



UNIVERSIDAD DE COSTA RICA  
FACULTAD DE CIENCIAS ECONOMICAS  
ESCUELA DE ESTADISTICA  
MAESTRIA PROFESIONAL EN POBLACION Y SALUD

Meta-Evaluación del “Programa Antidrogas y Salud Sexual y  
Reproductiva para Adolescentes”, Jamaica 1999-2001:  
Lecciones aprendidas y Contribución Metodológica

Proyecto de Graduación para optar al título de Master Profesional en  
Población y Salud

*Helen Larissa Centeno Monge*

San Pedro Montes de Oca, Costa Rica  
Julio, 2002

# **Tabla de Contenidos**

- I. Introducción**
- II. Antecedentes**
  - a. Ubicación del Programa
- III. Justificación de la Metaevaluación para este programa**
- IV. Referentes Teórico-conceptuales**
  - a. Concepto de “Metaevaluación para Proyectos y Programas Sociales”
  - b. Necesidad de la Metaevaluación
  - c. Etapas de una Metaevaluación
  - d. Pasos de una Metaevaluación
  - e. Criterios de una Metaevaluación
  - f. Limitaciones de la Evaluación
- V. Objetivos**
- VI. Metodología**
  - a. Fuente de información
  - b. Descripción de la Metodología
- VII. Limitaciones del estudio**
- VIII. Resultados y Discusión**
- IX. Conclusiones y lecciones aprendidas**
- X. Anexos**
- XI. Bibliografía**

## **I. Introducción**

En el presente trabajo se pretende realizar una Meta evaluación –nueva evaluación de la evaluación - del Programa Antidrogas y Salud Sexual y Reproductiva para Adolescentes, que se implementó en escuelas públicas de la ciudad de Kingston, Jamaica, durante los años de 1999 y 2001.

Meta evaluación que persigue evaluar la planificación e implementación de este programa así como también evaluar las estrategias de análisis estadístico con las que se trabajaron los datos y las interpretaciones que se les dieron a los resultados obtenidos durante el proceso de Evaluación de Impacto que se realizó en la Universidad de Tulane, New Orleans. Investigación que se realizó como respuesta a la inquietud de la Organización patrocinadora, de saber el impacto logrado con su programa y así promover su continuidad.

A la vez este proyecto pretende legar una contribución metodológica basada en la Teoría Psicométrica, que fortalezca los conocimientos de futuros Evaluadores y los incentive a darle importancia tanto a los análisis cuantitativos como los cualitativos dentro de una Evaluación de Impacto, dejando en claro que si se cuenta con otras alternativas metodológicas y técnicas adecuadas se pueden obtener resultados más claramente interpretables –estadísticamente hablando- y al final del proceso Evaluaciones confiables, útiles y eficaces para la toma de decisiones – cualitativamente hablando-.

Para finalizar se pretende enunciar cuales fueron las lecciones aprendidas durante el proceso de la Evaluación de Impacto y la Metaevaluación, como muestra de lo que hasta hoy se realiza en este tipo de Investigaciones y los retos y desafíos que nos quedan a los que pretendemos continuar trabajando en esta área de Evaluación de Impacto de Programas de Salud y Proyectos Sociales.

## II. Antecedentes

### a) Ubicación del Programa evaluado

El nombre original del programa es “Addiction Alert Adolescent Drug Prevention Program” tiene como propósitos primordiales el de “promover un estilo de vida saludable, libre de drogas, prevenir conductas autodestructivas y/o antisociales -incluyendo abuso en el uso de drogas- y aumentar el conocimiento sobre temas de Salud Sexual y Reproductiva entre la juventud de alto riesgo que habita en el área Metropolitana de Kingston-Jamaica” (Ver Anexo#).

Los “tópicos” que se incluyeron en el programa fueron “abuso en el uso de drogas, salud sexual y reproductiva y resolución de conflictos personales y sociales”; y los “componentes” principales de la intervención eran: “poder crear un plan de estudios que contemplara la mayor participación de los alumnos, la capacitación y entrenamiento de adolescentes más viejos cuyo trabajo sería ser facilitadores del programa (edad promedio de 18 años), y la creación de Grupos de Apoyo en las escuelas participantes que funcionarían como entes de retroalimentación del programa una vez éste caducara”. (Ver anexo#)

El programa buscaba que mientras se estuviera implementando durante todo el año escolar se lograra influir y se generaran cambios significativos, en dimensiones tales como comportamiento, actitud y conocimientos respecto al uso y abuso de drogas y la salud sexual y reproductiva, en adolescentes entre los 14 y los 17 años, que asistían a las escuelas públicas del área Metropolitana de Kingston, específicamente. Dicha influencia pretendía aumentar los conocimientos y mejorar las actitudes y comportamientos de estos adolescentes.

#### ORGANIZACIÓN QUE LO PATROCINA

La Organización patrocinadora de dicho programa es la “Addiction Alert Organization” (AAP), que nace en 1993 y se establece en la ciudad de Kingston, Jamaica.

Dicha Organización nace con propósitos de ser una Institución especialista en problemas sociales de los adolescentes, como son uso y abuso de drogas, sus implicaciones familiares, sociales, económicas y políticas, deserciones en las escuelas, así como el estudio de su incidencia y prevalencia. Y no es sino hasta finales de 1997 que dicha Organización

introduce dentro de sus propósitos de intervención tópicos relacionados con Salud Sexual y Reproductiva.

