

## Blueprints para Europa: Promoviendo Programas Basados en la Evidencia en los Servicios de Atención a la Infancia

Nick Axford. The Social Research Unit, Dartington, UK

Delbert Elliott. University of Colorado at Boulder, and Founding Director of Blueprints, USA

Michael Little. The Social Research Unit, Dartington, UK

Ésta es un versión traducida al español del original en inglés Blueprints for Europe: Promoting Evidence-Based Programs in Children's Services.

Para referenciar, citar el original como: Nick Axford, Delbert Elliott y Michael Little (2012). Blueprints for Europe: Promoting Evidence-Based Programs in Children's Services. *Psychosocial Intervention*, 21, 205-214. <http://dx.doi.org/10.5093/in2012a11>

**Resumen.** En el artículo se describe la base de datos Blueprints de programas basados en la evidencia (EBPs) y su aplicación potencial en servicios de atención a la infancia en los países europeos. Se abordan los aspectos relevantes del contexto europeo, y se señala una cierta tendencia al escepticismo ante programas importados de los EEUU, así como la necesidad de una fuente de información paneuropea sobre EBPs en todas las áreas de resultados. A continuación, se describen los criterios de evidencia empleados en los Blueprints, tales como la especificidad de la intervención, la calidad de la evaluación, el impacto de la intervención y la disponibilidad del sistema. Se examinan los criterios que se siguen para que un programa sea considerado como “Modelo” y “Prometedor”. A continuación, se resume el proceso mediante el que se desarrollaron los criterios y algunos de los aspectos de más difícil resolución. También se describen las líneas generales del proceso de incorporación de un programa a la base de datos de Blueprints, y se proporcionan tres ejemplos de programas aprobados por los Blueprints e implantados en Europa: un programa de visitas domiciliarias a madres de menores, un programa de enseñanza de habilidades parentales, y una intervención terapéutica en familias de delinquentes reincidentes. También se indica brevemente que la mayor parte de programas revisados no cumple con los criterios de evidencia. Finalmente, el artículo resume las orientaciones futuras de trabajo en este campo, con especial énfasis en las distintas estrategias para extender el uso de Blueprints por toda Europa.

*Palabras clave:* bienestar infantil, Europa, evaluación del programa, programa basado en la evidencia.

Durante la última década ha surgido un interés cada vez mayor en los programas basados en la evidencia (EBPs) en los países desarrollados. Un “programa” es un paquete de prácticas organizadas y diferenciadas (en ocasiones llamado manual o protocolo) que sirve de guía y que explica qué debe ser administrado y a quién, cuándo, dónde y cómo. Un programa está basado en la evidencia cuando ha sido evaluado exhaustivamente, normalmente mediante un ensayo controlado aleatorizado (RCT) o un diseño cuasiexperimental (QED), y ha demostrado de forma inequívoca que posee un efecto positivo en uno o más resultados infantiles relevantes (Social Research Unit, 2012a).

Existen en la actualidad muchos programas de este tipo que cubren todas las áreas de las vidas de los niños, incluyendo educación, conducta, salud, relaciones y bienestar emocional. Sin embargo, su penetración en el mercado es bastante escasa. Se adoptan en pocas ocasiones y, cuando se hace, la implementación es con frecuencia deficiente y los programas se diluyen en cuanto las subvenciones iniciales se agotan (Bumbarger y Perkins, 2008). Esto ocurre en los EEUU, donde se originan la mayoría de EBPs, e incluso en mayor medida en Europa, donde los distintos países importan cada vez más estos programas.

Existen varias razones posibles que explican esta situación (Little, 2010). En primer lugar, hay una falta generalizada de conocimientos entre los legisladores y profesionales de mayor edad sobre la existencia y la naturaleza de los EBPs. Algunos quizás otorguen escaso valor a la evidencia. En segundo lugar, existe cierta confusión sobre lo que significa “basado en la evidencia” y qué es un estándar apropiado. En tercer lugar, existen ciertas dudas sobre si estos programas pueden ser implementados en entornos reales, y sobre los recursos humanos y económicos que requiere su implementación.

Dichos obstáculos para la implantación de EBPs se magnifican en Europa, donde además existe cierta resistencia tanto ante el concepto de un programa como por el hecho de que la mayoría de los EBPs más conocidos se originan en los EEUU (se suelen mezclar ambas ideas en los debates). Los estados en regímenes de bienestar social-demócratas o católicos tienden a mostrar cierto escepticismo ante programas procedentes de los EEUU (Grietens, 2010). Esto es en parte debido a que algunos programas han obtenido resultados poco claros o decepcionantes al ser implantados en Europa. Por ejemplo, los resultados obtenidos con la *Terapia Multisistémica* fueron tan sólo ligeramente mejores que los de resultados de servicios habituales en el Reino Unido (Butler, Baruch, Hickey, y Fonagy, 2011), mientras que en Suecia los servicios regulares obtuvieron similares resultados (Sundell et al., 2008).

Estas dudas también reflejan las diferencias culturales entre Europa y los EEUU en cuanto a la provisión de servicios. Los EEUU tienen un sistema mínimo de prestaciones sociales. Por el contrario, los países del norte de Europa poseen más políticas de redistribución, así como un sistema universal de prestaciones sociales (Rowlands, 2010). En estos países también invierten en la profesionalización del personal de los servicios de atención a la infancia, principalmente a través de pedagogos sociales entrenados en desarrollo infantil que administran terapia a niños y familias en muchos entornos (Petrie, Boddy, Cameron, Wigfall, y Simon, 2006). Existe cierta resistencia a adoptar las prácticas de un país (EEUU) que habitualmente permanece en los puestos bajos de las clasificaciones internacionales de bienestar infantil, particularmente porque muchos países europeos, especialmente los países escandinavos, ya obtienen unos resultados relativamente buenos (UNICEF, 2007).

A este escepticismo se une el hecho de que no existe una fuente de información pan-Europea sobre los EBPs para una audiencia europea que cubra una variedad de áreas de intervención. Las bases de datos más conocidas de EBPs se publican en inglés y están principalmente dirigidos a proveedores norteamericanos. Apenas se incluyen programas europeos. Los proyectos de similar naturaleza en Europa tienden a ser de carácter exclusivamente nacional<sup>1</sup>, o se centran en sistemas específicos (por ejemplo, el sistema educativo)<sup>2</sup>, o se centran en un sólo aspecto, como por ejemplo la prevención y tratamiento del consumo de drogas<sup>3</sup>, o en una sola clase de programas (por ejemplo, intervenciones de prácticas parentales)<sup>4</sup>. Se ha reconocido la necesidad de la difusión de estos programas en distintos idiomas y para distintas culturas (Soydan, Mullen, Alexandra, Rehnman, y Li, 2010). Los proveedores de servicios de atención a la infancia también necesitan asesoramiento para seleccionar y adaptar programas a un contexto distinto. Esto incluye información acerca de cuáles son las poblaciones de niños y familias con las que los programas han demostrado tener éxito.

