

# Theory Driven Assessment

Ray Pawson & Sanjeev Sridharan

## *Abstract*

Este capítulo ofrece una introducción al uso de enfoques dirigidos por la teoría en la evaluación de programas de salud pública. Comienza con algunas ideas y principios, continúa con las principales estrategias de investigación y diseños asociados a este enfoque, y concluye indicando algunos beneficios notables y limitaciones tangibles. Posteriormente, se utiliza la investigación sobre una política de atención preventiva en Escocia (enfaticando la medicina preventiva y la atención preventiva comunitaria) para ilustrar algunos detalles de la aplicación del método. Este estudio de caso destaca el rol de la teoría como el competente pero falible “motor” del programa, mientras el diseño y el método de evaluación juegan el rol del modesto pero astuto “navegante”.

## **Introducción – la idea en pocas palabras**

Los programas de salud pública pueden ser conceptualizados de varias maneras. Para el auditor, son flujos de gastos. Para el político, son “buques insignia” a nivel ideológico. Para el administrador, son responsabilidades que precisan recursos humanos y gestión. Para el médico, son una forma de ganarse la vida. Para el público, son uno más de varios episodios cotidianos que claman por su atención. Aunque todas estas percepciones son justificables, este capítulo comienza desde un punto de vista distinto, uno que creemos esencial para la evaluación – concretamente, que los programas son *teorías*.

¿En qué sentido podemos considerar *teorías* a las intervenciones en salud pública? Para comprender la idea, es útil considerar el devenir de las intervenciones desde su nacimiento hasta su adultez. Éstas nacen en las mentes de los arquitectos de políticas, pasan a manos de los médicos y –es de esperar– a los corazones y mentes de los sujetos del programa. Ese viaje es una propuesta del tipo “*si es que... entonces...*” La idea preliminar, ambición, expectativas, hipótesis, o “teoría del programa”, es que *si* ciertos recursos (a veces materiales,

a veces sociales, a veces cognitivos) son proporcionados, *entonces* afectarán el razonamiento de las personas hasta el punto de producir cambios hacia un comportamiento más saludable.

Como todas las hipótesis, estas especulaciones resultan ser verdaderas o falsas (o más comúnmente, un poco de ambas). Estas eventualidades componen la lógica subyacente para la “evaluación dirigida por la teoría” (*theory-driven evaluation*). La investigación comienza revisando las teorías claves asumidas en la construcción de programas, y después pone a prueba su precisión y su alcance – se supone que el programa debiera funcionar de tal manera, pero ¿qué ocurre en la realidad? El trabajo empírico tiene la tarea de descubrir dónde se han justificado las expectativas previas y dónde éstas se han derrumbado – y por qué. Es más, estas investigaciones terminan con más teoría; es decir, una mejor comprensión de cómo funcionan realmente estas intervenciones. Si la evaluación avanza de acuerdo al plan, debería conducir al desarrollo de una teoría más refinada y discriminatoria, especificando *para quién* y *bajo qué circunstancias* y *en qué sentidos* funciona un programa. Las teorías (a diferencia de los programas y las poblaciones) son portátiles, y tienen un beneficio adicional: muchos los hallazgos de la investigación dirigida por la teoría son transferibles.

Para explicar esta abstracta visión en pocas palabras, y ver cómo puede ser adoptada como método de evaluación, consideremos la historia de una de nuestras teorías del programa favoritas. En la década de los ‘90, en el Reino Unido, un grupo de educadores de salud comenzaron a estudiar la idea de que los estilos de vida poco saludables de muchos adolescentes (en especial de las niñas) estaban siendo formados por la cultura de masas, y particularmente por los muy poco saludables pasatiempos del cine, las telenovelas y la música rock. Esto generó una teoría del programa que intentaba insertar “modelos de conducta” igualmente atractivos, pero decididamente más saludables (especialmente estrellas deportivas) en las páginas y los espacios de transmisión de los medios de comunicación juveniles. Los editores de estos espacios no necesitaron muchos incentivos, y así nació una valiente conjetura, conocida en la comunidad de la educación en salud como la “Teoría Dishy David Beckham”.

Y aquí hace su aparición en escena la metodología de evaluación. Una buena forma de poner a prueba el programa en cuestión es poner a prueba su teoría central. Quizás los lectores desearán revisar las detalladas investigaciones sobre este tema en Mitchell (1997), pero baste con decir aquí que la evidencia de hecho reveló un entusiasmo desenfrenado de las

adolescentes por observar las fotografías de Beckham y compañía en las revistas juveniles. Sin embargo, este pasatiempo continuó dando más ejercicio a las *mentes* que a los *cuerpos* de estas niñas. Pocos entrevistados informaron de un renovado interés en el deporte y el ejercicio. Así, somos testigos del fracaso de la “teoría del modelo de conducta juvenil”, según su implementación en este caso – aunque todavía queda la posibilidad de que pueda surtir efecto bajo otras circunstancias y en otros sentidos. La imagen de Beckham sigue siendo utilizada para vender perfume y ropa interior, pero su único aporte a la educación en salud fue marcarse un gran autogol. Y así, nuestros educadores de salud tuvieron que “volver a la pizarra” y reflexionar acerca de cómo diseñar una evaluación de la teoría del programa.

## **Diseñando una evaluación dirigida por la teoría**

En esta sección, ofreceremos una síntesis simple, de cinco pasos, del método. Los lectores observarán que el modelo sigue fastidiosamente el clásico ciclo de la investigación científica, concretamente: construcción de la hipótesis, recolección de datos, análisis de datos, y revisión de la teoría (Franklin & Ebdon, 2005). Durante este proceso, las investigaciones nunca son tan lineales ni ordenadas, y los evaluadores encontrarán que se deberán trasladar varias veces de un lado al otro antes de completar el ciclo.

### **1. Indagando y develando las teorías del programa subyacentes**

Las teorías del programa son fáciles de detectar. Revise Ud. cualquier discusión, debate, manifiesto, línea guía o lista de orientaciones en salud pública, y allí estarán, luchando por salir a la superficie. Su forma propositiva es fácilmente descifrable. Las teorías de salud pública tienden a seguir la siguiente fórmula:

- (A) Un intento de comprender los problemas sociales e institucionales que dan origen a un comportamiento poco saludable o a las desigualdades sanitarias.
- (B) Esto, seguido por conjeturas sobre los cambios que deben realizarse a estos sistemas para disminuir la mala salud, lo cual entonces conduce a...
- (C) Ideas para llevarle recursos frescos a los individuos y/o comunidades, en la esperanza de que logren cambiar y mejorar su comportamiento.

Seguramente los lectores serán capaces de recitar sus propias versiones de A, B y C, pero para ilustrarlas comenzaremos con un pequeño ejemplo:

Las muertes por incendios en hogares aún son significativas:

- (A) Porque las personas le dan poca prioridad al peligro de incendio, desconocen la utilidad de las alarmas para escapar oportunamente de los incendios, y encuentran que éstas son difíciles de instalar y mantener.
- (B) La solución están en vencer la flojera, aumentar el conocimiento de los peligros, estar dispuestos a gastar un poco más, y facilitar el reemplazo de las baterías.
- (C) Un programa que combine mayor información con una libre distribución e instalación de alarmas aumentará su uso y reducirá el peligro.

Por supuesto, el *contenido* de tales teorías será variable, dependiendo de si la cadena de conjeturas termina con “zanahorias” (como la oferta de alarmas), “garrotes” (por ejemplo, la prohibición de fumar) o “sermones” (por ejemplo, consejos sobre las dietas). Su *forma*, sin embargo, es bastante similar – conjeturas sobre el problema, lo que se puede o debe hacer al respecto, y cómo se propone hacerlo. Esta forma de pensar proporciona a la evaluación dirigida por la teoría su primera tarea concreta, que es indagar y develar las teorías del programa subyacentes.

Los creadores de una teoría son los más adecuados para ayudarnos a develarla, y esto puede involucrar:

- i) Una cuidadosa lectura y análisis de la documentación, orientación, regulaciones del programa, etc., sobre cómo el programa busca obtener sus metas.
- ii) Entrevistas con los arquitectos y/o administradores y/o médicos del programa, sobre cómo su intervención busca generar los cambios requeridos.