Para lo cual cuenta con clínicas asistenciales tanto para el individuo que adolece de problemas de adicción como para sus familiares, personal especializado para tratar este tipo de problemas, folletería educativa, programas educativos, etc.

Conjuntamente con Hope Enterprises Corporation y FOCUS on Young Adults Program/Family Health International, la Universidad de Tulane, a través de Measure Evaluation realizó la Evaluación de Impacto y Monitoreo de este programa; y es con base en el proceso que implicó esta evaluación de impacto (conformación del equipo evaluador, planificación de la evaluación, determinación de las estrategias de análisis estadísticos que se utilizarían y las interpretaciones que se le dieron a los resultados finales) que nace la inquietud y se desarrolla la Metaevaluación que contempla mi proyecto de graduación.

### **III. Justificación del proyecto**

Ante la evidencia de que los resultados finales, obtenidos en la Evaluación de Impacto realizada en Tulane para el Programa Antidrogas y Salud Sexual y Reproductiva para Adolescentes” realizada en Tulane, no fueron uniformes para las tres dimensiones (comportamiento, actitudes y conocimientos) que buscaba modificar el programa y que éstos dependían del análisis estadístico que se les hiciera (análisis bi-variable o análisis multivariable) se crea y justifica la necesidad de realizar una Metaevaluación, es decir una nueva evaluación de dicha evaluación.

Justificación que encuentra su sustento en el hecho de que cuando se realizó el análisis bivariable (tablas cruzadas), en Tulane, los resultados mostraron una disminución en algunos de los indicadores de las dimensiones: conocimientos y actitudes, que nos harían pensar en un total fracaso del programa al menos en estas dimensiones, y los resultados obtenidos realizando análisis multivariable reflejaron un efecto positivo del programa para estas mismas dimensiones de conocimientos y actitudes pero *sólo* para los indicadores que median el uso de alcohol y drogas, no así para los indicadores pertenecientes a la dimensión denominada “comportamiento”, en la cual se observaron resultados poco convincentes para catalogar un efecto positivo del programa, situación que se repite con los indicadores de

SSR que se tomaron en cuenta para el estudio, para los cuales se puede decir que en ninguna de las tres dimensiones se observaron resultados aplastantes que permitieran catalogar el programa como efectivo. El único cambio en esta área del programa, que reflejaba la dirección deseada al implementar cualquier programa de Salud (disminución de conductas de riesgo, aumento en el conocimiento y actitudes positivas) se observó en un pequeño porcentaje de alumnos, los cuales reportaron haber usado el condón en su última relación sexual.

También se esperaba encontrar diferencias de impacto según el tiempo de implementación del programa, es decir, esperábamos que a más tiempo de recibido el programa mayores conocimientos y cambios más significativos en el comportamiento y actitudes de los adolescentes lastimosamente en un 90% de los indicadores (variables) estudiados lo opuesto fue lo que obtuvimos.

Es así como se llegó a la conclusión durante dicha evaluación de que el programa Antidrogas y de Salud Sexual y Reproductiva para los adolescentes, llevado a cabo durante los años 1999-2001 en las escuelas públicas de Kingston Jamaica, por medio de la Addiction Alert Organization (AAO) resultó no ser el 100% satisfactorio, es decir no tener un impacto significativo en todas las dimensiones que se pretendían.

Es entonces ante dicha evidencia (por análisis *únicamente* estadístico) que surge, como se mencionó anteriormente, la necesidad tanto de la organización patrocinadora como la mía propia, de hacer una Metaevaluación, es decir, una evaluación de la evaluación (Stufflebeam, 1974 y Scott, 1998) de dicho programa, mediante la cual se pudiera analizar de forma integral el proceso que implicó la planeación, implementación y resultados de dicho programa.

Oportunidad que aprovecho para poner en práctica y reforzar los conocimientos adquiridos en esta maestría, en cuanto a Evaluación de Impacto de Programas de Salud se refiere, y que a la vez considero adecuada para enumerar lecciones aprendidas durante la etapa de Evaluación y también durante la Metaevaluación, las cuales considero importantes de tomar en cuenta, tanto por las organizaciones que implementan este tipo de programas como por los que pretendemos continuar trabajando en el campo de la Evaluación en un futuro.

La Metaevaluación realizada en este proyecto contempla cuatro etapas:

#### Diseño del programa

Se plantea dentro de este proyecto Metaevaluar dicho programa, comenzando con la etapa de diseño del programa, en el cual nos interesaría constatar la presencia de una “*valoración de necesidades*” de acuerdo al entorno socio-económico y político dentro del cual se estableció el programa, ya que según la bibliografía consultada y la experiencia de Evaluadores de Impacto “La efectividad de un programa cualquiera que sea su énfasis y que pretenda incrementar las condiciones sociales de una población meta, estará fuertemente determinada en la medida en que el mismo este dirigido realmente a las necesidades sociales y que lo haga de una manera que responda a las verdaderas circunstancias de aquellos en necesidad”( Rossi y colaboradores, 1999) .