En este artículo se analizan los Blueprints, un recurso que proporcionará a los políticos y profesionales en Europa información de alta calidad sobre los programas que cumplen los altos estándares de evidencia y que están disponibles para su implementación en sistemas de servicios. Blueprints ha sido diseñado para promover un mayor conocimiento y uso de los EBPs y una mejora en el bienestar de los niños y familias que los reciben. Creemos que puede llenar un importante nicho en este área. Antes de continuar, resultaría útil proporcionar algunos antecedentes históricos.

Blueprints se inició en los EEUU y originalmente estaba enfocado a la prevención de la violencia tras la masacre en la escuela de Columbine en 1993, en la que dos estudiantes de instituto mataron a compañeros estudiantes y un profesor. Este fue el elemento catalizador para que el Center for the Study and Prevention of Violence de la Universidad de Colorado, en la cercana población de Boulder, comenzara a recopilar una lista de los programas basados en la evidencia específicamente dirigidos a prevenir la violencia (Elliott, 2010). En 2010, la Annie E. Casey Foundation dotó de apoyo económico a la Social Research Unit (SRU) en Dartington, en Reino Unido, y al Social Development Research Group (SDRG) en la Universidad de Washington, EEUU, para desarrollar un método que ayudara a los gestores de sistemas y las comunidades a trabajar mejor juntos para implementar los EBPs a escala. La idea era unir los programas de *Common Language* y *Communities that Care* de ambos centros (Axford y Morpeth, 2012; Hawkins y Catalano, 2002). El proyecto requería del desarrollo de una selección de programas que cubriesen todos los resultados evolutivos clave, y por ello se comenzó a elaborar una base de datos de EBPs.

Los investigadores de Blueprints para la Prevención de la Violencia se dedicaron a desarrollar los estándares de evidencia que sustentaban la base de datos de *Evidence2Success* (véase abajo). Además reconocie-

ron el valor de ampliar el espectro de Blueprints, tanto en términos geográficos como de resultados, y de ahí surgió la colaboración con Casey y la Social Research Unit. Tras lograr el respaldo económico de Casey, Blueprints para la Prevención de la Violencia:

- Fue rebautizado como “Blueprints para Conductas Problemáticas y Desarrollo Saludable de los Jóvenes”, convirtiéndose en la base de datos de *Evidence2Success* (no existe una base de datos separada para E2S).
- Adoptó básicamente el estándar de evidencia desarrollado para *Evidence2Success* (esto conllevó ciertos cambios tanto en los estándares ya existentes de Blueprints como en los estándares de *Evidence2Success* con el fin de alinearlos).
- Amplió su campo de acción e incluyó programas dirigidos a mejorar la salud, la educación, las relaciones, el bienestar emocional y la conducta de los niños.
- Estableció una oficina europea en Londres con personal de la Social Research Unit.
- Nombró a su primer representante europeo en el Consejo.
- Rediseñó su página web norteamericana y comenzó a trabajar en una página web europea (asegurándose de que existiera coherencia entre ambas en el contenido clave).

En el resto de este artículo, se describen los estándares de evidencia utilizados en la selección de programas para el nuevo Blueprints, cómo fueron desarrollados estos estándares, el proceso mediante el cual un programa es incorporado a la página web, la clase de programas que han sido aprobados, qué resultados obtienen los programas al ser contrastados con los estándares, y las futuras orientaciones de trabajo con especial referencia a Europa.

## Los estándares de evidencia

Los estándares de evidencia que respaldan a los Blueprints en la actualidad cubren cuatro dimensiones:

1. *Calidad de la Evaluación* – ¿el estudio sobre la eficacia y efectividad del programa produce resultados válidos y fiables?
2. *Impacto de la intervención* – ¿qué proporción de cambio positivo en resultados evolutivos clave puede ser atribuida al programa?
3. *Especificidad de la intervención* – ¿está el programa bien enfocado? ¿es práctico y lógico?
4. *Disponibilidad de su disseminación* – ¿cuenta el programa con el apoyo e información necesarios para posibilitar una implantación exitosa en las comunidades y en los sistemas de servicios públicos?

Dentro de cada dimensión, los Estándares contienen criterios “Suficientemente buenos” y “Óptimos”. Los criterios “Prometedores” establecen un estándar básico mínimo, mientras que los criterios “Modelo” refuerzan la confianza en el rigor científico de un programa, su impacto en los resultados, la especificidad de la intervención o su transportabilidad a escala. A continuación, abundamos en cada una de las cuatro dimensiones.

### *Calidad de la Evaluación*

Para que un programa aparezca en Blueprints, debe haber sido evaluado con al menos un buen ensayo controlado aleatorizado (RCT) o con dos buenos estudios de diseños cuasiexperimentales (QED). “Bueno/s” se refiere a aspectos de la calidad metodológica, específicamente que:

- El método de asignación a la intervención y el control se ajusten a su nivel (por ejemplo, individuo, centro escolar).
- Los instrumentos de medición sean apropiados para la población de la intervención en cuanto al enfoque y los resultados deseados.
- El análisis esté basado en intención de tratamiento.
- Los procedimientos y pruebas estadísticas sean los apropiados.
- Los grupos de la intervención y los grupos de control presenten una equivalencia de resultados clave en el periodo basal o que se incluyan en el análisis controles apropiados de las diferencias.

Tales estudios de evaluación también deben cumplir los siguientes criterios. Debe quedar claro con quién se ha testado el programa y qué tipo de tratamiento recibieron tanto los grupos de intervención como de referencia. Los instrumentos empleados deben ser válidos y fiables, capturar un resultado relevante, y no estar ligados al programa examinado. Con el fin de minimizar el sesgo, las observaciones, de mediciones y las evaluaciones deben ser aplicados por alguien *sin* ningún interés económico directo en el programa.

También se examina el nivel de abandonos por parte de los participantes en el estudio. Una cierta cantidad de abandonos es habitual, pero puede resultar problemático que muchos jóvenes abandonen si se dan más abandonos en unas categorías de jóvenes que en otras (por ejemplo, más en chicos que en chicas), o si el índice de abandonos y el tipo de persona que abandona difieren significativamente entre los grupos del programa y los grupos de control.

Hay varios criterios “Modelo” de evaluación de la calidad. Uno es simplemente si se han empleado como mínimo dos RCTs o un RCT y un QED que cumplan con los criterios de calidad. Generalmente, los resultados de un programa generan mayor confianza si este programa ha sido correctamente evaluado en más de una ocasión. Muchas evaluaciones sólo examinan el impacto al final del programa, lo cual resulta un tanto problemático ya que el impacto frecuentemente desaparece con el paso del tiempo. Por este motivo, también puntúa que se presente evidencia de un efecto a largo plazo. Se deben cumplir estos dos criterios para obtener la calificación de Programa Modelo.

