tres meses se trabajó con dos submuestras con datos en los tres momentos de evaluación: el grupo experimental aportó $n = 16$ ($n = 12$ en Consenso/Corresponsabilidad). El grupo control aportó $n = 95$ ($n = 76$ en Consenso/Corresponsabilidad). El reducido tamaño del experimental sitúa el análisis en un plano principalmente descriptivo, si bien disponer del control permite añadir un contraste entre grupos en el momento de seguimiento.

Para cada dimensión se calcularon medias y desviaciones típicas en los tres momentos. Sobre la trayectoria del experimental se estimó un índice de retención del efecto $((MSEG - MPRE) / (MPOST - MPRE)) \times 100$: valores próximos al 100% indican mantenimiento del efecto inmediato; superiores, mejora adicional a tres meses; e inferiores, pérdida parcial. El índice se interpretó con cautela dada la *n*. Complementariamente, se aplicó el Reliable Change Index (RCI; Jacobson y Truax, 1991) a la transición posttest–seguimiento en ambos grupos, clasificando a los participantes en mejora adicional, estabilidad o recaída fiable; el control proporciona aquí una tasa basal de cambio fiable atribuible a fluctuación natural en ausencia de intervención.

El contraste entre grupos en seguimiento se estimó mediante ANCOVA (SEG ~ POST + Grupo), tomando la puntuación post-test como covariable; de ese modelo se obtuvieron la diferencia ajustada entre grupos, su intervalo de confianza al 95% y el tamaño del efecto *d* (estandarizado con la DT residual). Como prueba de robustez se reportó adicionalmente el contraste *t* de Welch sobre las medias de seguimiento sin covariable. Finalmente, se representaron las trayectorias individuales pre-posttest-seguimiento por dimensión y grupo.

Todos los análisis se realizaron en R 4.5.3 (R Core Team, 2025) con los paquetes *psych*, *car*, *emmeans* y *effectsize*. El nivel de significación se fijó en $\alpha = .05$. Para la interpretación de los tamaños del efecto se toman como referencia los umbrales clásicos de Cohen (1988: $d \approx .20$ pequeño; $\approx .50$ moderado; $\approx .80$ grande), con la advertencia explícita de que, en el campo de los programas universales de prevención familiar, efectos de magnitud pequeña-media sobre el funcionamiento familiar son habituales y sustantivamente relevantes (Sandler et al., 2011; Van Ryzin et al., 2016).

4.RESULTADOS.

El análisis de eficacia pretende a comprobar la evolución del grupo experimental tras la intervención, así como su comparación con el grupo control. Las hipótesis planteadas son: que el grupo experimental mostrará mejoras significativas en las dimensiones de funcionalidad familiar tras la intervención (hipótesis 1), mientras que no se esperan cambios relevantes en el grupo control (hipótesis 2). Asimismo, se esperan diferencias significativas entre ambos grupos (experimental y control) en la medida tras la intervención, lo que permitirá valorar el efecto del programa, así como diferenciarlo de la *no* intervención (Hipótesis 3). Se espera que las mejoras observadas en el grupo experimental se mantengan parcialmente en la medida de seguimiento a tres meses (Hipótesis 4). Por último, se espera que las diferencias entre el grupo experimental y el grupo control observadas en post-test se mantengan, al menos parcialmente, en la medida de seguimiento a tres meses (Hipótesis 5).

4.1. Análisis descriptivos por grupo y dimensión.

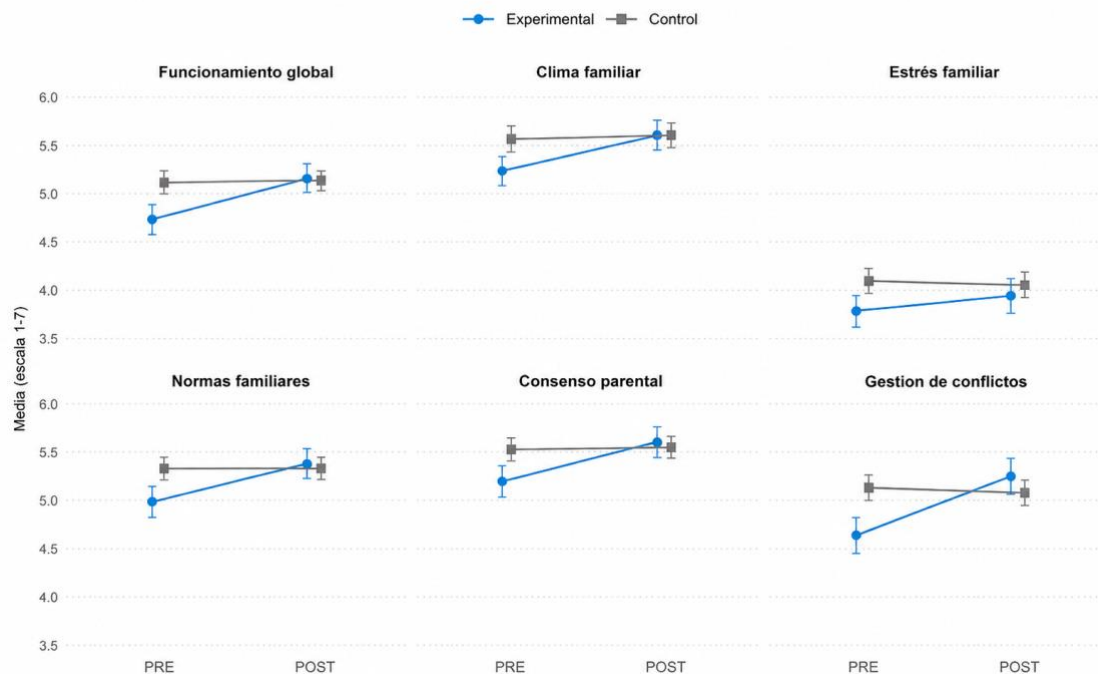
La Tabla 3 recoge las medias y desviaciones típicas pretest y posttest de las variables evaluadas desglosadas por grupo.

Tras la intervención, al analizar la evolución del grupo experimental, los resultados señalan un aumento en todas las dimensiones analizadas, a excepción de la dimensión Estrés Familiar ($M_{pre} = 3.83 - M_{post} = 3.95$). Las dimensiones con cambio más visible son las de Gestión de Conflictos ($M_{pre} = 4.67 - M_{post} = 5.24$) y Funcionamiento Global ($M_{pre} = 4.79 - M_{post} = 5.14$). Respecto al grupo control, no existen cambios entre medidas.

Tabla 3. Diferencias en puntuaciones pre-post intragrupo					
		Grupo Experimental (n = 129) Consenso (n=119)		Grupo Control (n = 286) Consenso (n=237)	
		Media	DT	Media	DT
Funcionamiento global	Pre	4.79	.74	5.12	.60
	Post	5.14	0.69	5.12	.58
Clima familiar	Pre	5.31	1.03	5.61	.89
	Post	5.61	.93	5.64	.87
Estrés familiar (invertida)	Pre	3.83	1.17	4.09	1.03
	Post	3.95	1.10	4.08	1.13
Normas	Pre	5.01	.95	5.33	.84
	Post	5.39	.89	5.35	.81
Consenso	Pre	5.27	1.04	5.54	.97
	Post	5.64	.91	5.57	.89
Gestión de conflictos	Pre	4.67	1.03	5.08	.79
	Post	5.24	.94	5.04	.76

La Figura 2 muestra la información de forma visual.