El otro aspecto importante que se debe determinar dentro de la etapa del diseño de dicho programa es si existió la articulación de una conceptualización o “*teoría del programa*” de acuerdo a las personas que lo realizaron, teoría que tendría que haber constado con el establecimiento de las relaciones entre los insumos, actividades y resultados intermedios del eventual programa (Hipótesis de acción), y la identificación de una concepción de lo que debería hacerse o en que aspectos debería intervenir para lograr los beneficios sociales que se pretenden alcanzar, es decir, saber como se supone que el programa logrará los resultados deseados (Hipótesis conceptual) (Rossi et al, 1999).

#### Instrumento

Tomando en cuenta que durante el diseño de un programa una parte importante la constituye la construcción y validación de un instrumento con el cual se medirán diferencias entre un *antes* del programa y un *después* del mismo, se consideró necesario analizar esta etapa y hacer una evaluación del instrumento (cuestionario utilizado en el programa) con el fin de examinar la validez y la confiabilidad Psicométrica de éste. “La validez esta asociada al grado en que el instrumento, como un todo, y los ítems que lo componen, representan adecuadamente al constructo teórico que se quiere medir y a sus componentes, así como es indispensable el manejo de los referentes teóricos y su correcta operacionalización” (Anastasi, 1999). “La confiabilidad significa la precisión del instrumento, su consistencia en la medición” (Anastasi,1999).

### Análisis de datos llevada a cabo en la Universidad de Tulane

En esta etapa, se pondría especial énfasis en la valoración de las estrategias y metodologías cuantitativas y cualitativas utilizadas para el análisis de los datos con los que se contaba, ya que de ello podrían depender los resultados que se obtuvieron y se tendría la oportunidad de plantear una metodología estadística diferente que reflejara cambios importantes en esos resultados.

Aporte metodológico que se basaría en los principios de la Teoría Psicométrica de modo que ayudaría a incrementar el poder estadístico de las pruebas de significancia utilizadas durante la evaluación, con el único fin de confirmar o modificar resultados ya existentes.

### Valoración de la interpretación

En la que se evaluaría la confiabilidad y veracidad de los datos, su precisión y fiabilidad, su validez externa para ser generalizados y la objetividad con la que se publicaron.

## **IV. Referentes Teórico-conceptuales**

### **a) Conceptos de Metaevaluación**

Para entender la Metaevaluación en el contexto de la Evaluación de Programas y Proyectos se utilizarán acepciones de este concepto.

La primera define la Metaevaluación como “una fase o etapa, usualmente terminal, del trabajo o proyecto de evaluación: es una evaluación de la evaluación previa” (Díaz, 2000) Aunque el prefijo “meta” refleja connotaciones temporales (después de), no necesariamente se realiza la Metaevaluación una vez finalizado el proceso de evaluación. Esta puede realizarse al mismo tiempo.

Es importante resaltar que no se presume como condición necesaria la existencia de un metaevaluador externo, ajeno al proyecto y a la evaluación. La Metaevaluación bien puede ser ejecutada por los mismos agentes bajo un replanteamiento de las finalidades que se buscan.

La segunda acepción define la Metaevaluación como “un proceso que pretende fijar los criterios que nos permitan determinar la calidad de la evaluación y una vez diseñados dichos criterios el meta-evaluador tendrá que elegir los que más ajusten al caso concreto, y

a las necesidades y pretensiones de las organizaciones patrocinadoras de programas y proyectos de salud” (Sida Advisory Team –SAT- Consultants, 2001).

La tercera acepción hace énfasis a un proceso que busca evaluar y describir una actividad de evaluación, así como también juzgarla bajo ciertos criterios que caracterizan una “buena evaluación” (Stufflebeam, 1969).

## **b) Necesidad de la Metaevaluación**

En los últimos años el tema de la Metaevaluación ha estado en auge, debido a la creciente necesidad de los expertos en Evaluación de mejorar su trabajo, fundamentados en nuevas conceptualizaciones y herramientas dentro del área de la Evaluación de Programas y Proyectos Sociales.

Otras razones:

- Porque la Metaevaluación es una valoración al mérito por los esfuerzos realizados en la Evaluación previa.
- Porque la Evaluación sirve para tomar decisiones y responsabilidades, y la Metaevaluación proporcionará información para apoyar activamente dichas decisiones así como información retroactiva a los futuros evaluadores para que se concienticen en sus siguientes trabajos de Evaluación.
- Porque la Metaevaluación es formativa y retroalimenta con nuevos conocimientos.
- La Metaevaluación tiene como propósito ayudar a los evaluadores a mejorar su práctica, ya que se centra en los problemas metodológicos, provee información en relación con la calidad de los estudios evaluativos bajo ciertos estándares y conduce a incrementar la responsabilidad de los evaluadores.
- Porque la Metaevaluación evalúa la importancia de los objetivos de la evaluación, la adecuación de planes y diseños de evaluación dependiendo del programa social a evaluar, el proceso de una evaluación (tomando en cuenta todos los problemas que se pueden presentar y el impacto que pueden representar en los resultados esperados) así como la calidad e importancia de los resultados obtenidos gracias a la evaluación.
- Porque mientras más ajenos al programa sean los agentes evaluadores, la objetividad de los resultados es mayor, y por lo general es en la Metaevaluación que se contratan equipos externos.