También puntúa cualquier indicio de que los evaluadores hayan dedicado esfuerzos a entender más exhaustivamente el impacto del programa. Por ejemplo, podrían haber comprobado la relación entre la fidelidad en la implementación y el resultado, o entre la cantidad de programa recibido y el resultado. Si un programa ha sido administrado correctamente, o si algunos jóvenes o familias han recibido una mayor dosis, se espera un efecto más fuerte. Algunos estudios examinan si el programa funciona mejor para algunos de los subgrupos que para otros, centrándose en el género, la etnia y el nivel socioeconómico. Los estudios también pueden probar si la lógica en la que se basa el programa realmente se sostiene; ¿tienen lugar los efectos por las razones que se esperaban? Estas dos cuestiones son consideradas también criterios “Modelo”.

### ***Impacto de la intervención***

Los programas presentes en Blueprints deben tener un efecto positivo en un resultado relevante, cuyo tamaño sea conocido, y no presentar ningún efecto perjudicial. Sólo los estudios de evaluación que cumplen con todos los criterios de calidad “suficientemente buenos” pueden ser tenidos en consideración cuando se realice esta valoración.

La mayoría de los estudios que alcanzan este umbral deben probar que el programa tiene un efecto positivo en un resultado relevante para que se considere que posee un efecto positivo general. Un “efecto positivo” implica que los jóvenes o familias del grupo del programa obtuvieron mejores resultados que los jóvenes o familias del grupo de referencia. Es importante que este efecto no sea el resultado de la casualidad, de manera que debe ser “estadísticamente significativo”. También debe proporcionarse el tamaño de este efecto positivo.

No debería existir *ninguna* evidencia de que el programa posea un efecto perjudicial en los jóvenes o las familias. Esto incluye todos los resultados y a todos los subgrupos. Por ejemplo, un programa no será aprobado si mejora las relaciones de los adolescentes con iguales pero a expensas de que aumente su consumo de sustancias ilícitas. De igual forma, si el programa reduce el consumo de sustancias en chicos pero no en chicas, sería aprobado tan sólo para su uso con chicos.

Existen dos criterios “Modelo” para el impacto de la intervención. El primero es la existencia de varios estudios que cumplan los criterios “Prometedores” de calidad de evaluación y que la mayoría de estos muestren un impacto positivo estadísticamente significativo. El segundo es que exista evidencia de que los niños que recibieron una mayor dosis de programa obtuvieron mejores resultados que aquellos que recibieron menos: en otras palabras, que se dé una relación positiva entre dosis-efecto.

## ***Especificidad de la intervención***

Los programas deben establecer claramente cuáles son los resultados en los que pretenden incidir y a qué grupo de niños beneficiarán. Deben incluir una descripción precisa de los distintos componentes del programa y una explicación de por qué y cómo funciona el programa... dicho de otra manera, de qué forma el programa abordará los factores de riesgo y de protección como un medio de lograr los resultados.

Hay un criterio “Modelo” para la especificidad de la intervención, que es la existencia de evidencia empírica convincente que apoye la lógica del programa. Esta evidencia debe explicar por qué y de qué forma el programa probablemente beneficie a los niños y jóvenes a los que va dirigido. Por ejemplo, si un programa de enseñanza parental incita a los padres a practicar ciertas habilidades para enfrentarse a la conducta negativa de sus hijos, se deben presentar otros estudios que hayan demostrado que estas prácticas funcionan.

## ***Disponibilidad de la disseminación***

Los programas aceptados en la base de datos también deben probar que pueden ser implementados de manera generalizada en las comunidades y los sistemas de servicios. En el nivel más simple, se debe probar que el programa evaluado debe estar todavía disponible. Además, debería especificarse claramente de qué forma se puede hacer llegar el programa a los niños, jóvenes y familias adecuados. También son necesarios un manual y materiales de formación e implementación, porque ayudarán a confirmar que el programa es implantado coherentemente (o con fidelidad). Además, debería ofrecerse un listado de los recursos económicos y humanos necesarios para la implementación.

Existen varios criterios “Modelo” de disponibilidad de disseminación, comenzando por la disponibilidad de apoyo técnico para la implementación y de una lista de comprobación para supervisar la fidelidad.

Así mismo, se identifican los programas que están siendo disseminados ampliamente en ese momento, o que hayan sido probados y hayan probado su efectividad cuando son administrados por profesionales regulares en entornos normales. Muchos programas son inicialmente testados en unas condiciones especiales, por ejemplo, en clínicas universitarias y administrados por personal investigador. Los legisladores pueden tener una mayor confianza en programas que han sido probados cuando estos son administrados por la clase de personas que normalmente proporcionan servicios similares en su trabajo diario.

## **Desarrollo de los estándares**

En el curso de su labor en *Evidence2Success*, la SRU, el SDRG y la Fundación Annie E. Casey descubrieron que ya hay más de 25 bases de datos sobre programas basados en la evidencia y, por lo tanto, varios conjuntos de estándares<sup>5</sup>. En un intento de lograr cierto consenso, se tomó la decisión de desarrollar los estándares de evidencia de *Evidence2Success* (desde entonces, adoptados por Blueprints) y para ello se contó con la colaboración de expertos que anteriormente habían desarrollado conjuntos de estándares de evidencia y que listamos a continuación, todos los cuales a excepción del último han sido utilizados para informar las bases de datos de EBPs:

- Best Evidence Encyclopedia<sup>6</sup> (Robert E. Slavin, Johns Hopkins University, US).
- Blueprints for Violence Prevention<sup>7</sup> (Delbert S. Elliott, University of Colorado, US).
- LINKS (Lifecourse Interventions Nurturing Kids Successfully)<sup>8</sup> (Kristen Moore, Child Trends, US).
- Communities that Care<sup>9</sup> (J. David Hawkins y Richard F. Catalano, SDRG, University of Washington, US).
- Greater London Authority Project Oracle<sup>10</sup> (Michael Little, SRU en Dartington, UK).

Estos expertos mantuvieron reuniones periódicas durante 6 meses y probaron los estándares prototipo empíricamente para observar el nivel de dificultad de su aplicación y qué tipo de programas los cumplirían. También se prestó atención a otros conjuntos de estándares, como los desarrollados por la Society for Prevention Research (Flay et al., 2005) y las directrices CONSORT sobre informes de RTCs<sup>11</sup>. Durante

este periodo surgieron varios temas que precisaron debate. La mayoría se refieren a la calidad de la evaluación y el impacto de la intervención. A continuación, describimos brevemente algunos de los más importantes.

Se planteó un debate acerca de si el nivel de deserción de un estudio era importante y si debía especificarse un nivel en el que ya resultara problemático. Sin embargo, se argumentó que este enfoque penalizaría a los estudios de seguimiento, defendidos en los estándares, ya que tienden a presentar una deserción más alta. El grupo de expertos decidió en consecuencia centrarse principalmente en la deserción *diferencial*, en lugar de la general, aunque el Consejo de Blueprints sigue teniendo en cuenta los niveles de deserción y espera ciertos controles o ajustes (por ejemplo, puntuación de propensión) cuando los niveles son altos.