Figura 2. Trayectoria pre-post de grupo experimental y control por dimensión.



Nota. Líneas: medias observadas pre-test y post-test por grupo, en la escala original 1–7. Barras de error: IC 95% de la media en cada momento. En todas las dimensiones positivas, valores más altos indican mejor funcionamiento. En Estrés familiar (escala invertida), valores más altos indican menor estrés percibido. $n = 415$ para todas las dimensiones excepto Consenso parental ($n = 354$).

4.2. Efectos del programa sobre el funcionamiento familiar (pretest-postest)

4.2.1. Análisis ANCOVA

La Tabla 4 recoge los resultados ANCOVA completos. Los resultados muestran que, tras el ajuste inicial de las puntuaciones postest por el valor del pretest del análisis ANCOVA, el grupo experimental obtiene puntuaciones significativamente más altas en cinco de las seis dimensiones evaluadas.

El grupo experimental mejoró respecto a sus propias puntuaciones y también en relación al grupo control. Los tamaños del efecto oscilaron entre pequeños y moderados. Las mayores diferencias se observaron en gestión de conflictos ($p < .001$; $d = .69$) y funcionamiento global ($p < .001$; $d = .58$), seguidas de Establecimiento de Normas y Límites ($p < .001$; $d = .40$), consenso parental ($p < .01$; $d = .31$) y clima familiar ($p < .01$; $d = .29$). Respecto a la dimensión de Estrés Familiar, no se encontraron diferencias significativas entre grupos.

Tabla 4. Análisis ANCOVA: Efecto del programa EFEPV sobre las dimensiones de la escala EF-FAD

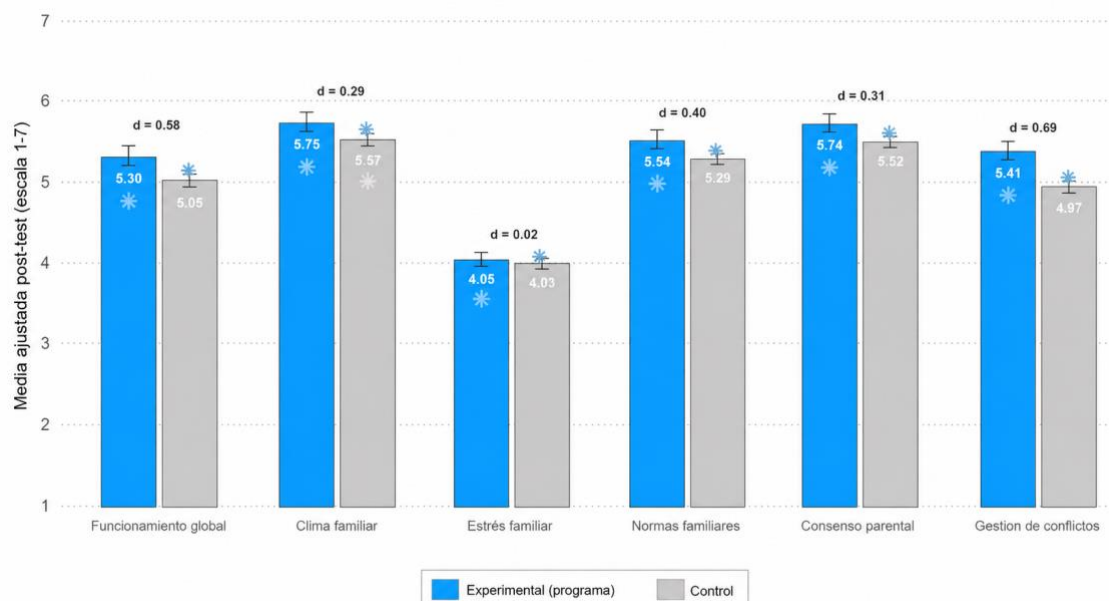
	n	EMM Exp	EMM Ctrl	Dif. ajustada [IC 95%]	p	η^2p	d
Funcionamiento global	415	5.30	5.05	.25 [.16; .34]	***	.064	.58
Clima familiar	415	5.75	5.57	.18 [.05;.31]	**	.017	.29
Estrés familiar (invertida)	415	4.05	4.03	.02 [-.18; .21]	n.s.	< .001	.02
Normas	415	5.54	5.29	.25 [.12; .38]	***	.033	.40
Consenso	352	5.74	5.52	.22 [.06; .37]	**	.021	.31
Gestión de conflictos	415	5.41	4.97	.44 [.30; .58]	***	.089	.69

* $p < .05$; ** $p < .01$; ***: $p < .001$; (n.s.): no significativo; d: Estadístico d de Cohen (tamaño de efecto); EMM: media marginal estimada del modelo ANCOVA ajustado por la puntuación pre-test correspondiente. Diferencia ajustada = Experimental – Control; p: significación.

La n para Consenso parental es inferior porque los ítems P30–P32 no se administran a quienes no comparten crianza.

En Estrés familiar (escala invertida), valores positivos de la diferencia indicarían menor estrés percibido en el grupo experimental

Para interpretar correctamente estos resultados, es importante tener en cuenta que las diferencias observadas entre grupos en el posttest, el análisis ANCOVA compara las puntuaciones finales ajustando estadísticamente las diferencias iniciales (pretest). En otras palabras, el análisis estima si ambos grupos seguirían mostrando diferencias al finalizar la intervención en el caso de haber partido exactamente del mismo nivel inicial. Así, el foco interpretativo principal no se sitúa tanto en si cada grupo mejoró respecto a sí mismo entre el pretest y el posttest, sino en si el grupo experimental presenta resultados superiores al grupo control una vez ajustadas las diferencias en la medida inicial.

Figura 3. Medias marginales estimadas del ANCOVA por grupo


Nota. Barras: medias ajustadas posttest (ANCOVA con la puntuación pretest como covariable). Barras de error: IC 95%. Los asteriscos (*) sobre cada barra indican la media pretest observada del grupo correspondiente. La distancia entre el asterisco (PRE) y la parte superior de la barra (POST ajustado) refleja la magnitud del cambio. En todas las dimensiones excepto Estrés Familiar (escala invertida), valores más altos indican mejor funcionamiento. En Estrés Familiar, valores más altos indican menor estrés percibido. d: Estadístico de Cohen (tamaño del efecto entre grupos).

4.2.2. Mejora fiable.

Para conocer el porcentaje de grupo experimental que mostraba una mejora fiable, junto con el análisis de las medias grupales (efectos medios estimados), resultaba importante examinar también el patrón de respuesta a nivel individual mediante Reliable Change Index (RCI).

La Tabla 5 y la Figura 4 muestran, para cada dimensión, el porcentaje de participantes que presentan una mejora o un empeoramiento suficientemente grandes como para considerarse cambios reales y no fluctuaciones debidas al error de medida del cuestionario.