- Porque la Metaevaluación describe y juzga los trabajos de evaluación previos y brinda recomendaciones de cómo deben mejorarse y como los hallazgos pueden usarse apropiadamente.
- La Metaevaluación busca encontrar la utilidad de los reportes de evaluaciones realizadas, es decir, si éstos contienen información útil para practicantes o posteriores evaluadores, y de no ser así plantearlas en esta etapa.

Las primeras tres citadas por Scriven (1969) y las siguientes por Stufflebeam (1969) ambos considerados los padres y creadores de la Teoría de Evaluación de Programas y Proyectos Sociales.

La Metaevaluación ha pretendido conocer cuál ha sido el desarrollo de la evaluación en todas sus etapas, tanto en el diseño general como en los instrumentos utilizados en la recogida de información, así como también analizar las metodologías utilizadas para obtener los resultados y sus interpretaciones desde la perspectiva de las personas que conforman el equipo de evaluación.

### **c) Etapas de una Metaevaluación**

El proceso de la Metaevaluación de una Evaluación para un programa cualquiera que este sea debe incluir las siguientes etapas (SAT, 2001):

1. La organización de la Evaluación
2. La selección del equipo que realizará la Evaluación.
3. El material teórico (cualitativo) y los datos (material cuantitativo) con las que se realizará la Evaluación.
4. Los métodos que se utilizaran para obtener resultados en la Evaluación.
5. La interpretación de dichos resultados.

### **d) Pasos de una Metaevaluación**

Para Michael Scriven, citado por Stufflebeam en 1969, los pasos para conducir una buena Metaevaluación son:

1. La caracterización del programa evaluado
2. Verificación la relación causa-efecto (impacto del programa en la población meta)
3. Metaevaluar las consecuencias de esta relación
4. Metaevaluar las conclusiones obtenidas

5. Metaevaluar los costos de la implementación de un programa
6. Identificar y evaluar si los objetivos del programa se lograron con la planificación de actividades e implementación del mismo.
7. Metaevaluar la valoración de necesidades y la teoría del programa para poder brindar un juicio de la importancia del programa.
8. Metaevaluar la crítica que pueden hacer los diferentes integrantes del equipo evaluador.
9. Metaevaluar la formulación de la valoración final que se hace al programa.

### **e) Criterios de aplicación de una Metaevaluación**

Además hay que tomar en cuenta ciertos criterios importantes que la Metaevaluación considera esenciales para una adecuada técnica en el proceso (Stufflebeam, 1974):

1. Validez interna: criterio que hace referencia a la veracidad de los hallazgos o resultados obtenidos en la Evaluación y responde a las siguientes preguntas:
  - ¿Se cuenta con suficiente información sobre el programa a ser evaluado?
  - ¿Será que el diseño de Evaluación planteado para el programa respondió a los objetivos que buscaba la organización que lo implementó?
  - ¿Será la metodología cualitativa y cuantitativa la correcta para obtener resultados verídicos?
  - ¿Son los resultados certeros e inequívocos?

Puesto que parte de la consideración que todo estudio, ya sea este una investigación o una evaluación, tiene un número mínimo de respuestas exactas o resultados esperados ante los objetivos planteados al inicio de cualquier programa.

2. Validez externa: criterio que se refiere a la generalización de los resultados obtenidos durante la evaluación. ¿A qué objetivos del programa responden los resultados obtenidos? ¿A qué población favorecen estos resultados obtenidos? ¿La información (resultados) obtenidos responden únicamente a la muestra estudiada o se pueden generalizar a otros grupos dentro del mismo contexto en el que se llevó a

cabo la implementación del programa? ¿Son los resultados predictivos de lo que puede ocurrir en posteriores aplicaciones de programas de este tipo?

Conocer la validez externa significa que uno puede generalizar con seguridad los resultados a ciertas poblaciones de interés, algunas de las condiciones del programa y algunas circunstancias de desarrollo. Sin embargo hay que dejar en claro la extrapolación que el evaluador hace, entre los resultados obtenidos y lo que se quiere demostrar.

3. **Fiabilidad:** criterio que hace referencia a la confiabilidad y exactitud de los datos que se utilizan para obtener los resultados de una evaluación. Pero también hace referencia a la consistencia interna de los resultados obtenidos, es decir, que la consistencia que presentaron los resultados en la primera evaluación sea la misma en la Metaevaluación. También hay que considerar que si los resultados obtenidos carecen de precisión y reproducibilidad, puede deberse a : que se deben al azar o se aplicó una metodología errada o incompleta para obtenerlos y por ende carecen de sentido y fiabilidad.
4. **Objetividad de los resultados:** se refiere a como se publican los resultados. ¿Fueron los resultados publicados con juicios competentes y documentados sustantivamente? O simplemente responden a la experiencia de los que participaron en la evaluación, a las percepciones de cada uno de ellos o están sesgados por los evaluadores. Y también pueden responder a los intereses de las organizaciones financiadoras, intereses propios de las organizaciones ejecutoras de los programas e incluso a ciertas predilecciones de los diferentes integrantes del equipo evaluador.