Las teorías del programa normalmente fluyen con bastante facilidad a partir de estas investigaciones, aunque debemos mencionar un par de dificultades relacionadas. La primera, ubicada en lo más alto del “palo encebado” político, es una tendencia a la ambigüedad –o incluso a la amnesia– del discurso en torno a las políticas (Shackley & Wynne, 1997; Pollitt,

2000). Estos autores observaron que las intervenciones prometidas para solucionar los problemas del mundo pueden producir mayores ganancias si se mantiene una deliberada imprecisión. Inicialmente, esto podría ayudar a reunir a más inocentes participantes en la mesa política. Y más adelante en el camino, una cierta ambigüedad o “capacidad de olvido” con respecto a los objetivos originales puede ser útil a la hora de proclamar sus éxitos.

El segundo problema, que está más relacionada con la ardua cotidianidad de la puesta en práctica de los programas, ocurre cuando la teoría central es, o aparentemente tan obvia, o tan tácitamente instalada en las mentes de los ejecutores del programa, que no logra “salir a flote” en la entrevista. En tales situaciones, se requiere una cierta capacidad persuasiva para incentivar a los médicos a clarificar de qué manera sus acciones buscan progresivamente afectar las decisiones de los participantes (Pawson & Tilley, 1997, Ch 6).

Más allá de estos obstáculos, develar una teoría es relativamente simple. De hecho, dado que entre la clase política no hay escasez de ideas en cuanto a nuevas reformas, también podríamos decir que suele ser *demasiado* simple. Lo que ocurre invariablemente en este tipo de ejercicio es que la develación de la teoría opera en plural, culminando en el descubrimiento de no *una*, sino *varias*, potenciales teorías del programa. Así, se requiere una segunda fase preparatoria antes de comenzar a investigar.

## **2. Mapeo y selección de las teorías a investigar**

La perspectiva de las teorías del programa tiene la inmensa ventaja de penetrar en la “caja negra” de una intervención. Rápidamente, nos lleva más allá del paradigma “*un problema – una cura*” para los programas sociales. Habiendo encontrado una forma de develar el funcionamiento de las teorías del programa, el próximo paso es comenzar a codificarlos, o “mapearlos”. Existe una confusa batería de técnicas para lograr esta tarea, conocidas con los nombres de: mapeo conceptual, modelo lógico, mapeo de sistemas, teoría de cambio, árbol de problemas y soluciones, construcción de escenarios, mapeo de configuración, y así sucesivamente (Trochim, 1989; Kaplan & Garrett, 2005; Snowden et al, 2008). Todas ellas tienen que ver con intentar presentar (casi siempre en forma de diagrama) el proceso a través del cual el programa busca lograr sus objetivos. Estos mapas pueden centrarse en la identificación: las variadas causas del problema, los pasos administrativos a seguir, el desarrollo de una secuencia de factores (*inputs*) y actividades del programa, los sucesivos

cambios en la disposición de los participantes, la focalización progresiva hacia los beneficiarios del programa, la secuencia de productos (*outputs*) y resultados de la intervención. No intentaremos aquí describir todos estos caminos; más allá de eso, perseguimos un tema fundamental y previo.

Nunca es una tarea fácil producir las listas, tablas o diagramas que se requieren como palancas generadores de cambios en un programa. Aún en nuestro ejemplo inicial de las alarmas de humo, los lectores habrán notado tres o cuatro ideas distintas en el debate sobre el problema y su resolución. Si se examina una intervención compleja, con múltiples agencias y objetivos, como las “Zonas de Acción Saludable” del Reino Unido, los investigadores se enfrentarán inmediatamente con una sobre-abundancia de teorías del programa (Judge & Bauld, 2001). Esta iniciativa en particular buscaba promover soluciones comunitarias a los problemas locales de salud. Así, muchos actores tuvieron la oportunidad de convertirse en teóricos del programa, y varias listas de actividades emergieron inmediatamente, haciendo imposible describir la intervención en términos de un modelo único.

La ejemplificación suprema del mapeo complejo es el famoso “diagrama de espagueti”, desarrollado para el proyecto “Previsión para Atajar a la Obesidad”. Los autores intentaron presentar un mapeo integral de la multiplicidad de factores que estaban contribuyendo a la epidemia de obesidad, como prelude al montaje de un “contra-ataque” igualmente integral. Detrás de la simple observación de que las medidas de nuestras cinturas están aumentando, hay una compleja red de factores, que van desde la psicología y fisiología individual, las actitudes hacia la actividad física, las normas sociales sobre la forma de los cuerpos, las culturas de consumo de alimentos, las economías de producción de alimentos, las estructuras del medio ambiente construido, etc. (Vandenbroek et al, 2007) En total, se describieron 108 factores – “los motores de la obesidad” – como líneas de influencia en el mapa<sup>1</sup>.

Los programas de salud pública buscan persuadir a los sujetos para cambiar, y el enfoque dirigido por la teoría acerca al evaluador a las ideas que lo logran. Confiamos en que hemos demostrado, sin embargo, que nunca hay una única teoría que comprobar. Siempre hay múltiples teorías incorporadas en una sola intervención, y éstas también irán cambiando a medida que el programa madura, aprende de otras experiencias, cambia sus ejecutores, sufre rivalidades, se fatiga, etc. Así, un paso inicial, esencial y –en nuestra opinión– subestimado en

---

<sup>1</sup> Varias versiones del “modelo de sistemas de obesidad” se encuentran disponibles en:  
<http://www.foresight.gov.uk/Obesity/12.pdf>  
[http://www.visualcomplexity.com/vc/project\\_details.cfm?index=622&id=622&domain=](http://www.visualcomplexity.com/vc/project_details.cfm?index=622&id=622&domain=)

la evaluación de las teorías del programa es escudriñar el mapa de ideas y seleccionar el subgrupo de teorías que se pretende poner a prueba.

La garantía de esta estrategia es nunca afirmar que se está poniendo a prueba *la* teoría del programa. Por el contrario, la lógica investigativa debiera ser que se investigará *un* subgrupo de las actuales teorías del programa relevantes. Y es en este punto que entra en juego una apuesta o decisión informada. Los investigadores finalmente deciden que han identificado algunas de las ideas dominantes que infunden un aspecto específico del programa, y que son dignas de evaluar por si mismas, sin suponer que están perfectamente ejemplificadas en una aplicación específica. Esta justificación es aún más pertinente considerando que el objetivo final del ejercicio es pulir la teoría del programa, más que emitir un veredicto sobre algún episodio en particular.

Entonces, ¿cómo concentra sus esfuerzos esta evaluación dirigida por la teoría? Existen diversos buenos consejos sobre este aspecto (Stame 2004: Pawson, 2006b), incluyendo el trabajo de Weiss (2000), apropiadamente titulado “*¿Cuáles eslabones de cuáles teorías debemos evaluar?*” No hay una sola fórmula para realizar esta apuesta; la siguiente lista ofrece sólo algunos indicadores preliminares de posibles prioridades:

- Podría ser que una intervención particular haya sido escrita junto con un marco legislativo claro – en el cual estas ordenanzas sobre quién-debe-hacer-qué-a-quién constituyen la teoría a ser comprobada.
- Las teorías del programa se imitan mucho entre sí; las mismas ideas surgen en distintos sectores de políticas – en este caso, es la “gran idea” la que debe ser clarificada y refinada a través de una comparación de su eficacia en distintos ámbitos.
- Podría existir un choque de ideas sobre cómo un programa específico debiera funcionar y cómo funciona concretamente – en este caso, podría estudiarse esta disputa acerca de la teoría o punto álgido del programa.
- Podría ser que el programa fluye por distintas manos – en este caso, se deberá abordar el tema de la consistencia de la teoría aplicada por distintos actores.

- Dada la cultura de “medición por resultados”, podría ser que el programa conlleve un conjunto de calendarios y metas – en este caso, éstas pueden orientar a las teorías que están siendo puestas a prueba.
- Podría ser que un creador o auspiciador de políticas tiene responsabilidades claves pero limitadas en el montaje de un programa – en este caso, una fuerte dosis de pragmatismo podría guiar la selección de la hipótesis a inspeccionar.

La evaluación dirigida por la teoría está guiada por una claridad conceptual y, sea cual sea la ruta, la selección del marco teórico para la investigación debe hacerse, y debe hacerse de manera clara.

### **3. Formalizando las teorías que se pondrán a prueba**

Este objetivo es apoyado por el paso final de lo que es todavía la etapa preparatoria de la investigación. Después de la selección y el mapeo de las teorías del programa, ha llegado el momento de formalizarlas. Deben ser traducidas a una forma propositiva, transformadas en una hipótesis apropiada para la investigación empírica. Las teorías del programa surgen como perspicacia, inspiración, ideas brillantes e intuiciones bien informadas. A veces demuestran ser poco más que pensamientos optimistas y “sueños de hojalata”. La investigación y la evaluación, por su parte, requiere de propuestas demostrables.