Tabla 5. Porcentaje de participantes con cambio fiable pre-post (Reliable Change Index) por dimensión y grupo

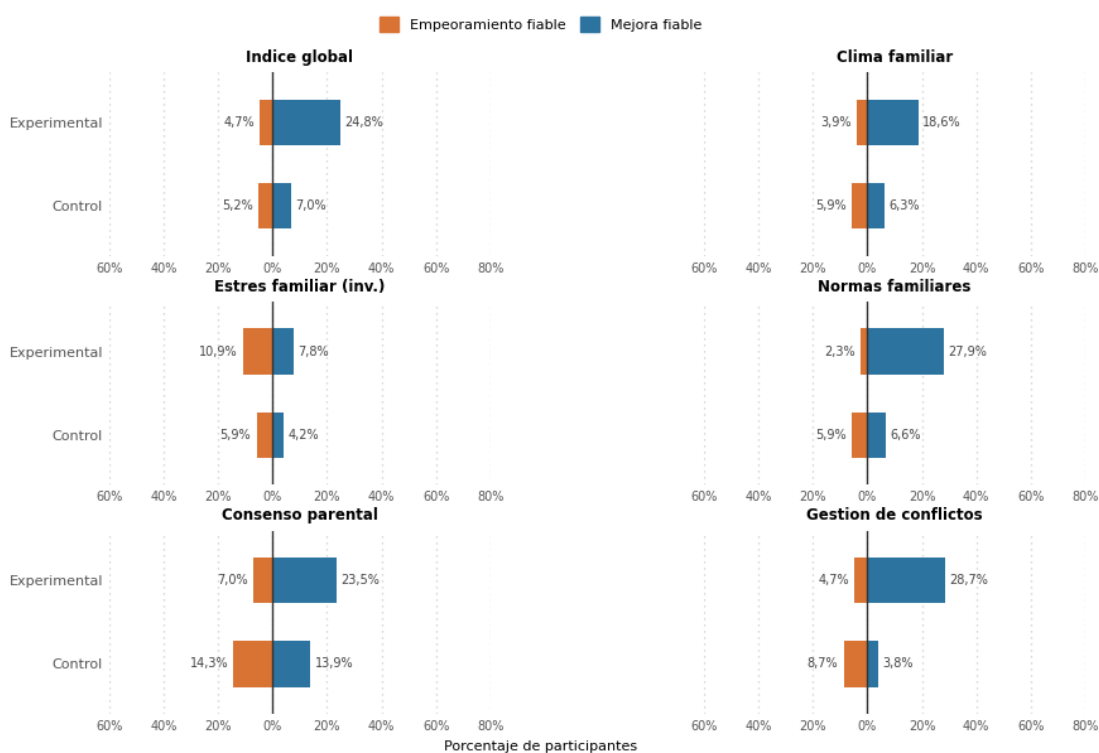
Dimensión	Mejora fiable Experimental (%)	Mejora fiable Control (%)	Empeoramiento fiable Experimental (%)	Empeoramiento fiable Control (%)
Funcionamiento global	24.8	7.0	4.7	5.2
Clima familiar	18.6	6.3	3.9	5.9
Estrés familiar (invertida)	7.8	4.2	10.9	5.9
Normas	27.9	6.6	2.3	5.9
Consenso	23.5	13.9	7.0	14.3
Gestión de conflictos	28.7	3.8	4.7	8.7

Nota. El Reliable Change Index (Jacobson y Truax, 1991) clasifica a cada participante en mejora fiable, sin cambio fiable o empeoramiento fiable a partir del error de medida específico de cada dimensión. Para Consenso parental las n efectivas son inferiores por la restricción de diseño de la escala. El porcentaje complementario al 100% en cada celda corresponde a participantes sin cambio fiable. Las fiabilidades empleadas (α de Cronbach) se calcularon sobre los ítems pre-test en la muestra pareada.

Los resultados refuerzan el patrón observado en los análisis grupales, mostrando que las familias del grupo experimental presentan porcentajes claramente superiores de mejora fiable respecto al grupo control, especialmente en Gestión de Conflictos (28.7% vs. 3.8%), Establecimiento de Normas Familiares (27.9% vs. 6.6%) e índice global de Funcionamiento Familiar (24.8% vs. 7.0%).

Así, por poner algunos ejemplos, respecto a Gestión de Conflictos, el 28.7% de los participantes del programa experimentaron una mejora individualmente detectable, frente al 3,8% del grupo control. En cuanto a Establecimiento de Normas, un 27.9% del grupo experimental mejoró de forma fiable a nivel individual frente al 6.6% del grupo control. Respecto al Funcionamiento Familiar Global, un 24.8% de los participantes experimentales mejoraron individualmente, frente a un 7% del grupo control. Además, los porcentajes de empeoramiento fiable en el grupo experimental son similares y en la mayoría de dimensiones inferiores a los del grupo control.

Figura 4. Distribución de participantes por dirección del cambio fiable, por dimensión y grupo



Nota. Barras divergentes: porcentaje de participantes con mejora fiable (a la derecha del eje) y con empeoramiento fiable (a la izquierda del eje) según el Reliable Change Index. El complemento al 100% en cada barra corresponde a participantes sin cambio fiable.

3.2.3. Variables moderadoras

Para conocer si existen perfiles en los que el programa EFEPV funcione de forma distinta, se evaluaron dos potenciales moderadores: a) el sexo de la persona participante y b) la comunidad autónoma de implementación.

4.2.3.1. Sexo del participante

La interacción Grupo × Sexo no resultó estadísticamente significativa en ninguna de las variables evaluadas (F entre .001 y 1.10; todos los $p \geq .295$). De este modo, el programa produce efectos comparables en hombres y mujeres.