A menos que los resultados de la evaluación puedan ser interpretados de igual forma por diferentes y competentes expertos en esta área, el verdadero significado de los resultados serán siempre sometidos a cuestionamientos.

5. **Relevancia:** criterio que hace referencia a si los resultados obtenidos responden a los propósitos de la evaluación.

El criterio de relevancia de los resultados obtenidos tiene mayor fuerza si éstos van a servir para otros fines y no solo los académicos.

La relevancia de una Metaevaluación se ve determinada por la comparación de cada dato o resultado obtenido con las preguntas u objetivos planteados al principio de la evaluación.

6. **Importancia:** criterio que evalúa si existió determinación correcta de cuales deberían ser los datos o resultados a obtener, ya que en toda investigación o evaluación existen una gran cantidad de datos o información que parece tener relevancia potencial en cuanto a los propósitos de los programas, sin embargo el evaluador debería escoger los datos que el considere más útiles para el propósito de su evaluación, tomando en cuenta las prioridades de las audiencias a las que llegara el reporte final de la evaluación y basado en su propios juicios sustantivos.
  
7. **Alcance:** hasta donde pueden llegar los resultados obtenidos (reporte final) y que tanto beneficiarán a las organizaciones patrocinadoras y la población meta definida.
  
8. **Credibilidad:** hace referencia en la confianza y credibilidad que ponen las organizaciones sobre el equipo evaluador de sus programas. Hay que recordar que las organizaciones, por lo general, depositan mayor confianza en equipos de evaluadores externos por su independencia con la institución y por la objetividad que esta condición le pueda brindar al reporte final de una Evaluación.  
El meta-evaluador tiene que conocer que tanta confianza ha depositado la organización patrocinadora en la evaluación, ya sea que esta haya sido realizada por un agente interno o externo. El evaluador se encargará de asegurar una gran confiabilidad de su evaluación en la medida que sea abierto y transparente en mostrar los resultados tal cual se obtuvieron y consistente en demostrar su profesionalismo.
  
9. **Tiempo:** criterio muy importante puesto que en muchas ocasiones las organizaciones limitan el tiempo para obtener algún resultado. Por lo que el equipo

evaluador tiene que tenerlo presente a la hora de planificar las actividades de la evaluación.

Es importante tener en mente que “aun así se obtenga la mejor información o resultados impactantes, éstos serán inútiles si son entregados demasiado tarde para los propósitos que buscaba la organización patrocinadora” (Scriven, 1969). El evaluador tiene que esforzarse para presentar información razonable y buena a las personas interesadas en el tiempo que ellas la requieran.

10. Difusión: criterio que hace referencia a la posibilidad que las personas que están interesadas en los resultados obtenidos de la evaluación verdaderamente la reciban y la utilicen.
11. Relación costo-efecto: hace referencia a que el metaevaluador tiene que analizar si la realización de la evaluación requirió de bajos costos, sin que ello significara sacrificar la calidad de la misma. Al final del proceso el evaluador tiene que demostrar que los resultados de la Evaluación valen la pena para el costo que implicó su realización.

En resumen las Evaluaciones tienen que ser técnicamente adecuadas, útiles y eficaces, y la Metaevaluación persigue demostrarlo, para lo cual los once criterios anteriormente mencionados se hacen indispensables para que los metaevaluadores, valoren evaluaciones de diseños de programas y proyectos sociales y de Evaluaciones de Impacto, confiabilidad y validez de instrumentos psicométricos utilizados durante las mediciones de impacto, estrategias de análisis de datos o resultados obtenidos, valoraciones de las interpretaciones que se les dé a dichos resultados, reportes finales, etc.

### **f) Limitaciones de una Evaluación**

Cualquiera que sea el propósito de realizar una Metaevaluación, hay que tener en cuenta que el proceso previo a ésta (Evaluación en sí) pudo haberse visto sometido a ciertos problemas, ya sea antes, durante o después de la Evaluación. Básicamente la noción de estos problemas sirve para mejorar el trabajo que se realiza durante una Evaluación y si se

pasaron por alto durante esta fase, en el proceso de la Metaevaluación se pretende resaltarlos y dejarlos documentados (Objetivo importante de la Metaevaluación formativa) lo cual demuestra la utilidad de evaluar la evaluación y hacer un mejor diagnóstico de estas investigaciones.

Se hace una breve recopilación de ellos con el propósito de que próximos evaluadores los tengan en cuenta en futuros estudios de evaluación y se generen medidas preventivas adecuadas para afrontarlos.

1. Problemas conceptuales: que hacen referencia a como los diferentes integrantes del equipo evaluador conciben la evaluación, es decir, ¿qué entienden por evaluación?, ¿para qué sirve?, ¿quiénes tienen que realizarla? ¿cuáles son los propósitos de ella? ¿Qué información se requiere para hacer una evaluación de calidad? y ¿a quiénes beneficiaran con la evaluación?.
2. Problemas sociopolíticos: hacen referencia al contexto social, económico y político en el que se implementó el programa a evaluar, puesto que los objetivos y propósitos de los programas y proyectos sociales implementados por las organizaciones patrocinadoras pueden estar determinados por dicho contexto y la necesidad de una evaluación responder a intereses particulares de máximas autoridades. Por lo que el evaluador tendrá que diseñar una evaluación que incluya tanto a los involucrados en el programa como a las personas que se beneficiaran con él. Sin embargo hasta el momento ni los programas de entrenamiento en evaluación ni los libros de texto preparan a los evaluadores para que enfrenten este tipo de problemas.
3. Problemas legales y de contrato: Para que una evaluación de Impacto sea efectiva tiene que haber mutuo acuerdo entre las partes involucradas así como un compromiso de ética profesional: los evaluadores, los responsables de la investigación y el personal que intervino en la implementación del programa. Acuerdo que consiste en intercambio de información, disponibilidad de trabajar en conjunto cuando sea necesario (reuniones planificadas bajo agenda), etc.