Una vez más, hay diversas maneras de organizar las teorías del programa como hipótesis de investigación empírica. Aquí presentaremos dos formas centrales, comenzando con una constatación de sus elementos comunes, antes de observar sus diferencias. Estos dos modelos se muestran en la Figura 1. El enfoque dirigido por la teoría intenta adentrarse en la “caja negra” de las intervenciones, alejándose de métodos de “caja negra” de investigarlas. En un experimento clásico, el programa se considera un tratamiento único (X) y la hipótesis única a investigar (Figura 1.1) trata de averiguar si este tratamiento conduce al resultado esperado (Y). Esta proposición se pone a prueba con una comparación experimental con un grupo de control, que examina la diferencia en los resultados en la presencia y ausencia del tratamiento. No nos referiremos a las críticas ya bastante conocidas a esta estrategia; simplemente recordaremos que aún cuando las diferencias en los resultados aparentemente confirman la validez del tratamiento, este método específico no nos logra decir nada acerca de *por qué* ha



funcionado el tratamiento. Está demás decir que la pregunta del “¿por qué?” es la razón de ser del enfoque dirigido por la teoría.

En la Figura 1, distinguimos entre las hipótesis de “teoría-de-cambio” (Connell et al, 1995) y las hipótesis “realistas” (Pawson & Tilley, 1997). Ambas pretenden investigar el proceso interno que conlleva un programa. Ambos enfoques nos recuerdan que los programas funcionan cuando logran persuadir a los sujetos (individuos o comunidades) para que cambien. Y los sujetos, desde el mismo comienzo, serán relativamente reacios o dispuestos al cambio. Los sujetos en el umbral de un programa ponderarán, esperarán, se preguntarán, investigarán, y cambiarán de opinión. Los sujetos que hayan sobrepasado el umbral “se tirarán el piquero”, entrarán en él de a poco, se saldrán, vacilarán, apoyarán, sabotearán, se lo tomarán, se harán los enfermos, harán proselitismo, y así sucesivamente. Las teorías del programa se proponen cambiar esta inercia, trasladando a una cantidad suficiente de personas del sector marginal y reacio hacia el cumplimiento y compromiso con los objetivos de la intervención.

Las hipótesis de “teoría-de-cambio” intentan describir este tipo de secuencias (Figura 1.2) como un conjunto de pasos a seguir, o “piedras de cruce”. La planificación del programa literalmente se traslada al papel, y entonces se convierte en la hipótesis a ser demostrada. Típicamente, esto puede implicar el mapeo de las decisiones tomadas con respecto a los *inputs* o recursos sobre los cuales se debiera construir el programa, y cómo éstos se traducen en actividades o *outputs* para los participantes del programa, y cómo éstos se incorporan a las acciones de los sujetos (primero a través de una comprensión inmediata de las ideas del programa, después en cumplimiento intermedio de sus actividades, y finalmente en cambios de comportamiento a largo plazo). La composición exacta de los “hitos” de esta secuencia será distinta de acuerdo al tipo de intervención y el conjunto de sub-teorías seleccionadas para ser investigadas. El elemento clave es que la posterior investigación interroga este proceso y llega así a una comprensión de por qué el programa ha funcionado según el plan, o por qué ha fallado.

## FIGURA 1 AQUÍ

Las hipótesis realistas (Figura 1.3) comienzan con la selección e identificación de un “mecanismo” clave del programa. Un mecanismo no es el programa, el nombre del programa,

ni la presencia o ausencia del programa. Un mecanismo describe lo que sucede *dentro* del programa para cambiar el comportamiento – lo que se ofrece y que podría persuadir a las personas a cambiar. Los mecanismos adoptan diversas formas en las intervenciones en salud pública: pueden ofrecer incentivos o regalos (esperando que las personas continúen utilizando lo regalado); pueden ofrecer instrucción (esperando que las personas aprendan las lecciones); pueden ofrecer consejos u orientación (esperando que las personas sigan los consejos); pueden imponer legislación y fiscalización (esperando que las personas la respeten).

Re-conceptualizar el ingrediente activo de un programa como una “oferta” (o incluso como un mandato) inmediatamente da origen a la posibilidad de que algunas personas rechacen la oferta (o desobedezcan el mandato). Las hipótesis realistas, por lo tanto, intentan diferenciar a los sujetos de los programas y sus circunstancias (normalmente abreviadas como “contextos”), respecto a cómo podrían responder ante el mecanismo del programa. La Figura 1.3 ilustra este tipo de situación, en la cual el mismo mecanismo conduce a resultados diferentes (designados por los distintos estilos de flecha) en tres contextos diferentes. Los programas siempre son focalizados, y las hipótesis realistas atienden a las teorías sobre por qué ciertos objetivos pueden o no ser adecuados. Para poner un ejemplo sencillo, consideremos una vez más para quién y bajo qué circunstancias una oferta de alarmas podría reducir las muertes por incendio. En el primer contexto, la alarma se ofrece a una casa que ya cuenta con este aparato – en este caso la oferta sería un desperdicio. En el segundo contexto, por ejemplo un hogar de pensionados, la alarma podría ofrecer protección sólo si se ofrece ayuda con su instalación y esos molestos cambios de batería. En el tercer contexto, una casa recién construida, esta protección sería perfecta para nuevos residentes con muchas otras cosas por hacer. Los programas se construyen con muchas más eventualidades en mente, y este tipo de hipótesis genera una oportunidad para ponerlas a prueba.

Por supuesto, es posible utilizar estos dos modos de hipótesis conjuntamente y en paralelo (Blamey & Mackenzie, 2007). Sin embargo, sea cual sea la fórmula utilizada, la evaluación dirigida por la teoría realiza un esfuerzo considerable para generar y organizar formalmente las conjeturas claves sobre por qué funcionan los programas.

#### **4/5. Recolección y Análisis de Datos**

Ahora nos centraremos en la recolección y análisis de datos para interrogar las teorías del programa escogidas. A pesar de que ambas son actividades fundamentales dentro del ciclo investigativo, aquí las trataremos juntas –y de hecho en orden inverso– para revelar mejor sus principios esenciales. Aunque pueda ser apropiado seguir la costumbre y referirse a esta etapa como “la puesta a prueba de la teoría”, preferimos designar a estos enfoques como “el refinamiento de la teoría”. El objetivo no es aceptar o rechazar teorías del programa; la misión es mejorarlas, una diferencia que se ilustra en la Figura 2.

## FIGURA 2 AQUÍ

Un análisis de teoría-de-cambio inspecciona las “piedras de cruce” en la cadena de implementación del programa, esperando que alguna piedra se vuelva inestable, que el espacio entre una y otra sea irregular, o que haya aguas peligrosas que cubren esta política. Para expresarlo de manera menos metafórica, la presuposición es que las teorías del programa siempre se lograrán de manera imperfecta. Se enfrentarán a ambivalencias o resistencias, generarán consecuencias no intencionadas, y otras prioridades y programas interferirán con ellas. El objetivo de este análisis, por lo tanto, es revisar la fidelidad de la cadena de implementación hipotética (Figura 1.2) para descubrir sus flujos, sus obstáculos y sus filtraciones (representadas en la Figura 2.1).

Este análisis puede desarrollarse de dos maneras. La primera es a través de un anexo explicativo en una evaluación sumativa. Si el objetivo de la evaluación es descubrir “lo que funciona”, entonces la atención se centra en los resultados finales de la secuencia de las teorías-de-cambio. Una comparación de los resultados esperados y concretos proporciona una cierta medida de éxito. Pero, igual de significativamente, la exploración empírica del camino anterior ofrecerá una explicación detallada de este éxito o fracaso relativo. Los programas exitosos habrán establecido una fuerte base de actividades, generado victorias rápidas, produciendo resultados intermedios, y así sucesivamente; por otro lado, los programas menos exitosos habrán enfrentado un camino con mayores obstáculos.

El otro uso del análisis de las teorías-de-cambio es en la planificación de programas o la evaluación formativa. En este caso, el mapa lógico proporciona el diagrama arquitectónico para la construcción del programa. En vez de construir un programa con la ambición amplia de que genere el cambio de comportamiento requerido, los mapas de teorías-de-cambio

exigen una justificación para cada paso y decisión en ese camino. Una planificación de este tipo debiera considerar la viabilidad, verosimilitud y demostrabilidad de cada componente de la teoría. Esta última característica es particularmente importante, ya que permite algo de “ensayo y error” en la construcción del programa, corrigiéndolo progresivamente a medida que van surgiendo las imponentes filtraciones, emergentes obstáculos y consecuencias no intencionadas, según se muestra en la Figura 2.1.