Otro de los temas se refería al valor de la replicación independiente, en otras palabras, el valor de un estudio que no es dirigido por el desarrollador del programa y que, sin embargo, prueba un impacto. Se pensaba que la participación del desarrollador podría introducir un sesgo (Eisner, 2009). El grupo de expertos acordó que, aunque la réplica independiente es la más deseable, y debiera ser promovida, insistir en ello en estos momentos provocaría que la lista quedara excesivamente reducida y eliminaría programas más sólidos, tales como el Nurse Family Partnership.

Hubo bastante debate acerca de si se debía aceptar evidencia del impacto derivada únicamente de medidas de auto-informe. En la justicia penal, tales medidas son por lo general consideradas aceptables. En educación, son inaceptables si son las únicas que se emplean; se prefieren las observaciones, las valoraciones del profesor y las calificaciones de pruebas académicas. El grupo de expertos decidió que debían centrarse en la adecuación de las medidas: ningún tipo de medida es incorrecta per sé. Sin embargo, insistieron que las medidas no pueden ser aplicadas únicamente a las personas que no ignoran la condición experimental o que suministran la intervención, porque esto podría introducir un sesgo.

También se desarrolló un extenso debate sobre la necesidad de que se presente evidencia de un impacto sostenido a los 12 meses tras la finalización de la intervención. Se expresó cierta preocupación de que se pudieran incluir muy pocos programas educativos, ya que la última medida que se realiza en estos programas es habitualmente al final de la intervención. La evidencia de un impacto sostenido fue considerado por lo tanto como un criterio “Óptimo”, en lugar de un criterio “Suficientemente bueno”.

El principal aspecto del impacto de la intervención que se trató se refiere al requisito de que se proporcione un tamaño del efecto estadísticamente significativo. Se argumentó que en muchos estudios con clusters amplios (como, por ejemplo, los centros escolares), es difícil obtener tamaños de muestra lo suficientemente grandes para permitir el análisis a nivel de clúster. Sin embargo, se pueden obtener cálculos significativos e imparciales mediante análisis a nivel de participante cuando los tamaños de muestra del nivel del participante son grandes. Por lo tanto, el grupo de expertos acordó como alternativa un tamaño del efecto de media ponderada de tamaño de la muestra de 0.2 y un tamaño de la muestra de más de 500 individuos en todos los estudios<sup>12</sup>.

Finalmente, era necesario tomar una decisión sobre los criterios para determinar el nivel general de un programa. Blueprints clasifica los programas en “Modelo” o “Prometedores”. Se acordó que un programa debe cumplir todos los criterios “Suficientemente buenos”<sup>13</sup> en las cuatro dimensiones para ser considerado “Prometedor” y que un programa puede ser designado “Modelo” si cumple estos criterios y:

- Tiene (a) dos o más ensayos controlados aleatorizados suficientemente buenos o (b) al menos un ensayo controlado aleatorizado bueno y una evaluación de diseño cuasiexperimental suficientemente buena.
- Hay pruebas de que se da un impacto sostenido (al menos hasta 12 meses tras la finalización del programa).

## **Cómo se incorporan los programas a Blueprints**

Estos estándares son aplicados a programas dirigidos a obtener resultados en áreas de la conducta, el bienestar emocional, las habilidades y logro académico, la salud (en especial la relacionada con aspectos comportamentales, como por ejemplo fumar, comer, beber), y las relaciones (principalmente con padres e iguales). El proceso de Blueprints para aprobar un programa consta de cuatro pasos.

En primer lugar, se identifica toda la literatura científica sobre el programa que resulte de relevancia. (Actualmente, esta revisión está restringida a publicaciones en lengua inglesa). El equipo de investigación revisa las principales publicaciones periódicas en todas las áreas de conducta problemática, salud y desarrollo infantil de forma regular para identificar la literatura sobre nuevos programas que pudieran ser incluidos en Blueprints, y añadir o cuestionar programas de la base de datos. También se estudia la información enviada por desarrolladores o proveedores de programas.

En segundo lugar, esta literatura es valorada según los estándares de evidencia por un equipo de reseñadores especializados de la Universidad de Colorado en los EEUU y la Social Research Unit en el Reino Unido. El resultado es una descripción narrativa estructurada de cada estudio y un resumen cuantitativo en el que se señala si en general el programa cumple con cada uno de los criterios de los distintos estándares. Las revisiones se centran principalmente en la especificidad de la intervención, la calidad de la evaluación y el impacto de la intervención. La calidad de cada revisión debe ser aprobada por un coordinador de revisiones.

En tercer lugar, los programas que se considera que puedan tener bastantes posibilidades de cumplir con los estándares de evidencia son enviados al Consejo de Blueprints para ser examinados. El Consejo está formado por ocho científicos punteros en el campo de la prevención de los EEUU y Europa y se reúnen dos veces al año. El Consejo decide qué programas cumplen con los estándares en cuanto a calidad de evaluación e impacto de la intervención y que, por lo tanto, pueden ser recomendados para su diseminación.

En cuarto lugar, el equipo de revisión comprueba la disponibilidad del sistema para los programas aprobados por el Consejo de Blueprints. Esto se hace consultando las páginas web de los programas y pidiendo a desarrolladores o proveedores que completen un cuestionario escrito. El cuestionario cubre aspectos como la disponibilidad de materiales y formación, procedimientos de supervisión de fidelidad, y necesidades de recursos humanos. Si se requiere mayor información tras presentar este cuestionario, entonces se envían preguntas de seguimiento y, con frecuencia, se mantiene una conversación telefónica. Dos miembros del grupo de revisión (al menos uno de ellos posee vasta experiencia en la administración y gestión de servicios) analizan la información recibida y determinan si el programa está “disponible para el sistema”. En esta fase un programa es oficialmente aprobado y se informa al desarrollador que será incluido en la página web de Blueprints.

La lista de programas aprobados se actualiza regularmente. Se realizan búsquedas regulares en la literatura utilizando un proceso consistente para identificar nuevos estudios que muestren resultados positivos, o incluso resultados negativos en el caso de programas que ya están en la lista. De manera similar, se revisan los estudios sobre nuevos programas si se considera que pueden cumplir con los estándares de evidencia. Se están realizando considerables esfuerzos para identificar los programas originados en Europa, ya que la mayoría de programas aprobados hasta el momento han sido desarrollados en los EEUU. Los desarrolladores y evaluadores de programas pueden remitir sus programas para su comprobación.

## **Programas aprobados por Blueprints**

En el momento en que se escribe este artículo existen 11 programas “Modelo” y 22 programas “Prometedores” en Blueprints. Estos son los programas aprobados de un total de 1000 programas revisados. Sin embargo, varios programas nuevos serán añadidos a la lista en breve, en gran parte debido a la decisión descrita más arriba sobre ampliar el espectro de Blueprints para acoger programas que van más allá de la prevención de la *violencia* (enfoque inicial) para abarcar otras áreas de desarrollo infantil y juvenil, como por ejemplo la educación y la salud.

A continuación se describen algunos ejemplos de programas que aparecen en la página web de Blueprints. Estos representan distintos tipos y niveles de intervención y todos ellos han sido implementados en Europa. Las descripciones ofrecen unas líneas generales sobre cómo cada programa cumple los estándares y dónde son administrados.