4.2.3.2. Comunidad autónoma de residencia

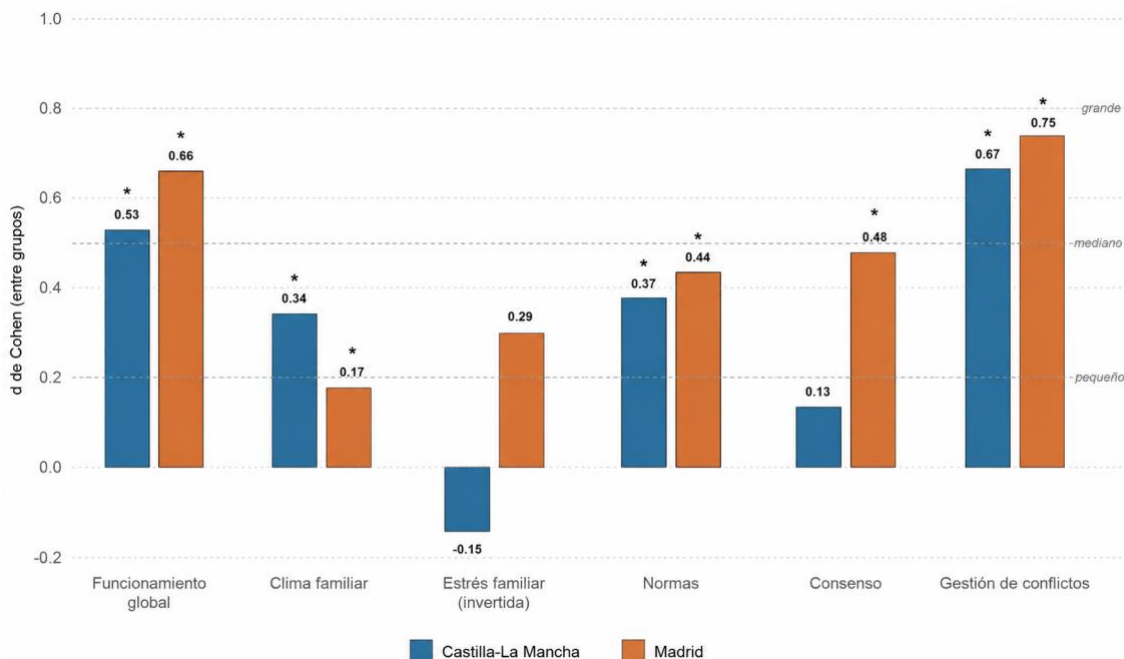
La interacción Grupo × Comunidad autónoma tampoco fue significativa en ninguna variable (F entre .00 y 2.73; todos los $p \geq .099$). En términos inferenciales, no hay evidencia de que el patrón de mejora difiera entre comunidades.

A título descriptivo y exploratorio, se aporta la magnitud de efecto del programa en cada comunidad autónoma, con el objetivo de comprobar tendencias de estabilidad territorial. Así, la Tabla 6 y la Figura 5 presentan los tamaños del efecto entre grupos estimados mediante el mismo ANCOVA del análisis principal, estratificados por comunidad.

Tabla 6. Tamaño del efecto entre grupos (d de Cohen) por dimensión y comunidad autónoma

	Castilla La Mancha p	Castilla La Mancha (n = 227) d	C.A. Madrid p	C.A. Madrid (n = 188) d
Funcionamiento global	< .001	.53	< .001	.66
Clima familiar	.017	.34	.313	.17
Estrés familiar (inv.)	.293	-.15	.095	.29
Normas familiares	.009	.37	.012	.44
Consenso parental	.381	.13	.008	.48
Gestión de conflictos	< .001	.67	< .001	.75

Nota. Tamaños del efecto entre grupos (Experimental vs. Control) estimados mediante el ANCOVA del análisis principal (POST ~ PRE + Grupo), ejecutado por separado en cada subgrupo territorial. En Consenso parental las n efectivas son inferiores por la restricción de diseño de la escala (n = 186 en Castilla-La Mancha y n = 166 en Madrid). Los contrastes formales de interacción Grupo × Comunidad no son significativos en ninguna dimensión (todos los $p \geq ,099$); las diferencias entre comunidades visibles en esta tabla son por tanto descriptivas y no soportan inferencia sobre patrones diferenciales por territorio.

Figura 5. Tamaño del efecto entre grupos por dimensión y comunidad autónoma


Nota. *: $p < ,05$; Tamaños del efecto entre grupos (Experimental vs. Control) estimados mediante ANCOVA con la puntuación pre-test como covariable, estratificados por comunidad autónoma. Líneas discontinuas: umbrales convencionales de Cohen (pequeño = .20, mediano = .50, grande = .80).

En ambas comunidades se observan tamaños del efecto positivos en Funcionamiento Global, Establecimiento de Normas y Gestión de Conflictos. En este sentido, el programa funciona en estas tres dimensiones de forma comparable en los dos territorios evaluados. Los mayores tamaños del efecto se registran en gestión de conflictos, tanto en Castilla-La Mancha ($d = .67$) como en Madrid ($d = .75$), seguidos de funcionamiento global ($d = .53$ y $d = .66$, respectivamente). Respecto a Establecimiento de Normas Familiares, los tamaños del efecto son moderados en ambas comunidades (Castilla-La Mancha: $d = .37$; Madrid: $d = .44$).

Por su parte, Consenso/Corresponsabilidad Familiar, Clima Familiar y Estrés Familiar se comportan de forma divergente en la muestra evaluada. Así, para Consenso/Corresponsabilidad se observan diferencias entre comunidades, con un tamaño del efecto reducido en Castilla-La Mancha ($d = .13$) y moderado en Madrid ($d = .48$). En cuanto a Clima Familiar, Castilla-La Mancha presenta un tamaño del efecto pequeño de $d = .34$, mientras que Madrid muestra un tamaño menor ($d = .17$). Respecto al estrés familiar (invertida), mientras que Castilla-La Mancha presenta un tamaño del efecto negativo ($d = -.15$), Comunidad Autónoma de Madrid muestra un tamaño del efecto positivo reducido ($d = .29$).

4.3. Seguimiento a tres meses

Una submuestra de participantes completó una medida de seguimiento aproximadamente tres meses tras la finalización del programa. La n efectiva con datos en las tres olas (pretest, postest y seguimiento) es de 16 participantes en el grupo experimental ($n = 12$ en Consenso/Corresponsabilidad) y de 95 participantes en el grupo control ($n = 76$ en Consenso/Corresponsabilidad). El desequilibrio entre grupos es relevante para la interpretación: el grupo control aporta n suficiente para caracterizar la trayectoria natural de las variables, pero el grupo experimental queda limitado en potencia estadística para los contrastes entre grupos en seguimiento. Los análisis que siguen combinan descripción de la trayectoria de cada grupo y una comparación inferencial entre grupos en el punto de seguimiento, leída con la prudencia que exige la asimetría muestral (Tablas 7 y 8).

Como puede observarse en la Tabla 7, el grupo control apenas muestra cambios entre el pretest y el postest y seguimiento. Las variaciones observadas son mínimas en todas las dimensiones, lo que indica una elevada estabilidad de las puntuaciones a lo largo del tiempo (cambios entre $+0.01$ y $+0.03$ en cinco dimensiones, $+0.02$ en Clima, -0.05 en Consenso).

Tabla 7. Descriptivos pre-test, post-test y seguimiento a tres meses en la submuestra con tres olas. Grupo control.

	n	$M (DT)$ PRE	$M (DT)$ POST	$M (DT)$ SEG	Δ PRE→POST	Δ POST→SEG	Reten. (%)
Funcionamiento global	95	5,04 (0,58)	5,05 (0,58)	5,00 (0,61)	+0.01	-0.05	n.p. ^b
Clima familiar	95	5,62 (0,90)	5,64 (0,88)	5,57 (0,85)	+0.02	-0.07	n.p. ^b
Estrés familiar (inv.)	95	3,66 (0,90)	3,69 (0,96)	3,72 (0,99)	+0.03	+0.03	n.p. ^b
Normas familiares	95	5,37 (0,77)	5,38 (0,71)	5,33 (0,72)	+0.01	-0.05	n.p. ^b
Consenso parental	76	5,69 (0,93)	5,64 (0,84)	5,66 (0,96)	-0.05	+0.02	n.p. ^b
Gestión de conflictos	95	5,04 (0,76)	5,04 (0,73)	4,94 (0,74)	+0.00	-0.10	n.p. ^b

Nota. El % de retención se calcula como $((M \text{ SEG} - M \text{ PRE}) / (M \text{ POST} - M \text{ PRE})) \times 100$; valores cercanos a 100% indican mantenimiento del efecto inmediato; > 100%, mejora adicional a tres meses; < 100%, retroceso parcial. ^b No procede: el cambio inmediato PRE→POST del grupo control fue prácticamente nulo en todas las dimensiones, por lo que la ratio de retención no es interpretable.