La atención ahora se vuelve hacia el análisis realista y la Figura 2.2. El objetivo es el mismo –refinar la teoría del programa inicial (Figura 2.3). En este caso, la atención se centra en aprender más sobre el alcance y la focalización de una intervención. El análisis realista comienza con la idea de que la oferta del programa tendrá un atractivo variable. Profundizar la comprensión del programa implica un entendimiento más detallado de ese atractivo. Ciertos subgrupos habrán sido hipotetizados como grupos objetivos óptimos en la teoría del programa original, y el objetivo del análisis es revisar el alcance de los hallazgos. Algunas expectativas serán confirmadas, algunas estarán equivocadas, y el producto final del análisis será una comprensión mejorada de la eficacia o ineficacia del programa (crudamente representada aquí por la configuración del resultado quintuple revisado del diagrama).

Es importante mencionar que este proceso de especificación contextual se aplica a los individuos, sus historias y características, pero también a sus instituciones, localidades y comunidades. El “mantra” de la evaluación realista es aprender “para quién, bajo qué circunstancias, con qué duración, y en qué sentidos funciona un programa” – y el análisis de contexto puede ser sensible a todas estas dimensiones. El otro objetivo realista, por supuesto, es comprender “por qué” el programa funciona, y esto implica una profunda comprensión del mecanismo del programa y su utilidad y atractivo exactos en las distintas etapas de una intervención. Los programas de salud pública no funcionan pasivamente; sólo funcionan con el consentimiento de los participantes. Descubrir por qué funciona un programa dependerá de nuestra capacidad de explicar por qué éstos tomaron sus decisiones.

Habiendo demostrado los objetivos del análisis de datos, ahora retrocederemos para ver qué forma asumirán los datos y cómo se obtienen. Por supuesto, gran parte de esto dependerá de la preocupación política, los detalles de la intervención, y la teoría del programa concreta. Sin embargo, se pueden establecer algunos principios básicos, de los cuáles la más fundamental es que el análisis dirigido por la teoría exige una base de evidencias de múltiples métodos.

Una breve mirada a las Figuras 1 y 2 nos permite comprender la necesidad de estos métodos mixtos. El mapa de teorías-de-cambio está basado en una simple e indiscutible verdad sociológica: las acciones traen consecuencias. Se describe el desarrollo de un programa en que un conjunto específico de actividades de intervención conduce a ciertos resultados; otro proceso generará otros resultados, y así sucesivamente. El diagrama realista contiene otra idea indiscutible: el mismo proceso no siempre conduce al mismo resultado. Un programa que funciona bien en un contexto no necesariamente será exitoso en otro.

Aquí, entonces, tenemos los ingredientes básicos y activos del cambio social. ¿Cómo se pueden detectar? Bueno, en términos amplios, los datos del proceso se generan utilizando métodos *cualitativos*; los productos y resultados se miden con enfoques *cuantitativos*; y la información de contexto requiere una observación y medición *comparativa*. La comprobación de cualquier teoría del programa requiere la acción conjunta o *triangulación* de estos tres elementos.

Detrás de estas estrategias gruesas de recolección de datos, hay un gran arsenal de técnicas potencialmente útiles. La investigación sobre los procesos de toma de decisiones depende de la observación, participación, entrevistas, auto-informes, grupos focales, análisis de documentos, etc. La medición de los resultados podría depender de la comprobación de actitudes, escalas, resultados de comportamiento, registros administrativos, etc. La información comparativa puede recopilarse a través de cualquiera de estas técnicas, y puede estar basada en estudios de caso de subgrupos y sub-secciones particulares de una intervención, a partir de sus distintas pruebas. Tal vez el elemento más significativo dentro de esta gama de métodos es que los datos no sólo se pueden obtener “en vivo” o “*in situ*”, sino que también de manera retrospectiva, sintetizando la evidencia de evaluaciones pre-existentes. Podemos consultar a Pawson (2006a) para algunos ejemplos de cómo se puede llevar a cabo una revisión sistemática, tanto en el modo de teorías-de-cambio como en el modo realista. Lo que conecta todos los enfoques, y lo que sigue siendo el principio fundamental de la recolección de evidencias, es que la selección de las técnicas de recolección de datos debiera ser dirigida por la teoría, y no condicionada por sus métodos.

## De la Teoría a la Práctica: “Ten Corazón, Paisley”

### El Programa

Nuestro ejemplo práctico de un método dirigido por la teoría se extrae de una reciente evaluación de un proyecto nacional escocés llamado “Ten Corazón, Paisley” (*Have a Heart Paisley*, o HaHP). De acuerdo a la reciente política sanitaria de Escocia, HaHP promovió un giro hacia un servicio de salud con un énfasis en la medicina preventiva (Scottish Executive, 2005a y 2005b). Los objetivos inmediatos fueron ofrecer un programa focalizado para la población en edad laboral, adultos de entre 45 y 60 años en riesgo de contraer enfermedades cardiovasculares, y aquellos de cualquier edad con enfermedades cardiovasculares, en la localidad de Paisley (Sloan et al, 2005). Se reconoció que se requeriría un esfuerzo adicional para atraer a aquellos individuos que típicamente tenían menos probabilidad de entrar al servicio de salud, especialmente aquellos que vivían en una de las áreas más empobrecidas de Escocia. El objetivo a largo plazo de HaHP era reducir la carga total y niveles de desigualdad en las enfermedades cardiovasculares coronarias (*CHD*) en el pueblo de Paisley, a través de programa integrado de prevención secundaria y primaria (Blamey et al, 2004). HaHP fue diseñado como un “laboratorio para la acción nacional”, que significaría “un recurso de aprendizaje para el resto de Escocia”. Se esperaba que los éxitos locales inspiraran una transformación nacional.

Las intervenciones de planificación e implementación para enfrentar problemas de desigualdades sanitarias en comunidades pobres presenta múltiples desafíos conceptuales y operacionales – desde un comienzo, implica teorías falibles. Por ejemplo, incluso las “comunidades empobrecidas” son heterogéneas; tanto indigentes como personas relativamente afortunadas pueden vivir en una misma localidad de este tipo. Es posible que una iniciativa diseñada para abordar el problema de la desigualdad pueda terminar exacerbándola; un mecanismo a través del cual esto podría ocurrir sería cuando son los “relativamente afortunados” de las comunidades pobres los que muestran una mayor tendencia a participar de la intervención. Y podrían existir otras razones de contexto por las cuales un programa diseñado para abordar las desigualdades sanitarias no funcione de acuerdo a lo esperado. Más arriba, describimos los programas como “ofertas”, y puede que la oferta de participar en una intervención en enfermedades coronarias no sea percibida con el mismo nivel de interés por todos los beneficiarios del programa: los obstáculos pueden incluir rutinas

rígidas de salud marcadas por factores familiares y de género; constructos psicológicos tales como la falta de confianza; escepticismo político frente al rol del gobierno en la solución de los problemas de salud; y limitaciones de acceso debido a los impedimentos de algunos beneficiarios.

En este contexto, deben tomarse decisiones concretas sobre cómo implementar el programa. La Tabla 1 describe las etapas claves de la dimensión de prevención primaria de HaHP. A medida que el programa se fue desarrollando, se realizaron actividades en cuatro áreas principales: contacto, exámenes médicos, educación en salud y micro-intervenciones. Recordemos nuestra discusión anterior sobre la complejidad interna de los programas de salud pública, y observemos la cantidad de sub-procesos en cada etapa.