Nurse Family Partnership es un programa de visita domiciliaria en el que enfermeras realizan visitas domiciliarias a padres primerizos jóvenes, frecuentemente adolescentes, desde los primeros momentos de embarazo hasta que los niños cumplen los 24 meses. El programa aspira a promover la salud prenatal, mejorar el bienestar y desarrollo infantil a través de una mejor crianza, y promover la auto-suficiencia parental a través de la educación, el empleo, o la planificación de futuros embarazos. Enfermeras especialmente formadas rea-

lizan visitas semanales o quincenales estructuradas al domicilio de las familias. Las visitas domiciliarias permiten a las enfermeras preparar a los jóvenes para ser padres y guiarles hacia estilos de vida más saludables, a cuidar de sus bebés y a planear su futuro. Es clave para el programa que se establezca una sólida relación terapéutica entre la enfermera y la familia.

Sólidas evaluaciones científicas muestran que el programa NFP propicia una variedad de mejoras en la salud y el desarrollo infantil, como por ejemplo, un mejor comportamiento y logro académico, prácticas parentales más positivas, una reducción de los índices de maltrato infantil, y un aumento de la autonomía parental, incluyendo una reducción en el uso de prestaciones sociales (eg. Olds, Henderson, Chamberlin, y Tatelbaum, 1986; Olds y Kitzman, 1990). Estos impactos se mantienen durante un largo periodo de tiempo tras la finalización del programa, por ejemplo, los niños en FNP tienen menos probabilidades de estar involucrados en el sistema de justicia juvenil durante su adolescencia (eg. Eckenrode et al., 2010).

Nurse Family Partnership viene acompañado de un extenso material de apoyo, incluyendo manuales, formación y asistencia técnica. Ha sido implantado en el Reino Unido y en Holanda. Por cada dólar invertido en la versión estadounidense del programa para familias de bajos recursos se obtiene un retorno de 3.23 dólares (Aos et al., 2011)<sup>14</sup>.

El segundo programa descrito aquí, Incredible Years BASIC, está diseñado para padres de niños entre 2 y 10 años con problemas de conducta. Con este programa se espera mejorar la interacción familiar y prevenir la conducta antisocial temprana y persistente en estos niños.

El programa consiste en un curso de 12 semanas de sesiones de dos horas administradas por un grupo de aproximadamente 12 padres por cada dos líderes de grupo especialmente entrenados. Se les enseñan a los padres estrategias que les ayuden a gestionar los problemas de conducta de sus hijos, tales como la agresión, las rabietas y el mal comportamiento. También aprenden cómo fomentar las habilidades sociales de sus hijos a través de la regulación de emociones. Las sesiones consisten en discusiones grupales, ejercicios de modelado con cintas de vídeo y el ensayo de técnicas parentales.

Incredible Years BASIC ha sido evaluado con RCT en varios países, incluyendo EEUU, Reino Unido y Noruega. Estas evaluaciones prueban de forma consistente que el programa aumenta el uso de estrategias parentales positivas, reduce el uso de disciplina severa e incoherente, y reduce las conductas desviadas de los niños (eg. Webster-Stratton, Kolpacoff, y Hollinsworth, 1988; Hutchings et al., 2007; Larsson et al., 2009; Scott et al., 2010; Little, 2012; McGilloway et al., 2012).

Incredible Years BASIC se acompaña de completos manuales para el líder de grupo, DVD's, libros, CD's, folletos y actividades y lecturas recomendadas entre sesiones. Los líderes de grupo reciben durante tres días una formación inicial y soporte técnico y supervisión continuados para ayudarles a implementar el programa con éxito. El programa ha sido administrado en agencias de salud mental, centros de salud pública y centros escolares en EEUU, Reino Unido, Irlanda, Noruega, Alemania, Dinamarca, Holanda, Portugal y Suecia. Por cada dólar invertido el programa produce un retorno de 1.20 dólares (Aos et al., 2011)<sup>15</sup>.

El tercer programa, la Terapia Multisistémica (MST), es una intervención intensiva basada en la familia y dirigida a jóvenes infractores crónicos; normalmente, estos jóvenes han cometido delitos graves y muestran problemas de abuso de sustancias. El principal objetivo de la MST es reducir la conducta antisocial y la actividad delictiva, así como mejorar las habilidades parentales, las relaciones familiares, los resultados escolares y la implicación con iguales y actividades positivas. Un terapeuta trabaja con el adolescente en su entorno cotidiano, y con su familia, amigos, en la escuela y en su comunidad. Junto a la familia, el terapeuta diseña un plan de tratamiento para abordar los riesgos identificados y promover las influencias protectoras en el entorno del adolescente. Se emplean varias estrategias, tales como el CBT o el asesoramiento. El terapeuta se convierte en el único punto de contacto para la familia y se encuentra disponible 24 horas siete días a la semana. Habitualmente, una intervención MST dura entre 3 y 5 meses y se desarrolla en 3-6 sesiones semanales, cada una de ellas de dos horas de duración.

La MST ha probado su efectividad en múltiples RCTs de alta calidad. Reduce los índices de reincidencia y la conducta antisocial, incluyendo problemas de conducta y agresión, y también mejora el manejo de las emociones y la cohesión familiar (eg. Henggeler, Melton, y Smith, 1992; Timmons-Mitchell, Bender, Kishna, y Mitchell, 2006). Algunos efectos perduran y siguen siendo detectables incluso varios años después de la intervención (eg. Henggeler, Clingempeel, Brondino, y Pickrel, 2002).

Los “Servicios MST”, ubicados en los EEUU, proporcionan: formación, apoyo técnico, supervisión, diversos materiales (por ejemplo, manuales de tratamiento) y licencias de uso. Estos servicios son administrados por terapeutas experimentados que reciben cinco días de formación y apoyo continuado. La MST ha sido ampliamente implementada en los EEUU y en varios países europeos, incluyendo Noruega, España, Suecia, Dinamarca, Holanda, Islandia y Reino Unido. Un análisis de costo-beneficio revela que cada libra esterlina gastada en MST produce un retorno de 1.7 libras (Social Research Unit, 2012b)<sup>16</sup>.

## **Qué resultados han obtenido los programas según los estándares de evidencia**

Resulta instructivo describir brevemente qué resultados han obtenido los programas revisados según los estándares de evidencia. En este caso nos centramos en los 100 programas revisados según los estándares de evidencia en 2011 como parte del proyecto *Evidence2Success* de la Fundación Annie E. Casey. Estos representan una amplia selección de etapas evolutivas infantiles y resultados meta, y fueron deliberadamente seleccionados porque el grupo de expertos los consideraron los mejores disponibles.