Por su parte, los resultados del seguimiento al grupo experimental (Tabla 8) muestran que el grupo no presenta pérdida de los efectos obtenidos inmediatamente tras la intervención en aquellas dimensiones en las que hubo efecto inmediato tras la intervención. Así, cuatro de las cinco dimensiones en las que el programa mostró efectos positivos inmediatos presentan porcentajes de retención del efecto iguales o superiores al 100% entre el posttest y el seguimiento, lo que sugiere un mantenimiento general de las mejoras observadas a corto plazo. Entre estas dos medidas, las puntuaciones se mantienen estables en algunas dimensiones (como Gestión de conflictos o Clima familiar) e incluso continúan mejorando ligeramente en otras (Funcionamiento Global, Establecimiento de Normas, Consenso/Corresponsabilidad y Estrés Familiar).

Tabla 8. Descriptivos pre-test, post-test y seguimiento a tres meses en la submuestra con tres olas. Grupo experimental

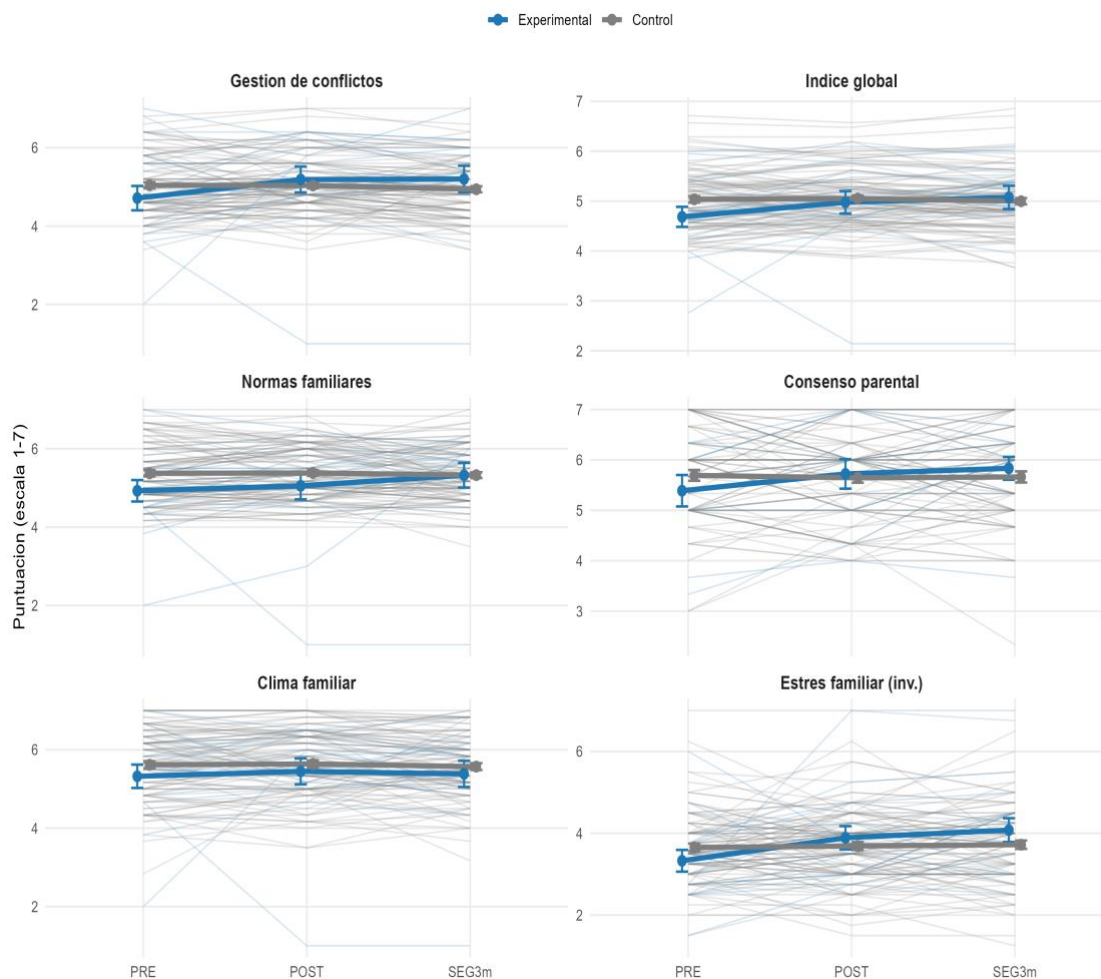
	<i>n</i>	<i>M</i> (DT) PRE	<i>M</i> (DT) POST	<i>M</i> (DT) SEG	Δ PRE→POST	Δ POST→SEG	Reten. (%)
Funcionamiento global	16	4,68 (0,80)	4,98 (0,90)	5,07 (0,93)	+0.29	+0.10	133,7
Clima familiar	16	5,32 (1,19)	5,45 (1,32)	5,39 (1,34)	+0.12	-0.06	50,0
Estrés familiar (inv.)	16	3,33 (1,06)	3,89 (1,14)	4,08 (1,16)	+0.56	+0.19	133,3
Normas familiares	16	4,93 (1,09)	5,05 (1,37)	5,32 (1,26)	+0.12	+0.27	316,7 ^a
Consenso parental	12	5,39 (1,08)	5,72 (1,01)	5,83 (0,78)	+0.33	+0.11	133,3
Gestión de conflictos	16	4,71 (1,24)	5,19 (1,32)	5,20 (1,35)	+0.47	+0.01	102,6

Nota. El % de retención se calcula como $((M \text{ SEG} - M \text{ PRE}) / (M \text{ POST} - M \text{ PRE})) \times 100$; valores cercanos a 100% indican mantenimiento del efecto inmediato; > 100%, mejora adicional a tres meses; < 100%, retroceso parcial. ^a Valor inestable por denominador pequeño: el cambio PRE→POST en Normas familiares en submuestra experimental (+0,12) es marcadamente menor que el observado en la muestra completa (+0,40); la transición POST→SEG (+0,27) indica consolidación tardía.

Es decir, la comparación de ambas tablas sugiere que las mejoras detectadas en el grupo experimental difícilmente pueden atribuirse a factores como la maduración natural de las familias entre puntuaciones, la regresión a la media o un simple efecto de repetición de la prueba.

Las trayectorias individuales (Figura 6) muestran cambios POST→SEG nullos o positivos en todas las dimensiones donde el programa produjo efecto inmediato, y la distribución del RCI individual confirma que la inmensa mayoría de los participantes mantiene sus niveles post-programa (Tabla 9).