TABLA 1 AQUÍ

### **Diseño y Entrega de la Evaluación**

Nuestro proceso ahora sigue el curso de la investigación real (Sridharan et al, 2008), comenzando con lo que sólo podemos describir como “la lucha por indagar, sacar a la superficie y formalizar la teoría del programa del HaHP”. El equipo de evaluación trabajó estrechamente con el equipo del programa y los planificadores de políticas para desarrollar la teoría del programa del HaHP. Cabe mencionar, antes que nada, el calce con lo que describimos anteriormente como el “ABC preliminar” (¿cuál es el problema?, ¿qué hacemos al respecto?, ¿cómo lo hacemos?) de la lógica de problema. En este caso, tenemos que:

- A. Individuos en múltiples situaciones de vulnerabilidad muchas veces otorgan una baja prioridad a su propia salud y no utilizan los recursos disponibles (por ejemplo, el Sistema Nacional de Salud) para monitorear y prevenir enfermedades y condiciones crónicas.
- B. La solución es proporcionar un complejo conjunto de servicios, incluyendo exámenes médicos, educación en salud, e intervenciones adicionales para informar y empoderar a los individuos que viven en situaciones de pobreza y exclusión.
- C. Un programa que combine la información de los exámenes con el apoyo de un monitor de salud, así como información y servicios adicionales de otras

intervenciones, generará mejores estilos de vida y menor riesgo de enfermedades coronarias en la población objetiva. Esto, a su vez, puede reducir las desigualdades sanitarias.

La clarificación de la lógica de programa se realiza por etapas, y la próxima fue traducir la lógica de programa en un mapa de teoría-de-cambio, según muestra la Figura 3. Nótese que la selección de “cajas negras” en esta Figura ilustra el hecho de que no solamente el programa mismo es una caja negra, sino que también pueden serlo los componentes de la intervención (por ejemplo: alcance, exámenes o educación). Un importante objetivo de la evaluación dirigida por la teoría es “desempacar” cada una de las “cajas negras”, una a la vez.

### FIGURA 3 AQUÍ

Comencemos con la aparentemente simple tarea de contactar y llegar a la comunidad. En la práctica, esto depende de una hipótesis algo más incierta sobre la eficacia de las cartas enviadas a través del correo. ¿No existe el riesgo de que éste sea un instrumento demasiado burdo para llegar a las personas del sector más vulnerable? ¿El mensaje caerá al tacho de la basura junto con el resto del “correo no deseado”? ¿Cómo responden en verdad las personas a las comunicaciones escritas? ¿Estará formulado apropiadamente el mensaje? ¿Deberían combinarse estas cartas escritas con otros métodos de promoción? ¿Hay alguna posibilidad de que los “bien preocupados por la salud” o los “pobres pero saludables” tengan mayores posibilidades de responder a la invitación a participar?

En la etapa de los exámenes médicos, los individuos fueron informados de su riesgo de contraer enfermedades cardiovasculares coronarias (CHD) en los próximos 10 años (se calculó su puntaje con el *Framingham Risk Score*), y recibieron su “estatus de riesgo” (bajo, moderado o alto). Una vez más, no hay nada simple ni automático en un proceso como éste. Claramente, la capacidad del examen de gatillar cambios en los individuos depende fundamentalmente de la capacidad de los examinadores de comunicar efectivamente el riesgo al individuo, y de explicar los cambios básicos que éstos podrían realizar para disminuir este riesgo. Además, este método de examinación depende completamente de la validez del examen Framingham. Investigaciones recientes han sugerido que el estatus socioeconómico puede ser un factor de riesgo independiente para las CHD (Hippisley-Cox et al., 2007;



Lyratzopoulos et al., 2006; Tunstall-Pedoe & Woodward, 2006), pero el estatus socioeconómico no está incluido en la escala de Framingham.

La educación en salud fue otro aspecto clave de la intervención HaHP. Las presuposiciones básicas sobre cómo esta educación en salud sería de ayuda no estaban claras al comienzo de la intervención. ¿Cómo se modificó el marco de educación en salud para centrarse en los individuos pobres? ¿Cómo exactamente ayudó esta educación a empoderar a los individuos? La relación entre monitor y usuario es importante para generar cambios; sin embargo, también puede conducir a una dependencia del monitor, en términos del apoyo permanente (lo cual es poco práctico como solución de largo plazo). Esta relación también podría interferir con las relaciones de los usuarios con sus médicos, ya que las relaciones con los monitores son más personales, mientras que la interacción con el médico puede ser más clínica. Si los individuos acuden preferentemente a sus monitores y no sienten la necesidad de discutir los temas con sus médicos, esto podría interferir con la capacidad de estos últimos de proporcionar una atención de salud efectiva.

Hay una pequeña “selva” de temas en los tres párrafos anteriores, pero incluso éstos representan sólo un pequeño sub-conjunto de las cuestiones que pueden surgir cuando se somete la lógica causal de un programa a una inspección microscópica. Como se explicó en la primera parte del documento, siempre se puede escarbar más profundamente en las teorías del programa. Dada la centralidad del “contacto” para enfrentar el problema de las desigualdades sanitarias, volveremos a la primera etapa para describir en mayor detalle la teoría que estamos trabajando. Un factor clave para esta intervención fue que los beneficiarios objetivos se decidieran a llamar por teléfono al número proporcionado para presentarse al examen. Los siguientes eventos debieron sucederse antes de que esta llamada se pudiera ocurrir:

Se envía por correo una carta de invitación → la carta es entregada (una cantidad de cartas, especialmente en las comunidades más pobres, fueron devueltas sin ser entregadas) → el receptor abre la carta → y la lee → toma la decisión de que el programa HaHP podría resultar beneficioso, a pesar del esfuerzo que implica → toma la decisión de hacer la llamada telefónica → realiza efectivamente la llamada telefónica.

Todas estas etapas podrían haber conducido a un resultado diferente. Los individuos más pobres podrían tener menos posibilidades de recibir la carta de invitación, debido a un inadecuado registro con los servicios públicos y censos, mayores tasas de transición entre

hogares, períodos de indigencia, etc. Podrían tener menos posibilidades de leer la carta y comprender el material, considerando las diferencias en la tasa de alfabetización o su familiaridad con los contenidos del mensaje. Finalmente, los individuos más pobres podrían carecer de los medios, experiencia y confianza para contactar al centro de llamados.

Todas estas “micro-teorías” (y más) emergieron en la discusión con el equipo del programa. Recordemos que en nuestra discusión anterior, establecimos que las teorías del programa también tienen consideraciones políticas más globales. Nuestro campo de acción es la salud pública, y podría decirse que el éxito del programa en términos de llegar a la población objetiva depende de dos grupos de teorías opuestas y bien establecidas. La teoría “optimista” considera a la salud pública como una “vía de salvación” para los pobres. En el actual contexto, los individuos más pobres podrían considerar la comunicación inicial como un reconocimiento de sus necesidades más significativas, y así presentarse voluntariamente a mayor escala para participar de la intervención. La idea de fondo es que los individuos desaventajados están muy concientes de sus permanentes problemas de salud, y considerarían esta intervención como una rara oportunidad de abordar estos problemas. Siguiendo esta lógica, las primeras cartas se enviaron directamente a los sectores geográficos más empobrecidos. El plan inicial del programa fue que los individuos más pobres (que vivían en los sectores más pobres) se auto-seleccionarían para la intervención y participarían del programa a mayor escala (cabe mencionar que no existía un mecanismo formal del programa para entregar los servicios basados en necesidad socioeconómica).

Una segunda perspectiva teórica es más bien pesimista. En términos generales, el argumento sería que, sin un mecanismo formal para restringir este servicio a los más pobres, los “bien preocupados por la salud” y los “mas afortunados” tenderían a registrarse voluntariamente para estos programas de salud a mayor escala que los más “desaventajados”. Estas personas no sólo tendrían una mayor comprensión de sus necesidades de salud, sino que tienden a tomar decisiones más inteligentes frente a una oportunidad “gratuita” de mejorar su salud. Otras fuerzas se suman a este escenario menos optimista. La literatura sobre desigualdades sanitarias informa extensamente sobre las personas “difíciles-de-contactar” y los fracasos de muchos programas en atraer a los indigentes, analfabetos o aquellos que están en situaciones transitorias o inestables. Por ejemplo, en Velupillai et al. (2008: 134) se descubrió un patrón repetitivo, en el cual los “bien preocupados por la salud” y los “pobres pero saludables” tendían a responder más a este tipo de llamado “voluntario”.