La *Especificidad de la intervención* en general fue buena, y la mayoría de programas revisados cumplían todos los criterios “suficientemente buenos” en esta dimensión. La *Calidad de la Evaluación* se mostró mucho más variable. Los criterios “suficientemente buenos” de calidad de la evaluación que tendían a ser mejor abordados en las evaluaciones de programas incluían: la adecuación de las medidas de evaluación (refleja los resultados, no está conectado con la intervención, no está únicamente revisado por el implementador); una clara exposición de los datos demográficos, y la asignación de casos a los grupos del programa y los grupos de referencia en el nivel adecuado (aunque no siempre son analizados en el nivel correcto). Se observó menos claridad en las siguientes cuestiones: el tipo de tratamiento que recibía el grupo de referencia; hasta qué punto la intervención que fue realmente administrada es fiel a la intervención según fue diseñada; si o de qué manera se controla el clustering en los análisis, por ejemplo cuando la unidad de ubicación es el colegio; si existe equivalencia entre el grupo del programa y el de referencia al inicio del programa en cuanto a las mediciones de resultados; si el análisis fue con intención de tratamiento; y si se observó una deserción diferencial.

Con relación a los criterios “Óptimos”<sup>17</sup> de calidad de la evaluación, el seguimiento a los 12 meses tras el tratamiento fue ofrecido en menos de la mitad de los casos, y se trataba de un análisis de subgrupos. Raras veces se presentaba un análisis sobre la relación entre fidelidad y resultados o sobre el papel desempeñado por los factores facilitadores. También eran extremadamente raros los análisis de respuesta a la dosis (en sentido literal, en estos análisis se planifican variaciones en la dosis de intervención y se compara, por ejemplo, una versión larga del programa con una corta).

En cuanto al *impacto de la intervención*, alrededor de la mitad de programas indicaban tamaño del efecto. En cuanto a la *disponibilidad del sistema*, fue por lo general difícil establecer la disponibilidad del sistema al no ser posible contactar con el desarrollador del programa; la información suministrada en las páginas webs de los programas, por ejemplo, es inapropiada para este fin. Normalmente, el programa evaluado todavía seguía disponible, aunque esto puede resultar difícil de detectar ya que los programas se metamorfosean con el paso del tiempo, principalmente para realizar mejoras o para adaptarse a una población o entorno diferente. La mayoría de programas revisados ofrecían un manual y formación, pero la información sobre recursos económicos y humanos no siempre estaba tan fácilmente disponible. El nivel de diseminación y comprobación en el “mundo real” con frecuencia no quedaba claro, y a pesar de que muchos de los programas hacían gala de poseer un protocolo de fidelidad, resultaba menos claro si este era apropiado para entornos distintos a los de un estudio de investigación, en otras palabras, para entornos de servicio ortodoxos.

## **Las páginas web de Blueprints**

Hasta ahora, la labor de Blueprints ha sido diseminada a través de una página web gestionada por el equipo para la Prevención de Violencia de Blueprints en la Universidad de Boulder, Colorado. El nuevo Blueprints ten-

drá dos páginas web. La página principal está dirigida principalmente a una audiencia de proveedores de los EEUU. Está diseñada para permitir a los legisladores y proveedores un acceso rápido a la información que necesitan sobre todos los programas aprobados. El enfoque adoptado es similar al usado en la revista “Consumer Report” (en los EEUU) o “Which?” (Reino Unido). En lugar de buscar y comparar por tipos de cámaras o lavadoras, los usuarios de la página web pueden realizar búsquedas sobre programas adecuados por resultado, por grupo objetivo, y por factores de riesgo y de protección. No nos es posible ofrecer aquí un listado exhaustivo de la información suministrada sobre cada uno de los programas, pero las áreas principales incluyen:

- Objetivos del programa.
- Receptores del Programa.
- Nivel de intervención (por ejemplo, prevención universal, prevención selectiva, tratamiento).
- Entorno (por ejemplo, escuela, comunidad, hogar).
- Factores meta de riesgo y de protección.
- Una breve descripción del programa.
- Una breve descripción de los resultados logrados mediante el programa.
- Una breve descripción de la metodología empleada en los estudios de evaluación de relevancia.
- Información económica (por ejemplo, coste por unidad, relación costo-beneficio, estrategias potenciales de financiación).
- Información sobre la formación y asistencia técnica.
- Información de contacto tanto del diseñador del programa como del proveedor.

Los usuarios pueden imprimir una lista de datos que incluye esta información y también comparar distintos programas revisados según criterios idénticos.

Una página web hermana dirigida a los proveedores de servicios europeos ofrecerá las mismas funciones y contenido coherente, pero contendrá menos texto, incluirá nuevas áreas (por ejemplo, evidencia de diseminación en Europa del programa en cuestión) y será traducida a distintos idiomas europeos (inicialmente español, francés y alemán). Los usuarios que deseen mayor información podrán enlazar con la página web estadounidense.

Ambas páginas web son proyectos en desarrollo. Se prevé que a su debido tiempo se añadan nuevas áreas, incluyendo un modelo de lógica visual, contenido de vídeo (por ejemplo, el vídeo de un desarrollador de programa resumiendo el programa), una herramienta para explorar el impacto probable de implementar un *dossier* de programas por costes, beneficios y resultados, y un feedback subjetivo de legisladores, profesionales, niños y familias que tienen experiencia con los programas.

Para la creación de las páginas web se precisa generar contenido de alta calidad para cada programa aprobado. Parte de este contenido procede directamente de las revisiones realizadas, incluyendo resultados meta, grupo meta y modelos lógicos. Sin embargo, se precisa recoger datos adicionales. En primer lugar, se genera un indicador consistente y equiparable del tamaño del efecto utilizando los métodos meta-analíticos empleados en el Washington State Institute for Public Policy (Aos et al., 2011; Lee, Drake, Pennucci, Bjornstad, y Edovald, 2012). En el caso de programas y resultados que no han sido examinados todavía por el centro de Washington, estos precisan estudios de codificación mediante meta-análisis y la aplicación de la fórmula del tamaño del efecto utilizada por el Washington State Institute for Public Policy. En segundo lugar, los datos de financiación también se crean sobre el coste de implementación. Este coste es desglosado en costes iniciales, de materiales, de administración, de formación y de asistencia técnica.

También se enumeran las estrategias y fuentes potenciales de financiación del programa. Esta información se obtiene por medio de cuestionarios y entrevistas con los desarrolladores/proveedores del programa. En tercer lugar, el centro de Washington utiliza el tamaño del efecto y los datos de coste por unidad para generar un ratio coste-beneficio, que será posteriormente usado por Blueprints.

## Desarrollos futuros

El trabajo aquí descrito se desarrollará al menos en seis direcciones. En primer lugar, se espera que los estándares de Blueprints, que ya colocan el listón muy alto en el campo de los servicios de atención a la infan-

cia, vayan aumentando de nivel a medida que aumentan los conocimientos científicos sobre los distintos programas basados en la evidencia y su implantación y que la calidad de los estudios mejore como consecuencia de estándares como estos. Por ejemplo, el requisito de que se observe un impacto sostenido en un seguimiento a los 12 meses tras el programa podría pasar de ser un criterio “Modelo” a ser un criterio “Prometedor”, y se podría añadir el requisito de una réplica independiente.