Figura 6. Trayectorias individuales pretest → postest → seguimiento por dimensión y grupo



Nota. Cada línea fina representa la trayectoria de un participante a través de los tres momentos; las líneas gruesas son las medias del grupo en cada momento (azul: experimental; gris: control). La escala original es 1-7. En Estrés familiar (escala invertida), valores más altos indican menor estrés percibido. Submuestras con datos en las tres olas: experimental n = 16 (12 en Consenso parental); control n = 95 (76 en Consenso parental).

La distribución del RCI POST→SEG en el grupo experimental (Tabla 9) es notablemente favorable al mantenimiento del efecto. Entre el 66.7% y el 93.8% de los participantes se mantienen estables según dimensión, las recaídas fiables son marginales (0% en cuatro de las seis dimensiones; un único caso en Gestión de conflictos y en Consenso parental) y, en Consenso parental, una cuarta parte de los participantes continúa mejorando tras el programa.

Tabla 9. Distribución de participantes según el Reliable Change Index en la transición post-programa → seguimiento. Grupo Experimental

	<i>n</i>	Mejora adicional (%)	Estable (%)	Recaída fiable (%)
Funcionamiento global	16	12,5	87,5	0,0
Clima familiar	16	6,2	93,8	0,0
Estrés familiar (inv.)	16	12,5	87,5	0,0
Normas familiares	16	12,5	87,5	0,0
Consenso parental	12	25,0	66,7	8,3
Gestión de conflictos	16	6,2	87,5	6,2

Nota. Reliable Change Index aplicado a la transición post-programa → seguimiento, usando las mismas fiabilidades (α de Cronbach) estimadas para la Tabla 5 del informe 'Mejora adicional': RCI > 1,96. 'Estable': $|RCI| \leq 1,96$. 'Recaída fiable': RCI < -1,96.

El grupo control (Tabla 10) muestra tasas de estabilidad similares (76.3%-94.7%) y tasas de cambio fiable algo más simétricas (entre mejora adicional y recaída fiable), lo que es consistente con fluctuación natural en ausencia de intervención.

Tabla 10. Distribución de participantes según el Reliable Change Index en la transición post-programa → seguimiento. Grupo Control

	<i>n</i>	Mejora adicional (%)	Estable (%)	Recaída fiable (%)
Funcionamiento global	95	4,2	90,5	5,3
Clima familiar	95	7,4	80,0	12,6
Estrés familiar(inv.)	95	3,2	94,7	2,1
Normas familiares	95	5,3	87,4	7,4
Consenso parental	76	10,5	76,3	13,2
Gestión de conflictos	95	5,3	90,5	4,2

Nota. Reliable Change Index aplicado a la transición post-programa → seguimiento, usando las mismas fiabilidades (α de Cronbach) estimadas para la Tabla 5 del informe 'Mejora adicional': RCI > 1,96. 'Estable': $|RCI| \leq 1,96$. 'Recaída fiable': RCI < -1,96. La distribución en el grupo control permite estimar la 'tasa basal' de cambios fiables atribuibles al error de medida o a fluctuación natural en ausencia de intervención.

Sin embargo, las comparaciones entre grupo experimental y grupo control en el seguimiento no alcanzan significación estadística en ninguna dimensión (ver Tabla 11). No obstante, este resultado debe interpretarse con cautela. Aunque las diferencias no son estadísticamente significativas, en cuatro de las seis dimensiones las diferencias ajustadas continúan siendo favorables al grupo experimental y los tamaños del efecto se sitúan, entre .27 y .41 en cuatro de ellas (magnitudes pequeño-medianas), y solo en Clima familiar la diferencia es prácticamente inexistente ($d = -.07$).

Tabla 11. Contraste entre grupos en seguimiento (ANCOVA ajustado por post-test)

Dimensión	n Exp	n Ctrl	M SEG Exp	M SEG Ctrl	Dif. ajustada [IC 95%]	p ANCOVA	d	p Welch
Funcionamiento global	16	95	5,07	5,00	+ .142 [-.043; .327]	.130 (n.s.)	.41	.754
Clima familiar	16	95	5,39	5,57	- .041 [-.365; .282]	.801 (n.s.)	-.07	.599
Estrés familiar (inv.)	16	95	4,08	3,72	+ .205 [-.169; .580]	.279 (n.s.)	.29	.261
Normas familiares	16	95	5,32	5,33	+ .226 [-.082; .535]	.149 (n.s.)	.40	.992
Consenso parental	12	81	5,83	5,64	+ .133 [-.340; .606]	.578 (n.s.)	.17	.461
Gestión de conflictos	16	95	5,20	4,94	+ .135 [-.134; .404]	.322 (n.s.)	.27	.461

Nota. ANCOVA con la puntuación post-test como covariable; el coeficiente del factor Grupo estima la diferencia entre experimental y control en la puntuación de seguimiento ajustando por nivel inmediatamente post-programa. d de Cohen estandarizado con la DT residual del ANCOVA. p Welch corresponde al contraste simple de medias en seguimiento sin covariable, reportado como robustez. Ningún contraste alcanza el umbral convencional $p < ,05$.

La explicación más probable de esta falta de significación estadística en la comparación entre grupos es el reducido tamaño de la muestra experimental disponible en seguimiento. La combinación de un número muy limitado de participantes experimentales y unos errores estándar amplios reduce considerablemente la potencia estadística del análisis,

Por tanto, de forma general, los descriptivos son compatibles con un efecto del programa que se sostiene a tres meses, sin recuperación de los niveles iniciales, mientras que el grupo control no presenta cambio en sus puntuaciones en las tres medidas, aunque el tamaño muestral actual no permite confirmarlo con el nivel habitual de robustez inferencial.

5.CONCLUSIONES, LIMITACIONES Y PROPUESTA DE MEJORA.