Es de esperar que lo anterior haya dado al lector una idea de cómo algunas de las presuposiciones del programa HaHP fueron indagados, formalizados y priorizados. Ahora, observaremos el trabajo con los datos, indicando cómo se fueron poniendo a prueba las teorías. Se interrogaron decenas de conjeturas, y para los propósitos de esta exposición nos remitiremos a los temas que tienen que ver con el contacto inicial del programa. La investigación se centró en las siguientes tres preguntas:

- ¿Cómo evaluó el equipo del programa la capacidad del programa de abordar el problema de las desigualdades sanitarias?
- ¿Cómo reaccionaron los beneficiarios esperados a los esfuerzos de contacto de la intervención de prevención primaria? Este tema requería entrevistar a los individuos que respondieron a la invitación de participar en la intervención, así como a aquellos que decidieron no participar. Sin embargo, este aspecto del estudio fue limitado, porque sólo se consiguió “permiso ético” para hablar con los individuos que estaban dispuestos a participar de la intervención.
- ¿Fue exitoso el programa en llegar a los individuos con múltiples desventajas (por ejemplo, los individuos más pobres que vivían en los sectores más pobres)? Los datos administrativos obtenidos sobre las características de los individuos (dónde vivían, su estatus socioeconómico según el sistema tributario, edad, género) en las múltiples etapas de la intervención (ver Figura 2) nos permitieron examinar si el programa tuvo éxito en llegar a aquellos con múltiples desventajas.

La Tabla 2 resume las preguntas de investigación, los métodos de obtención de datos, y el análisis posterior. Nótese el marcado enfoque de “métodos mixtos” de este enfoque dirigido por la teoría.

#### TABLA 2 AQUÍ

Las limitaciones de espacio no permiten un análisis más profundo de los resultados, pero ofrecemos aquí un breve resumen de éstos. Los hallazgos están organizados por orden de pregunta en la Tabla 2.

*¿Tuvo esta intervención el peso suficiente para lograr un impacto en las desigualdades sanitarias?*

El equipo del programa sintió que la estrategia de contacto había fallado en gran medida, no logrando atraer a la población objetiva “difícil-de-contactar”, y resultando en una sobre-representación de los “bien preocupados por la salud”. Hubo evidencias claves, como el hecho de que muchos usuarios declararon estar concientes de la necesidad de realizar cambios *antes* de recibir la carta de invitación, y así fueron motivados a participar por aquella conciencia pre-existente.

Una gran cantidad de médicos proporcionaron retroalimentación, lo cual ayudó a desarrollar un método más matizado de la focalización. Algunas ideas claves fueron la necesidad de un enfoque focalizado que tuviera múltiples escalas y niveles, y también la necesidad de poner especial atención al marco de muestreo y las provocaciones específicas requeridas para lograr un impacto en la población “difíciles-de-contactar”.

*¿Qué aspecto de la intervención la hizo atractiva para que los individuos de la población objetiva participaran en ella?*

Los análisis de los beneficiarios esperados fueron limitados, porque sólo se concedió permiso ético para hablar con los individuos que se habían decidido a participar del estudio. No pudimos conocer las circunstancias individuales de vida, las limitaciones, y los puntos de vista de aquellos que no respondieron a la invitación, o que no participaron de la intervención. Esto limita nuestra capacidad de contestar la pregunta, pero de todas maneras podemos aprender más sobre las razones por las cuales algunos individuos sí participaron de la intervención.

Los diez participantes entrevistados expresaron algún nivel de disposición al cambio. Muchos usuarios estaban concientes de los temas de salud antes de su participación, declarando haber respondido a la invitación porque querían estar más saludables, perder peso o dejar de fumar. Otras razones incluían una percepción de su propio proceso de envejecimiento y el consiguiente deterioro de la salud, o de una historia familiar de enfermedad cardiovascular.

La invitación personal pareció ser importante para los usuarios, y algunos reconocieron que hubiese sido improbable que enfrentaran sus necesidades de salud sin aquella invitación. Dos usuarios informaron que habían visto afiches en su centro de salud local, anunciando estas intervenciones, y que las habían considerado, pero no tomaron acción concreta hasta que les llegó la invitación directa. Su participación también estaba ligada a la idea que “alguien demostró interés”; esto fue así para la invitación al examen médico, pero aún más para el apoyo permanente de los monitores en salud.

*Describe la relación entre pobreza del sector, el estatus socioeconómico y la participación en la intervención.*

Se desarrollaron modelos de múltiples niveles para evaluar el impacto de medición de la pobreza del sector y de los individuos, a partir de su participación en la intervención. Si el programa hubiese sido exitoso en llegar a los “difíciles-de-contactar”, entonces se esperaría que los individuos de grupos socioeconómicos más bajos y de sectores más pobres participaran más de la intervención. Este patrón no se logró, y emergió un hallazgo algo paradójico.

La medición del nivel de pobreza por sectores no logró predecir quiénes participarían en la intervención. Después de controlar por diferencias en los niveles individuales de pobreza, los individuos que vivían en sectores más pobres no necesariamente participaron más en la intervención que los que vivían en sectores más “afortunados”. Sin embargo, la medición de los niveles individuales de pobreza sí logró predecir la participación en la intervención: los individuos más pobres tenían sustancialmente menos probabilidades de participar en el programa.

## **Implicancias y conclusiones**

En estos comentarios finales, consideramos las implicancias y la utilización de los hallazgos presentados aquí, utilizándolos para reflejar una de las afirmaciones más audaces del método, cual es que “no hay nada más práctico que una buena teoría” (Weiss, 1995: Pawson, 2003). La evaluación dirigida por la teoría no busca llegar a veredictos sobre las intervenciones.

Busca explicar su éxito o su fracaso, siendo ambas tremendamente relevantes para el desarrollo de los programas.

Como sucede a menudo en el afiebrado campo de la generación de políticas, alrededor de un año después del comienzo de la evaluación del HaHP, otro conjunto de intervenciones, llamado “Mantente Bien”, comenzó a ser aplicado en diversas comunidades de Escocia. Desde entonces, “Mantente Bien” se ha convertido en la plataforma central de las políticas preventivas escocesas. Aunque las dos intervenciones tienen importantes diferencias, comparten un enfoque en la medicina preventiva y el impacto sobre las desigualdades sanitarias. Así, las lecciones sobre el contacto inicial, sintetizadas más arriba, fueron capaces de nutrir los esfuerzos de planificación de “Mantente Bien”.

El desafío para las intervenciones que buscan abordar el problema de las desigualdades sanitarias es aprender específicamente cómo identificar y contactar a los individuos que se encuentran entre los más desaventajados a nivel individual y comunitario. La enorme ventaja de un enfoque dirigido por la teoría para el problema de las desigualdades sanitarias reside en la atención que otorga al tema de *quién* participa de la intervención. Evaluar la efectividad de intervenciones complejas centradas en las desigualdades sanitarias, sin prestar atención a quién se está tratando, no sería de mucha utilidad.

En este sentido, los hallazgos “negativos” del HaHP encontraron un uso inmediato. Un programa con un amplio enfoque en “sectores pobres” puede ser sólo el comienzo – otros enfoques más precisos serán necesarios para llegar efectivamente a los pobres de sectores pobres. Puede ser que los individuos que presentan las mayores necesidades, como aquellos con múltiples desventajas o condiciones co-mórbidas, no sean identificables a través de marcos de muestreo estandarizados (como las bases de datos centralizados). En este sentido, la investigación sobre HaHP debe conjugarse con la rica y variada literatura que comienza a emerger sobre las técnicas de muestreo que se requieren para llegar a la población “difícil-de-contactar” (Heckathorn, 1997, 2002; Salganik & Heckathorn, 2004).

Algunos de los hallazgos más “positivos” de la evaluación de HaHP también pudieron ser utilizados inmediatamente. Las entrevistas con el equipo de intervención revelaron la necesidad de enfoques “hechos a la medida” para llegar a los más pobres, enfatizando el uso de la invitación personal, seguido por un contacto personal. Estas lecciones sobre las

dinámicas del contacto fueron ampliamente difundidas a través de una serie de informes sobre cuidado preventivo, orientados hacia la comunidad de políticas sanitarias y atención médica, y en una presentación en el encuentro sobre Políticas de Cuidado Preventivo, organizado por el Departamento de Salud del gobierno escocés. Un aspecto notable de estos esfuerzos de difusión fue que los resultados se compartieron mientras la evaluación de “Ten Corazón, Paisley” aún estaba en curso, y los datos y resultados aún estaban siendo organizados.

Finalmente, para establecer conclusiones más generales sobre los métodos, quisiéramos terminar destacando una de las verdaderas fortalezas del enfoque dirigido por la teoría, en particular su visión dinámica del aprendizaje. Los programas se desarrollan a través de una compleja cadena de pasos; de manera similar, un sistema de aprendizaje en evaluación debe seguir la cadena de causas de manera dinámica.