En segundo lugar, se realizarán esfuerzos para analizar cómo los estándares y base de datos de Blueprints conectan con los utilizados por otros grupos. Por ejemplo, se desarrollarán discusiones periódicas con grupos como The Society for Prevention Research, y en Europa ya se ha comenzado a trabajar para establecer lazos con las comunidades de investigación e intervención de los países participantes (incluyendo aquellas que representan a los repositorios existentes). Dichas colaboraciones contribuirán a extender el uso de Blueprints y de otras fuentes. Por ejemplo, algunos repositorios, como MOVISIE de Holanda, priman la innovación, mientras que Blueprints se centra en programas probados. Ambos son importantes, y el trabajo del primero puede contribuir al segundo.

En tercer lugar, se aprobarán nuevos programas. Se deben contrastar más programas desarrollados en Europa con los estándares de evidencia. Se sabe de varios de ellos que tienen considerables posibilidades de cumplir con los estándares, pero deben ser revisados y, en caso de ser aprobados por el Consejo de Blueprints, diseminados más ampliamente (eg. Atria y Spiel, 2007; Faggiano et al., 2008; Salmivalli, Kärnä, y Poskiparta, 2011). Además, y desde un punto de vista a largo plazo, se está elaborando material que demuestra cómo emplear los estándares para ayudar a los profesionales a desarrollar programas desde el estadio de innovación hasta alcanzar el programa modelo. En este sentido, las innovaciones deberían ser contrastadas y probadas con el nivel de rigor científico adecuado a su etapa de gestación (Little, 2012a). Por ejemplo, los nuevos programas podrían justificar un estudio pre- y post-tratamiento o incluso un pequeño estudio con grupo de referencia, para luego progresar a RCTs más amplios en el caso de obtener resultados positivos.

En cuarto lugar, Blueprints necesitará promoción en Europa. Desde el primer instante, los estándares de evidencia deben ser compartidos con colegas de las organizaciones europeas más relevantes, incluyendo los centros pan-europeos, los repositorios y los institutos de investigación de cada país. Esto promoverá el debate sobre las diferencias y similitudes entre los distintos estándares/bancos de datos (por ejemplo, en base al tema, la función y la población meta) y las distintas formas de conectar las diversas iniciativas que permitan lograr una mayor sinergia. Finalmente, se puede fomentar en estados específicos y/o la Unión Europea la adopción de los estándares, la base de datos y el modelo económico. Se debería intentar reproducir en Europa las fuertes alianzas forjadas entre Blueprints con expertos y autoridades clave en los Estados Unidos. Así pues, se precisará de material escrito, así como la organización de conferencias y reuniones para llegar a entender las necesidades y preocupaciones de los potenciales usuarios y animarles a hacer uso de la página web. Se pueden desarrollar materiales de formación sobre los estándares y la promoción del uso de la página web entre los usuarios finales.

En quinto lugar, se precisa de investigación sobre la transportabilidad de programas de un contexto a otro; en concreto, de EEUU a Europa, o de un país europeo a otro. Como se señaló anteriormente, el éxito de los programas importados es un tanto irregular. Lo que resulta efectivo en un contexto podría no serlo en otro, y lo que es culturalmente apropiado en un contexto podría no serlo en otro. Varios factores podrían explicar esto, incluyendo el nivel de cuidado con el que los programas deben ser adaptados y son adaptados (Kumpfer et al., 2012) y el contexto político y de suministro de servicios en el nuevo lugar (Sundell et al., 2008). Pero se precisa más investigación, y los proveedores de servicios europeos precisan información acerca de si los programas importados ofrecen garantías de “ajustarse” o funcionar en su país y cómo se pueden aumentar las probabilidades de que esto ocurra. Además, los costes y beneficios de los programas podrían diferir de un país a otro, debido a las diferencias de sus sistemas de prestaciones sociales, lo cual refrenda la importancia de que se traslade el modelo económico de Washington, como ya se ha mencionado más arriba.

Finalmente, Blueprints se centra en programas, pero existe un interés cada vez mayor políticas de reconocimiento oficial y en identificar las prácticas efectivas (eg. Barth et al., 2012), teniendo en cuenta las dificultades y limitaciones que se dan con frecuencia durante la implementación de programas en los sistemas de

servicios de atención a la infancia (Little, 2010). Por este motivo, se desarrollarán y utilizarán nuevos estándares de evidencia para identificar las prácticas, políticas y procesos basados en la evidencia (por ejemplo, métodos de evaluación) que deberán ser recomendados para su amplia difusión.

## Referencias

- Aos, S., Lee, S., Drake, E., Pennucci, A., Klima, T., Miller, ... Burley, M. (2011). *Return on investment: Evidence-based options to improve statewide outcomes. Document No. 11-07-1201*. Olympia: Washington State Institute for Public Policy.
- Atria, M., & Spiel, C. (2007). The Viennese Social Competence (ViSC) training for students: Program and evaluation. In J. E. Zins, M. J. Elias, & C. A. Maher (Eds.), *Bullying, victimization and peer harassment: A handbook of prevention and intervention*. (pp. 179-198). New York, NY: Haworth Press.
- Axford, N., & Morpeth, L. (2012). The Common Language prevention operating system: from strategy development to implementation of evidence-based practice. In B. Kelly & D. Perkins (Eds.), *Handbook of Implementation Science for Educational Psychology*. Cambridge: Cambridge University Press.
- Barth, R. P., Lee, B. R., Lindsey, M. A., Collins, K. S., Strieder, F., Chorpita, B. F., ... Sparks, J. A. (2012). Evidence-based practice at a crossroads: The emergence of common elements and factors. *Research on Social Work Practice*, 22, 108-119.
- Bumbarger, B. K., & Perkins, D. F. (2008). After randomised trials: issues related to the dissemination of evidence-based interventions. *Journal of Children's Services*, 3, 55-64.
- Butler, S., Baruch, G., Hickey, N., & Fonagy, P. (2011). A Randomized Controlled Trial of Multisystemic Therapy and a statutory therapeutic intervention for young offenders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 50, 1220-1235.
- Eckenrode, J., Campa, M., Luckey, D. W., Henderson, C. R., Cole, R., Kitzman, H., ... Olds, D. (2010). Long-term effects of prenatal and infancy nurse home visitation on the life course of youths: 19-year follow-up of a randomized trial. *Archives of Pediatric Adolescent Medical*, 164, 9-15.
- Eisner, M. (2009). No effects in independent prevention trials: can we reject the cynics' hypothesis? *Journal of Experimental Criminology*, 5, 163-183.
- Elliott, D. S. (2010). Lessons from Columbine: effective school-based violence prevention strategies and programmes. *Journal of Children's Services* 4, 53-62.
- Faggiano, F., Galanti, M. R., Bohrn, K., Burkhart, G., Vigna-Taglianti, F., Cuomo, L., ... Wiborg, G., EU-Dap Study Group (2008). The effectiveness of a school-based substance abuse prevention program: EU-Dap cluster randomized controlled trial. *Preventive Medicine*, 47, 537-543.
- Flay, B. R., Biglan, A., Boruch, R. F., Castro, F. G., Gottfredson, D., Kellam, S., ... Peter, J. I. (2005). Standards of Evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science*, 6, 151-175.
- Grietens, H. (2010). Discerning European perspectives on evidence-based interventions for vulnerable children and their families. *International Journal of Child and Family Welfare*, 13, 6-17.
- Hawkins, J. D., & Catalano, R. F. (2002). *Communities that Care - Tools for Community Leaders: A Guidebook for Getting Started*. South Deerfield, MA: Channing-Bete.
- Henggeler, S. W., Melton, G. B., & Smith, L. A. (1992). Family preservation using multisystemic therapy: An effective alternative to incarcerating serious juvenile offenders. *Journal of Consulting and Clinical Psychology*, 6, 953-961.
- Henggeler, S. W., Clingempeel, W. G., Brondino, M. J., & Pickrel, S. G. (2002). Four-year follow-up of multisystemic therapy with substance-abusing and substance-dependent juvenile offenders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 41, 868-874.
- Hutchings, J., Bywater, T., Daley, D., Gardner, F., Jones, K., Eames, C., ... Whitaker, C. (2007). A pragmatic randomised control trial of a parenting intervention in sure start services for children at risk of developing conduct disorder. *British Medical Journal*, 334, 678-682.
- Larsson, B., Fossum, B., Clifford, G., Drugli, M., Handegard, B., & Morch, W. (2009). Treatment of oppositional defiant and conduct problems in young Norwegian children: results of a randomized controlled trial. *European Child & Adolescent Psychiatry*, 18, 42-52.