5.1. Conclusiones

5.1.1. Eficacia del programa EFEPV

- A nivel intergrupo, tras el ajuste estadístico de las diferencias iniciales mediante ANCOVA, el grupo experimental obtiene puntuaciones significativamente superiores al grupo control en cinco de las seis dimensiones evaluadas de la escala EF-FAD, lo que refuerza la interpretación de un posible efecto asociado a la intervención.
- Las mayores diferencias entre grupos se observan en Gestión de Conflictos ($d = .69$) y Funcionamiento Global ($d = .58$), seguidas de Establecimiento de Normas Familiares ($d = .40$), Consenso /Corresponsabilidad ($d = .31$) y Clima Familiar ($d = .29$), con tamaños del efecto pequeños y moderados.
- Las mejoras observadas en Gestión de Conflicto, Establecimiento de Normas y Consenso resultan coherentes con los objetivos específicos del programa relacionados con la organización familiar, la corresponsabilidad educativa y el establecimiento de límites claros.
- La dimensión que muestra un efecto más consistente es gestión de conflictos, tanto en el análisis ANCOVA como en el análisis de cambio fiable individual. Esto sugiere que el programa parece especialmente eficaz en proporcionar a las familias herramientas prácticas para afrontar desacuerdos, tensiones y dinámicas conflictivas cotidianas.
- Otro aspecto relevante es que los resultados muestran que el grupo experimental no solo mejora respecto a sí mismo, sino que iguala o supera al grupo control tras la intervención, pese a partir inicialmente de niveles inferiores de funcionamiento familiar.
- Este hallazgo resulta especialmente relevante desde una perspectiva preventiva, ya que sugiere que el programa podría estar funcionando precisamente en aquellas familias que presentan una mayor necesidad inicial de apoyo educativo y parental.
- En este sentido, los análisis de cambio fiable (RCI) muestran que una proporción relevante de familias del grupo experimental experimenta mejoras individuales detectables, especialmente en Gestión de Conflictos (28.7%), Normas Familiares (27.9%) y Funcionamiento Familiar Global (24.8%), superando claramente los porcentajes observados en el grupo control.
- Además, los porcentajes de empeoramiento fiable son reducidos en el grupo experimental y, en la mayoría de dimensiones, inferiores a los observados en el grupo control.
- En cuanto a la dimensión de la escala EF-FAD Estrés Familiar, que no observa cambios tras la intervención, consideramos que se asocia de forma transversal a las demás dimensiones evaluadas, actuando como una variable resultado producto del resto de dimensiones, y que a su vez puede influir en la percepción y evolución del clima familiar, la gestión de conflictos o el establecimiento de normas. La ausencia de cambios significativos en esta subescala podría deberse a varios factores: por un lado, al carácter más estructural y estable del estrés familiar, y por tanto, menos sensible a intervenciones breves; por otro, a la posibilidad de que las mejoras observadas en otras dimensiones aún no hayan tenido un impacto suficiente como para modificar los niveles de estrés auto-referidos en el corto plazo.
- Por otra parte, el estrés familiar percibido puede depender en gran medida de demandas externas (laborales, económicas, contextuales) que no son objeto directo del programa.

- La validación paralela de la EF-FAD (ver anexo 2) ya había sugerido esta lectura al mostrar que Estrés familiar se separa del factor general y se interpreta como una componente reactiva, más que estructural, del funcionamiento familiar. En este marco, la ausencia de efecto no debe leerse como un fallo del programa, sino como una delimitación honesta de su alcance.
- Así, los resultados sugieren que el programa produce mejoras especialmente en aspectos conductuales y organizativos de la dinámica familiar, como el establecimiento de normas, la resolución de conflictos y el funcionamiento familiar general.
- La estabilidad del grupo control entre el pretest y el postest refuerza la interpretación de que los cambios observados en el grupo experimental se encuentran asociados a la intervención y no únicamente al paso del tiempo o al efecto de repetición de la medida.

5.1.2. Eficacia en función de características de las familias

- No se observan diferencias significativas en la eficacia del programa en función del sexo de la persona participante ni de la comunidad autónoma de implementación.
- Aun así, la mayor presencia femenina de la muestra limita la potencia estadística para detectar posibles diferencias según sexo.
- A nivel descriptivo, el programa muestra un patrón relativamente estable entre Castilla-La Mancha y Comunidad de Madrid
- Respecto a las comunidades autónomas, ambas comunidades muestran mejoras de magnitud similar y estadísticamente significativa en relación a Gestión de Conflictos, Funcionamiento Global y Establecimiento de Normas Familiares.
- En cambio, las dimensiones de Clima familiar y Consenso parental muestran un patrón menos coherente entre comunidades, con tendencias en direcciones opuestas: en Clima familiar la mejora es de mayor magnitud en Castilla-La Mancha que en Madrid. En Consenso/Corresponsabilidad el patrón se invierte, con mayor mejora en Madrid que en Castilla-La Mancha. No obstante, las interacciones Grupo × Comunidad no resultan significativas en ninguna de las dos dimensiones, por lo que estas divergencias deben interpretarse como descriptivas y, con probabilidad, atribuibles a las propias características de la muestra.
- Respecto al estrés familiar, no se observan cambios significativos en ninguna de las dos comunidades, reproduciendo el patrón encontrado en los análisis generales.

5.1.3. Seguimiento

- En el grupo experimental, la mayoría de dimensiones muestran estabilidad o mantenimiento de las mejoras entre el postest y el seguimiento a tres meses.
- Las trayectorias de las puntuaciones individuales confirman que las puntuaciones se mantienen estables o incluso mejoran ligeramente entre el postest y el seguimiento a tres meses en las dimensiones donde el programa había mostrado efectos inmediatos. Además, el análisis del cambio fiable individual indica que la mayoría de participantes conserva en seguimiento los niveles alcanzados tras la intervención.
- Por su parte, el grupo control mantiene una estabilidad prácticamente total entre el pretest y el postest, con cambios mínimos en todas las dimensiones, lo que refuerza la atribución del cambio inmediato del grupo experimental a los efectos del programa.

- Sin embargo, la comparación entre grupo experimental y control no alcanza significación estadística en ninguna dimensión. Aun así, los tamaños del efecto se sitúan entre $d = 0,27$ y $d = 0,41$ en cuatro de las seis dimensiones a favor del grupo experimental, en dirección coherente con un mantenimiento del efecto.
- La razón más plausible de la no significación es la potencia limitada por la n experimental.

5.2. Limitaciones metodológicas

- Las conclusiones deben interpretarse con cautela debido a las limitaciones inherentes al diseño cuasi-experimental utilizado, especialmente la ausencia de asignación aleatoria entre grupos. Esta limitación se reconoce explícitamente y se compensa, en la medida de lo posible, con una estrategia analítica que controla las diferencias iniciales entre grupos y triangula varias fuentes de evidencia sobre el cambio.
- El grupo experimental presentaba inicialmente niveles inferiores de funcionamiento familiar y mayor presencia de problemáticas conductuales o de riesgo respecto al grupo control. Aunque estas diferencias fueron ajustadas estadísticamente mediante ANCOVA, no puede descartarse completamente el efecto de procesos de autoselección.
- La muestra presenta una mayor presencia femenina, aspecto habitual en programas de parentalidad, pero que limita la generalización de los resultados a figuras paternas u otros perfiles familiares menos representados.
- Asimismo, las medidas utilizadas son auto-informadas, por lo que las respuestas pueden verse influidas por factores como deseabilidad social, sensibilidad hacia la crianza o diferencias en la capacidad de autoobservación familiar.
- La duración relativamente breve de la intervención puede estar limitando la magnitud de los cambios observables, especialmente en dimensiones más emocionales o estructurales como el estrés familiar.
- En este sentido, el estudio evalúa principalmente cambios percibidos en funcionamiento familiar, pero no incorpora indicadores externos o conductuales relacionados con prevención de conductas de riesgo, lo que sería deseable en futuras evaluaciones.
- A nivel metodológico, la pérdida muestral fue especialmente elevada en el grupo experimental, lo que limita la potencia estadística y la representatividad de los análisis longitudinales. La elevada pérdida muestral observada parece responder a varios motivos. Por una parte, la disponibilidad de datos longitudinales completos en el grupo experimental fue limitada al cierre de la recogida de datos para la realización de este informe. Así, una parte importante de los casos clasificados como “solo PRE” corresponde a programas todavía en curso de implementación y, por tanto, sin posibilidad temporal de disponer aún de medida posttest. Otro porcentaje de datos no emparejados puede explicarse por el abandono real de participantes del programa. Por último, otro porcentaje se explica por dificultades logísticas en la recogida de medidas posttest.
- Esta interpretación se ve reforzada por la presencia de participantes con medidas pretest y seguimiento disponibles sin correspondencia emparejada en el posttest, lo que sugiere posibles pérdidas administrativas en la evaluación longitudinal. Se recomienda encarecidamente que en futuras evaluaciones se preste especial atención a estos aspectos metodológicos, ya que una mayor muestra aportaría mayor potencia estadística a la evaluación y representaría la diversidad de la muestra de forma más ajustada.
- Estos datos de pérdida muestral reflejan las dificultades inherentes a la evaluación longitudinal de programas comunitarios implementados en contextos naturales, especialmente cuando participan múltiples equipos, territorios y calendarios de ejecución heterogéneos.