Uno de los aspectos más interesantes de la literatura externa de validez general es lo poco que esta literatura ha abordado los desafíos de la generalización de las evaluaciones de intervenciones complejas. Muchas veces, el tema del aprendizaje en las evaluaciones se centra casi exclusivamente en los resultados de la evaluación de impacto—normalmente obtenidos al final del estudio, lo cual puede demorar hasta cinco años. En la realidad, dada la complejidad de muchas intervenciones, es necesario un cuerpo de aprendizajes mucho más rico, que pueda destacar la complejidad de la planificación, la implementación y los caminos por los cuales deben transitar las intervenciones. Según Sanderson, algunas otras funciones de las evaluaciones pueden ser: “influir la conceptualización de temas y el rango de opciones consideradas, y desafiar las presuposiciones que establecen cuáles con los objetivos y las actividades apropiadas” (Sanderson, 2003: 333). La evaluación dirigida por la teoría puede ayudar a crear una “hoja de ruta” que muchas veces no está presente al comenzar una intervención.

Las evaluaciones dirigidas por la teoría también nos ayudan a formular preguntas sobre los incentivos que deben proporcionar los programas para lograr sus objetivos de largo plazo – en este caso, llegar a individuos con múltiples desventajas. En una cultura centrada en los resultados (“se deben atender a 500 individuos este mes”), hay fuertes incentivos para tratar a los que se preocupan de su salud y a los pobres saludables, porque en general éstos son más fáciles de contactar y atraer. Como expusimos anteriormente, el enfoque dirigido por la teoría

nos ofrece mucha claridad en un tema muy sencillo – ¿quién realmente participa de las intervenciones que buscan reducir las desigualdades sanitarias, y *por qué?*



## Bibliografía

- Blamey A, Ayana M, Lawson L, Mackinnon J, Paterson I, and Judge K (2004). *Final report of the independent evaluation of the National Health Demonstration Project Have a Heart Paisley (Phase One)*. Public Health and Health Policy, University of Glasgow. Disponible en <http://www.scotland.gov.uk/Publications/2005/03/20836/54354>
- Blamey A and Mackenzie M (2007) Theories of change and realistic evaluation: peas in a pod or apples and oranges? *Evaluation* 13(4) 439-455.
- Connell J, Cubish A, Schorr L and Weiss C (eds.) *New Approaches to Evaluating Community Initiatives (Vol. 1)* Washington DC: Aspen Institute.
- Franklin A and Ebdon C (2005) 'Practical experience: building bridges between science and practice' *Administrative Theory and Praxis* 27(4) 628-649.
- Heckathorn, D. (1997). Respondent-driven sampling: a new approach to the study of hidden populations. *Social Problems* 44: 174-199.
- Heckathorn, D. (2002). Respondent driven sampling II: deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems* 49: 11-34.
- Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M and Brindle P (2007). 'Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study'. *BMJ* 335(136).
- Judge K and Bauld L (2001) 'Strong theory, flexible methods: evaluating complex community-based initiatives'. *Critical Public Health* 11(1) 19-38.
- Kaplan S and Garrett G (2005) 'The use of logic models by community based initiatives' *Evaluation and Programme Planning* 28(2) 167-172.
- Lyratzopoulos G, Heller P, McElduff M, Hanily M, and Lewis P(2006). 'Deprivation and trends in blood pressure, cholesterol, body mass index and smoking among participants of a UK primary care-based cardiovascular risk factor screening programme: both narrowing and widening in cardiovascular risk factor inequalities'. *Heart* 92: 1198-1206.
- Mitchell K (1997) 'Encouraging young women to exercise: can teenage magazines play a role?' *Health Education Journal* 56(2) 264-273.
- Nestle M (2003) *Food Politics: How the Food Industry Influences Nutrition and Health* Berkeley: University of California Press
- Pawson R (2003) 'Nothing as practical as a good theory' *Evaluation* 9(4) 471-490.

- Pawson R (2006a) *Evidence Based Policy: A Realist Perspective* London: Sage
- Pawson R (2006b) 'Simple Principles for The Evaluation of Complex Programmes' in A Killoran & A Kelly (eds.) *Evidence Based Public Health*, Oxford: Oxford University Press.
- Pawson R and Tilley N (1997) *Realistic Evaluation* London: Sage.
- Pollitt C (2000) 'Institutional amnesia: a paradox of the 'information society''. *Prometheus* 18(1) 5-12.
- Sanderson I (2003) Is it 'what works' that matters? Evaluation and Evidence-Based Policy Making, *Journal of Research Papers in Education*. 18(4): 329-343.
- Salganik, M, and Heckathorn D (2004). Sampling and estimation in hidden populations using respondent-driven sampling. *Sociological Methodology* 34: 193-239.
- Scottish Executive. (2005 a). *Delivering for health*.  
<http://www.scotland.gov.uk/Resource/Doc/76169/0018996.pdf>
- Scottish Executive (2005b). *Building a health service fit for the future*.  
<http://www.scotland.gov.uk/Resource/Doc/924/0012113.pdf>.
- Shackley S and Wynne B (1997) 'Global warming potentials: ambiguity or precision as an aid to policy?' *Climate Research* 8(1) 89-106.
- Sloan H, Samson, K, O'Suilleabhain J and Barrie L (2006). *Have a Heart Paisley Mapping and Needs Assessment Report: Winning Hearts and Minds*. Have a Heart Paisley.
- Snowdon W, Schultz J and Swinburn B (2008) 'Problem and solution trees: a practical approach for identifying potential interventions to improve nutrition' *Health Promotion International* 23(4) 345-353.
- Sridharan S, Gnich W, Moffat V, Bolton J, Harkins C, Hume M, Nakaima A, MacDougall I, and Docherty, P. (2008) *Evaluation of Primary Prevention Intervention: Have a Heart Paisley Phase 2*. Glasgow : NHS Health Scotland. 2008. Disponible en:  
<http://www.chs.med.ed.ac.uk/ruhbc/evaluation/hahpfinal>
- Stame N (2004) 'Theory-based evaluation and types of complexity' *Evaluation* 10(1) 58-76.
- Trochim W (1989) *Concept Mapping for Evaluation and Planning* Special issue Evaluation and Program Planning New York: Pergamon Press.
- Tunstall-Pedoe H and Woodward M (2006). By neglecting deprivation, cardiovascular risk scoring exacerbate social gradients in disease. *Heart* 92: 307-310.
- Vandenbroek P, Gossens J, Clemens M (2007) *Tackling Obesities: Future Choices – Building the Obesity Systems Map*. Government office for Science, Foresight Programme. London: Department of Innovation, Universities and Skills.

- Velupillai N, Packard C, Batty G, Bezlyak V, Burns H, Cavanagh J, Deans K, Ford I, McGinty A, Millar K, Sattar N, Shiels P and Tannahill C (2008). Psychological, social and biological determinants of ill health (pSoBid): Study Protocol of a population-based study. *BMC Public Health* 8:126. <http://www.biomedcentral.com/1471-2458/8/126>.
- Weiss C (1995) 'Nothing as practical as a good theory' in Connell J et al (eds.) *New Approaches to Evaluating Community Initiatives (Vol. 1)* Washington DC: Aspen Institute.
- Weiss C (2000) 'Which links in which theories should we evaluate?' in P Rogers et al *Program Theory in Evaluation: Challenges and Opportunities* (New Directions for Evaluation no 87) San Francisco: Jossey Bass.