- Lee, S., Drake, E., Pennucci, A., Bjornstad, G., & Edovald, T. (2012). Economic evaluation of early childhood education in a policy context. *Journal of Children's Services*, 7, 53-63.
- Little, M. (2010). *Proof Positive*. London: Demos.
- Little, M. (2012). *From Innovation to Proven Model*. Dartington: Social Research Unit.
- McGilloway, S., Mhaille, G., Bywater, T., Furlong, M., Leckey, Y., Kelly, P., ... Donnelly, M. (2012). A parenting intervention for childhood behavioural problems: A randomized controlled trial in disadvantaged community-based settings. *Journal of Consulting and Clinical Psychology*, 80, 116-127.
- Olds, D. L., Henderson, C. R., Chamberlin, R., & Tatelbaum, R. (1986). Preventing child abuse and neglect: A randomized trial of nurse home visitation. *Pediatrics*, 78, 65-78.
- Olds, D. L., & Kitzman, H. (1990). Can home visitation improve the health of women and children at environmental risk? *Pediatrics*, 86, 108-116.
- Petrie, P., Boddy, J., Cameron, C., Wigfall, V., & Simon, A. (2006). *Working with children in care: European perspectives*. Maidenhead: Open University Press.
- Rowlands, J. (2010). Services are not enough: child well-being in a very unequal society. *Journal of Children's Services*, 5, 80-88.
- Salmivalli, C., Kärnä, A., & Poskiparta, E. (2011). Counteracting bullying in Finland: The KiVa program and its effects on different forms of being bullied. *International Journal of Behavioral Development*, 35, 405-411. doi:10.1177/0165025411407457
- Social Research Unit (2012a). *An introduction to evidence-based programmes in children's services*. Dartington: Social Research Unit.
- Social Research Unit (2012b). *Investing in Children: Youth Justice 1.1*. Dartington: SRU.
- Scott, S., O'Connor, T. G., Futh, A., Matias, C., Price, J., & Doolan, M. (2010). Impact of a parenting program in a high-risk, multi-ethnic community: The PALS trial. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 51, 1331-1341.
- Soydan, H., Mullen, E. J., Alexandra, L., Rehnman, J., & Li, Y. P. (2010). Evidence-based clearinghouses in social work. *Research on Social Work Practice*, 20, 690-700.
- Stemmler, M., Kötter, C., Bühler, A., Jaursch, S., Beelmann, A., & Lösel, F. (in press). Prevention of familial transmission of depression through a family-oriented programme targeting parenting as well as the child's social competence. *Journal of Children's Services*.
- Sundell, K., Hansson, K., Lofholm, C. A., Olsson, T., Gustle, L. H., & Kadesjo, C. (2008). The transportability of Multisystemic Therapy to Sweden: Short-Term results from a randomized trial of conduct-disordered youths. *Journal of Family Psychology*, 22, 550-560.
- Timmons-Mitchell, J., Bender, M., Kishna, M. A., & Mitchell, C. (2006). An independent effectiveness trial of Multisystemic Therapy with juvenile justice youth. *Journal of Clinical Child and Adolescent Psychology*, 35, 227-236.
- Webster-Stratton, C., Kolpacoff, M., & Hollinsworth, T. (1988). Self-administered videotape therapy for families with conduct-problem children: Comparison with two cost-effective treatments and a control group. *Journal of Consulting Clinical Psychology*, 56, 558-566.

## Notas

<sup>1</sup> Algunos ejemplos son SFI Campbell (Dinamarca), MOVISIE (Holanda), Ungsinn (Noruega), Metodguiden y SBU (Suecia), y Prevención basada en la evidencia (España).

<sup>2</sup> Por ejemplo, The Best Evidence Encyclopaedia (Reino Unido)

<sup>3</sup> Por ejemplo, el mejor portal de prácticas adecuadas del European Monitoring Centre for Drugs and Drug Addiction (EMCDDA), conocido como el European Exchange on Drug Demand Reduction Action (EDDRA).

<sup>4</sup> Por ejemplo, el Instrumento de Evaluación de Programas de Prácticas Parentales (Reino Unido) desarrollado por la National Academy of Parenting Research (NAPR)

<sup>5</sup> Los autores pueden suministrar una lista de estos si se precisa.

<sup>6</sup> <http://www.bestevidence.org/>

<sup>7</sup> <http://www.colorado.edu/cspv/blueprints/>

<sup>8</sup> <http://www.childtrends.org/links/>

<sup>9</sup> The Prevention Strategies Guide that was informed by the Communities that Care standard is no longer available. CTC specifically direct sites to Blueprints.

<sup>10</sup> <http://www.london.gov.uk/priorities/crime-community-safety/time-action/project-oracle>

<sup>11</sup> <http://www.consort-statement.org/>

<sup>12</sup> Esto quedó acordado como disposición provisional hasta que pudiera quedar determinado cómo afectaba a los índices de los programas educativos. Desde que tras el proceso de revisión quedó probado que los programas educativos *no* eran rechazados en función del requisito de significación estadística, el Consejo de Blueprints no aplica esta disposición.

<sup>13</sup> Los criterios “Suficientemente Bueno” en *Evidence2Sucess* son básicamente los criterios “Prometedores” en Blueprints (con algunas modificaciones).

<sup>14</sup> Cifras correspondientes al momento de su aplicación en el estado de Washington.

<sup>15</sup> Cifras correspondientes al momento de su administración en el estado de Washington.

<sup>16</sup> Cifras correspondientes al momento de su administración en Inglaterra y Gales.

<sup>17</sup> Los criterios “Óptimos” en *Evidence2Success* son básicamente los criterios “Modelo” (con algunas modificaciones).