- En este sentido, no fue posible distinguir con precisión entre abandono real del programa y ausencia de respuesta en la evaluación posttest, debido a la falta de registros sistemáticos de asistencia.
- A pesar de estas limitaciones, el análisis permitió identificar un conjunto de participantes con medidas longitudinales válidas suficientes para evaluar preliminarmente los efectos del programa.
- En esta misma línea de limitaciones, el tamaño de la submuestra de seguimiento en el grupo experimental es reducido, mientras la submuestra control aporta n suficiente, lo que limita la potencia del contraste inferencial entre experimental y control en seguimiento, aunque sí permite caracterizar la trayectoria natural del grupo control.
- Por el propio diseño de la escala EF-FAD, los ítems del factor “consenso parental” no se aplicaban a las personas que no compartían tareas educativas y crianza, lo que disminuye la muestra para esa dimensión en los análisis.
- Por otra parte, la intervención se implementó en condiciones reales y con distintos equipos dinamizadores, pero no se dispone de indicadores sistemáticos de fidelidad de implementación. Por ello, no puede determinarse hasta qué punto todas las aplicaciones del programa reprodujeron de manera homogénea los contenidos y metodología originalmente diseñados.
- La recogida de variables no evaluadas en esta ocasión permitiría un mayor nivel de análisis. En este sentido, recoger información, por ejemplo, sobre el número de sesiones a las que efectivamente han asistido cada uno/a de los/las familiares sobre el total de sesiones, sería de interés para comprender la relación entre el grado de exposición al programa y los resultados obtenidos.

5.3. Propuestas de mejora y recomendaciones

- Mantener y reforzar especialmente los módulos relacionados con Gestión de Conflictos y Normas Familiares, al tratarse de las áreas donde el programa muestra efectos más sólidos y consistentes y con mayor estabilidad temporal de las mejoras observadas en el seguimiento.
- Reforzar contenidos relacionados con regulación emocional, manejo del estrés y sobrecarga parental, dado que la dimensión de Estrés Familiar no muestra mejoras relevantes tras la intervención.
- La elevada pérdida muestral en el grupo experimental limita la solidez inferencial del estudio y dificulta el seguimiento longitudinal de los participantes. En este sentido, aunque los análisis realizados no sugieren una autoselección asociada al nivel basal de funcionamiento familiar, sí parece recomendable reforzar los sistemas de seguimiento y emparejamiento de participantes, así como incorporar registros sistemáticos de asistencia y adherencia al programa, lo que permitiría diferenciar entre abandono real del programa y pérdida de respuesta en las evaluaciones.
- Se recomienda asimismo reforzar la coordinación con los equipos dinamizadores para garantizar la administración sistemática de las medidas posttest y seguimiento a tres meses.
- Respecto a la presencia femenina mayoritaria en la muestra, se sugiere la conveniencia de desarrollar estrategias específicas de captación e implicación de figuras paternas y otros referentes educativos familiares, con el objetivo de favorecer una participación más equilibrada y aumentar el potencial impacto sistémico del programa sobre la dinámica familiar.

- Como recomendaciones generales, se proponen:
 - Continuar consolidando la evaluación sistemática del programa mediante nuevas ediciones y muestras más amplias, con el objetivo de estabilizar la evidencia empírica acumulada sobre la eficacia del EFEPV.
 - Profundizar en el análisis diferencial de la eficacia del programa en función de distintas tipologías familiares y variables sociodemográficas.
 - Incorporar diseños multiinformante que permitan analizar discrepancias y coherencia entre distintos miembros del sistema familiar.
 - Incorporar en futuras evaluaciones variables complementarias relacionadas con asistencia efectiva, implicación familiar y posibles cambios en conductas de riesgo o convivencia familiar.
 - Incorporar indicadores complementarios del funcionamiento familiar y medidas externas que permitan reforzar la validez convergente de la escala EF-FAD.

6. REFERENCIAS BIBLIOGRÁFICAS

- Cabrera-Perona, V., Estévez-García, J.F., Blanquer, M. & Lloret-Irles, D. (en revisión). Psychometric Validation of the EF-FAD Scale: a Brief Measure of Family Functioning for Universal Prevention Contexts. *Journal of Child and Family Studies*.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2.ª ed.). Lawrence Erlbaum Associates. University of Toronto
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19. <https://doi.org/10.1037/0022-006X.59.1.12>
- Lord, F. M. (1967). A paradox in the interpretation of group comparisons. *Psychological Bulletin, 68*(5), 304–305. <https://doi.org/10.1037/h0025105> PubMed
- Maxwell, S. E., Delaney, H. D., & Kelley, K. (2018). *Designing experiments and analyzing data: A model comparison perspective* (3.ª ed.). Routledge. Designingexperiments
- R Core Team. (2025). *R: A language and environment for statistical computing* [Software]. R Foundation for Statistical Computing. <https://www.R-project.org/>
- Sandler, I. N., Schoenfelder, E. N., Wolchik, S. A., & MacKinnon, D. P. (2011). Long-term impact of prevention programs to promote effective parenting: Lasting effects but uncertain processes. *Annual Review of Psychology, 62*, 299–329. <https://doi.org/10.1146/annurev.psych.121208.131619> Annual Reviews
- Van Breukelen, G. J. P. (2006). ANCOVA versus change from baseline: More power in randomized studies, more bias in nonrandomized studies. *Journal of Clinical Epidemiology, 59*(9), 920–925. <https://doi.org/10.1016/j.jclinepi.2006.02.007> Maastricht University
- Van Ryzin, M. J., Roseth, C. J., Fosco, G. M., Lee, Y.-K., & Chen, I.-C. (2016). A component-centered meta-analysis of family-based prevention programs for adolescent substance use. *Clinical Psychology Review, 45*, 72–80. <https://doi.org/10.1016/j.cpr.2016.03.007> HealthEvidence

Si desea recibir más información:

Fad Juventud (FAD)
Av. Burgos 1
28036 Madrid
Teléfono: +34 91 302 69 47
familias@fad.es

Financiado por



"Las opiniones vertidas en este documento no son necesariamente las de la Delegación del Gobierno para el Plan Nacional sobre Drogas o las del Ministerio de Sanidad".