**Figura 1: Formalizando las teorías del programa**

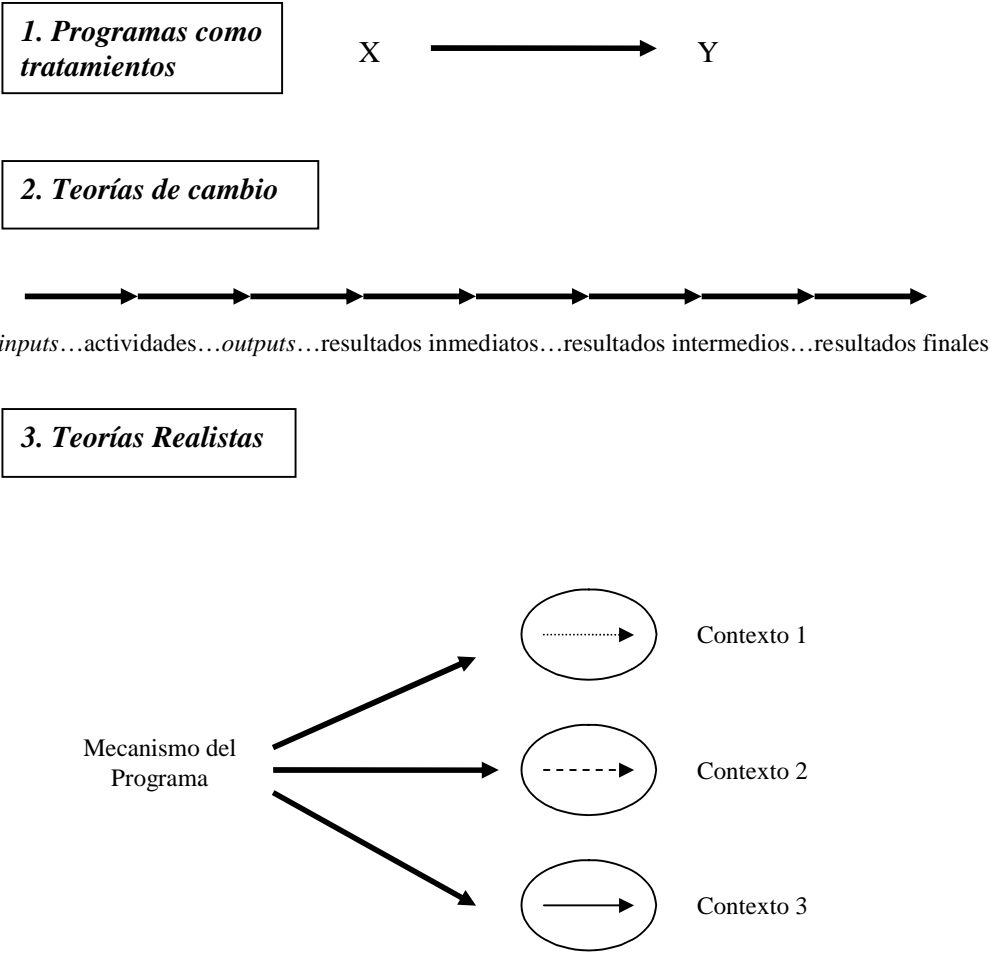
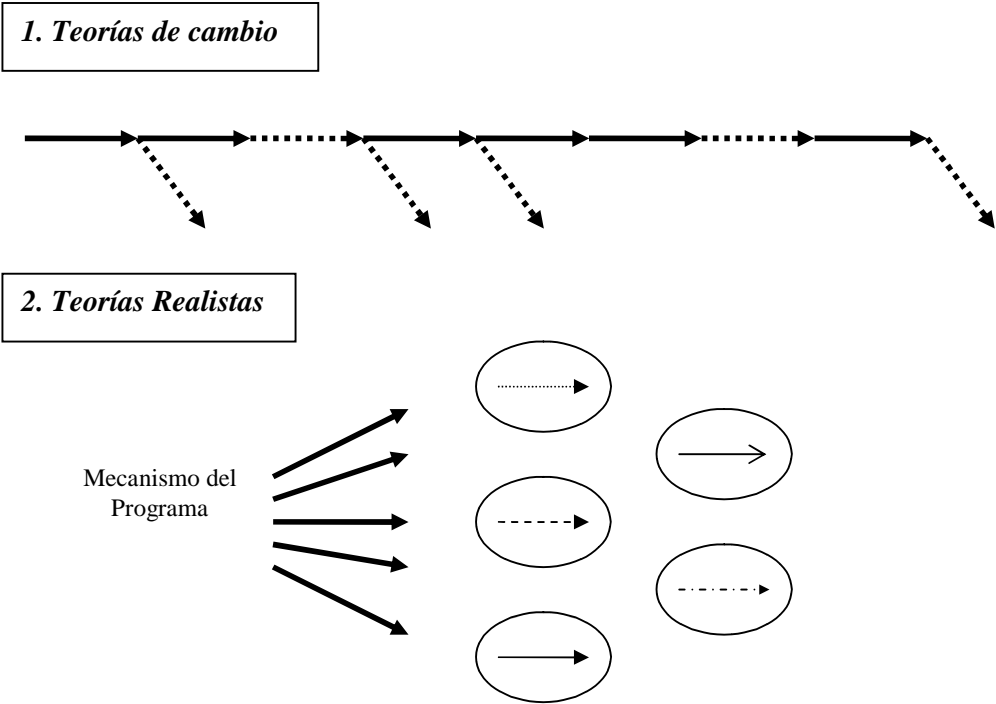


Figura 2: Refinando las teorías del programa



**Tabla 1. Medidas de Prevención Primaria del HaHP**

***Contacto***

El HaHP fue diseñado para individuos de entre 45 y 60 años, a través de la entrega de un sistema especial de prevención primaria. Una base de datos centralizada, que reunió datos de los centros de salud primarios, laboratorios, hospitales y estadísticas nacionales, fue utilizada para apoyar la entrega de prevención primaria, y como marco de muestreo para la intervención. A todos los individuos que reunían los requisitos necesarios (entre 45 y 60 años, sin historia de enfermedad cardiovascular, viviendo en Paisley, y registrados con un médico de Paisley) se les envió una invitación personal a través del correo para participar de “chequeo médico al corazón”. Las cartas fueron dirigidas, en primera instancia, a los sectores más empobrecidos. Posteriormente, se envió la intervención a otros sectores postales. Las personas que querían aceptar la invitación a participar de un chequeo médico al corazón tenían que contactar a un centro de llamados telefónicos, cuyo número fue enviado junto con la carta de invitación.

***Exámenes***

Todos aquellos que aceptaron esta invitación fueron examinados por las enfermeras del HaHP. Esto permitió calcular su probabilidad de desarrollar una enfermedad cardiovascular coronaria en los próximos 10 años (*Framingham Risk Score*), y a los individuos se les informó de su estatus de riesgo (bajo, moderado o alto).

***Educación en Salud***

Después, se les ofreció a todas las personas examinadas la oportunidad de reunirse con un “monitor de salud” del HaHP. Se pretendía que, en la segunda etapa del HaHP, la función del monitor de salud fuera un método para llegar a la población objetiva de manera individual, con un apoyo a los usuarios y orientación personalizada que empoderara a estos individuos para realizar cambios positivos en sus estilos de vida, con la intención de reducir su riesgo de desarrollar enfermedades coronarias.

***Micro-intervenciones***

De manera adicional, los monitores de salud tenían la opción de “derivar” a ciertos individuos hacia “micro-intervenciones” del HaHP u otros servicios comunitarios apropiados, identificados durante un ejercicio local de mapeo (Sloan et al., 2006). Los monitores de salud utilizaron una base de datos electrónica para orientar sus consultas y capturar detalles de los servicios apropiados para derivar a los individuos seleccionados.

**Figura 3. Lógica de Programa del HaHP**



**Tabla 2. Recolección y Análisis de Datos para explorar las actividades de alcance**

| <b>Fuente de Datos</b>                      | <b>Preguntas de Investigación</b>  | <b>Recolección y Análisis de Datos</b>  |
|---|--|---|
| <b>Actores del Programa</b>                 | <p>¿Tuvo la intervención el peso suficiente como para tener un impacto en las desigualdades sanitarias?</p> <p>¿Tuvo el programa la capacidad de llegar a individuos con múltiples desventajas que normalmente no participan dentro del sistema de salud? ¿Cómo podría haberse implementado el programa para maximizar su alcance?</p> | 15 entrevistas semi-estructuradas en profundidad con actores claves involucrados en la planificación e implementación inicial del HaHP.   |
| <b>Beneficiarios esperados del programa</b> | <p>¿Qué aspecto de la intervención hizo que fuera atractiva para que los individuos participaran en la intervención? ¿Cómo difiere esta intervención de las intervenciones normales del servicio de salud? ¿Era probable que esta intervención fuera atractiva para individuos con múltiples desventajas?</p>                          | Las entrevistas semi-estructuradas fueron llevadas a cabo con los usuarios de prevención primaria que llegaron a los exámenes. Solo se consiguió “permiso ético” para entrevistar a aquellos que estaban dispuestos a participar de la intervención.  |
| <b>Datos administrativos</b>                | <p>¿Era más probable que los individuos de sectores pobres (según la medición del Índice Escocés de Múltiples Desventajas) participaran en HAHP?</p> <p>¿Era más probable que los individuos más pobres (según la medición del consejo tributario local en el sector donde vivían) participaran en HAHP?</p>                           | Se enviaron cartas de invitación a un total de 11.277 individuos. Los datos sobre la participación en las múltiples etapas de la intervención fueron obtenidos por el equipo de evaluación interna del programa. Para contestar estas preguntas, se implementaron modelos de múltiples variables y niveles. |