#### 4. Problemas técnicos:

- Variables y objetivos del estudio: ¿cuáles serán las variables a incluir en la evaluación? ¿Cómo se operacionalizarán? Dado que existen tantas variables a poder ser estudiadas, se tendrán que tomar en cuenta aquellas que el evaluador considere bajo su juicio sustantivo importantes para sustentar los objetivos de la evaluación. Los metaevaluadores tendrán que revisar el diseño de la evaluación y las variables utilizadas, así poder incluir nuevas variables o sustraer algunas que consideren no necesarias. Para ello los metaevaluadores toman en cuenta la relevancia, el alcance y la importancia de las variables.

No hay que olvidar que se necesita que el personal involucrado en el programa plantee o defina sus objetivos específicos y así focalizar la evaluación en lo que ellos perciben como resultados deseables.

- El campo de trabajo y los instrumentos del evaluador: el cual incluye las condiciones bajo las cuales trabajará, la base de datos con la que cuenta (validez y confiabilidad), metodologías, instrumentos psicométricos y técnicas más apropiadas para obtener resultados, análisis de interpretación de resultados, dependiendo de los propósitos de la evaluación.
- La muestra: ¿cuál es la población que se evaluará? ¿Existe homogeneidad de condiciones al implementar un programa? O se parte de una muestra heterogénea? Responde esta población a una muestra representativa de las necesidades sociales que pretende cambiar el programa?
- La base de datos: ¿cómo fue almacenada, codificada y guardada? ¿Qué paquete estadístico se utilizó?, ¿bajo qué criterios se codificaron las variables que ella contiene?
- El análisis de los datos: tanto los métodos estadísticos como el análisis cualitativo son importantes en este tipo de evaluaciones. El evaluador tiene que tener presente que plan se utilizará para el análisis, puesto que de él dependerán los resultados, pero también tiene que tener en cuenta que tipo de análisis se puede obtener con los datos que cuenta.

- Estrategias de interpretación y reporte final: incluye organización y formato del reporte final, que resultados serán presentados, explicación de la interpretación que se hizo, definir las estrategias de comunicación de resultados.

#### 5. Problemas de dirección de la Evaluación:

- Organización del plan de Evaluación.
- Organización por jerarquías: Organización patrocinadora, director de la investigación, evaluadores, etc.
- Conformación del equipo de evaluadores.
- Facilidades: infraestructuras de las oficinas, equipo y material técnico con el que se cuenta para dar apoyo a la evaluación.
- Planificación de una agenda de trabajo.
- Entrenamiento para el equipo de evaluadores.
- Presupuesto para la evaluación.

#### 6. Problemas de ética y moral:

- Filosofía de los evaluadores
- Valores morales de los evaluadores
- Juicios éticos
- Objetividad de los evaluadores
- Perspectivas de utilidad: si la evaluación de programas y proyectos sociales es un ejercicio puramente académico o tendrá otra utilidad en el campo de la Evaluación.

En resumen, la Metaevaluación incluye el análisis, interpretación, explicación y la sistematización tanto del nivel teórico de la Evaluación como del práctico (metodológico), de los siguientes elementos:

- ❖ El proceso de la evaluación, desde el diseño hasta la presentación de resultados, incluyendo todas las actividades.

- ❖ Los resultados o productos de la evaluación, incluyendo metodologías para su obtención como las estrategias de interpretación.
- ❖ Los modelos de evaluación que se utilizan como referencia para realizar la evaluación.
- ❖ La metodología que nutre dichos modelos de evaluación.
- ❖ Los métodos, técnicas e instrumentos utilizados durante la evaluación.
- ❖ El equipo de evaluadores o sujetos que realizan la evaluación.

Sin embargo siempre hay que tener en cuenta y anticiparse a las principales limitaciones que un metaevaluador puede enfrentarse, ya que como es sabido la Metaevaluación proviene de los medios mismos con los que se cuenta, es decir, que se hace a la luz exclusivamente de los informes oficiales del proyecto y el reporte final de la evaluación, y es probable que en dichos documentos no se cuente con información explícita en varios de los aspectos que analiza la Metaevaluación.

## **V. Objetivos**

### **a) Objetivo General**

Realizar una Metaevaluación del Programa Antidrogas y Salud Sexual y Reproductiva que se implementó en los adolescentes de las escuelas públicas de Kingston, Jamaica en los años 1999-2001, y cuya fase de Evaluación se realizó en Enero-Febrero del 2002.

### **b) Objetivos Específicos**

1. Realizar un análisis de la etapa de *diseño* del programa, específicamente para la Teoría del Programa y la valoración de necesidades.
2. Realizar un análisis de la confiabilidad del instrumento (cuestionario) utilizado durante la *implementación* del programa
3. Valorar y en caso necesario modificar las estrategias que se utilizaron para el análisis de los datos, durante la fase de Evaluación de Impacto y que dieron pie a los resultados obtenidos.

4. Valorar y en caso necesario modificar la interpretación que se le dio a dichos resultados
5. Enumerar lecciones aprendidas de esta Evaluación de Impacto
6. Facilitar a futuros Evaluadores de Impacto y otras Organizaciones que implementan este tipo de programas una contribución metodológica en el área de la Metaevaluación.

## **VI. Metodología**

### **a) Fuente de información**

Este proyecto está basado en una serie de datos recolectados durante el ensayo comunitario de intervención y control que utilizó la Organización patrocinadora para la implementación del programa. Serie de datos que contiene información de todos los individuos que participaron en el programa tales como: datos personales, familiares, de la comunidad en la que habitan, de la escuela a la que asisten y los tópicos que se exploran están relacionados con el uso de drogas y con la salud sexual y reproductiva. La serie de datos fue otorgada por la Addiction Alert Organization, organización patrocinadora del programa, con el propósito que sirviera para la realización de una Evaluación de Impacto que les permitiera conocer la efectividad de su programa (Addiction Alert Adolescent Drug Prevention Program) en los adolescentes de Kingston, Jamaica.

Serie de datos que fue tabulada y codificada por personal de Hope Enterprises Ltd., organización que inició el proceso de Evaluación de Impacto, creando de esta manera una base de datos y posteriormente revisada y recodificada por el personal que complementó dicha Evaluación de Impacto en la Universidad de Tulane, equipo de evaluación en el que tuve la oportunidad de participar. La preparación de la base de datos, se llevó a cabo mediante la utilización del paquete computacional estadístico STATA, versión 6.

La base de datos cuenta con una muestra total de 16 escuelas, todas pertenecientes al área metropolitana de la ciudad de Kingston, Jamaica, el total de alumnos que participaron en la implementación del programa es de 4137, divididos de la siguiente manera:

	No. de escuelas	Alumnos
Escuelas control	8	1992
Escuelas intervención	8	2145
Total	16	4137

Es importante resaltar que del total de escuelas, 10 iniciaron la implementación del programa en el año de 1999 y únicamente 6 de ellas lo iniciaron en el año 2000. La medición para el seguimiento que evaluaría impacto se realizó para todas las escuelas en el año 2001, concluyendo que las primeras escuelas (10) recibieron el programa por dos años y el resto sólo por un año.

### **b) Descripción de la metodología**

Para valorar el primer objetivo que me había planteado en este proyecto (evaluar y analizar la teoría del programa y sus objetivos así como la valoración de necesidades, etapas correspondientes al diseño del programa) se esperaba contar con la información que nos pudiera brindar la Organización patrocinadora, desde Jamaica o en su defecto el director de la Investigación que realizamos en Tulane.

Para lo cual fue necesario hacer contactos vía telefónica y vía correo electrónico primero con la directora del programa en Jamaica, y al no obtener ninguna respuesta me puse en contacto via correo electrónico con el director de la Evaluación en Tulane, obteniendo respuestas negativas a mi petición, pues refirieron “desconocer si dicha información existía”.

Se constató que no había existido una validación previa del instrumento y una prueba piloto para él. Dato importante de tomar en cuenta pues con una prueba piloto se hubiese podido demostrar si dicho instrumento contaba con un nivel aceptable de calidad técnica que permitiese realizar análisis sustantivos en la mayoría de las dimensiones medidas (Anastasi, 1999 ).

Para efectos de la Metaevaluación, generación de los nuevos resultados y el análisis cuantitativo de la información, también se utilizó el programa estadístico STATA, pero versión 7.

Durante la Metaevaluación se realizó el análisis de confiabilidad del instrumento (utilizado en la medición de impacto) a través de metodologías psicométricas. En un principio se hizo indispensable decidir si los diferentes ítems del cuestionario respondían a algún constructo

o dimensión de las que el programa se había planteado medir y que ahora eran nuestro objeto de estudio: actitudes, comportamiento y conocimientos acerca de salud sexual y reproductiva y uso/abuso de drogas, si era así, a cual de ellos correspondían para poder agruparlos y configurar escalas (variables compuestas) en el cuestionario; pues la idea que se busca es disminuir el error de medición a través de la construcción de dichas variables. Luego correspondía hacer la validación de dichas escalas para lo cual se utilizaron dos parámetros psicométricos:

- 1) Índice de discriminación: para cada uno de los ítems que conformaban la escala, lo cual nos permitió identificar los ítems que diferenciaban a los individuos con puntajes bajos y aquellos con puntajes altos en la escala.

“Índice de discriminación: se define como la correlación entre el puntaje de cada ítem y el puntaje total del instrumento” (Anastasi, 1999).

“Normalmente un ítem se considera de alta calidad técnica si su discriminación es mayor o igual a 0.30” (Anastasi, 1999).

- 2) Coeficiente Alfa de Cronbach: el cual es una medida de confiabilidad total del instrumento, y estima la correlación entre el puntaje observado en el instrumento y el puntaje verdadero (hipotético) del individuo (Anastasi,1999), en este caso en particular, tomando en cuenta que se tratan de variables compuestas.

Se midió para cada uno de los ítems que conformaban las escalas construidas para el nuevo análisis de Metaevaluación.

Hay que tomar en cuenta que tanto la discriminación de ítems como la valoración del alfa de Cronbach, se fue realizando ítem por ítem, hasta lograr conformar escalas de alta confiabilidad. Es decir que las variables compuestas (escalas) quedaron conformadas con ítems que presentaban índices de discriminación mayores o iguales a 0.30 para cada uno de ellos y con un alfa de Cronbach mayor o igual a 0.7 tanto para cada ítem como para la variable compuesta (escala) en sí.

Para poder valorar el tercer objetivo de mi proyecto, referente a las estrategias para análisis y obtención de los resultados así como interpretación que se les dio a dichos resultados, hay que tomar en cuenta que una parte del análisis de los datos, dentro de la evaluación de impacto, se desarrolló con un análisis multivariado, utilizando modelos de regresión multinivel para la identificación de factores relacionados con el uso y abuso de drogas y

factores que influyen en el comportamiento y conocimiento acerca de salud sexual y reproductiva.

Sin embargo los indicadores de uso de drogas y comportamiento sexual que existen en el cuestionario son variables dicotómicas y se trabajaron sin hacerles ninguna transformación, luego se corrieron modelos de regresión logística que utilizaban cada uno de los indicadores de conocimiento, actitud y comportamiento como variables dependientes, para tratar de identificar los factores asociados a ellos. **Metodología estadística que no es la óptima?** para este tipo de análisis y por lo tanto podría reflejar errores de medición y significancia estadística de algunos resultados.

Por tal motivo en este proyecto propuse dentro de la metodología y para la parte de análisis multivariado, seguir el planteamiento de la Teoría Psicométrica, y considerar cada uno de estos indicadores como manifestaciones de un solo factor latente o constructo, al cual denominaremos según el caso. La medición de este constructo puede ser concebida como un puntaje en un “continuum”, de tal forma que permite desarrollar un modelo de regresión lineal para variables dependientes métricas. Por lo que por medio de la agregación de diferentes indicadores que teóricamente están representando un constructo “x” (construcción de variables compuestas) y considerando la mayor cantidad de ítems pertinentes en el cuestionario se buscaron cambios en los resultados obtenidos previamente en la evaluación de impacto de Tulane.

Tomando en cuenta que dicho “constructo” será considerado a partir de este momento como una “variable compuesta” y no dicotómica como se tenía inicialmente, y utilizándola como variable dependiente para las nuevas regresiones que se pretenden construir, se plantearon nuevos modelos de regresión y se buscó reducir el error de medición y proporcionar un mayor poder estadístico de los datos encontrados.

Importante tener en cuenta este nuevo indicador (variable compuesta) ya que las pruebas de hipótesis dependen del error estándar del estimador, el cual a su vez es función del error de medición asociado al indicador y el tamaño de la muestra, en nuestro caso en particular podría deberse al error de medición asociado al indicador que fue utilizado en Tulane.

“La obtención del valor verdadero de un constructo en una persona en particular, se logra por medio de la aplicación del dominio hipotético completo de reactivos o indicadores (universo o población de reactivos) que pueden construirse en relación con él mismo. El

propósito de cualquier medición particular es estimar ese valor verdadero” (Teoría Psicométrica, Nunnally y Bernstein, 1995).

La estimación de este valor puede estar afectado por un error de medición el cual consta de dos componentes:

- 1- Errores causados por procesos sistemáticos: puede cometerse cuando se usan instrumentos mal calibrados para efectuar las mediciones.
- 2- Errores aleatorios causados por el hecho de utilizar *solo una muestra aleatoria* de reactivos o ítems y no utilizar el conjunto completo de los mismos. “Una manera de disminuir este error es tratar de utilizar el mayor número posible de reactivos (ítems, o indicadores) para llevar a cabo la medición (Principio de Agregación, Nunnally y Bernstein, 1995)

Desde el inicio del proyecto se pretendía disminuir los errores de medición hasta donde fuese posible. “Los errores de medición aleatorios nunca se eliminan por completo, pero deben minimizarse tanto como sea posible, ya que limitan el grado de variabilidad de los indicadores y dificultan la exploración y verificación de relaciones sustantivas” (Nunnally y Bernstein, 1995)

Queriendo no limitar el grado de variabilidad de los indicadores a una medición dicotómica y facilitar la exploración y verificación de relaciones sustantivas planteé utilizar en esta ocasión diversos modelos de regresión lineal para variables métricas (a diferencia de los utilizados en Tulane: regresión logística) y análisis multinivel para la obtención de los nuevos resultados en la Metaevaluación.

Hacer mención que durante el proceso de la Metaevaluación decidí trabajar una nueva dimensión que no se valoró en Tulane, denominada “Percepción y Valoración del programa”.

Los modelos de regresión quedaron conformados así (se presentan únicamente las variables compuestas que se obtuvieron después del análisis de discriminación de los ítems).

#### **Constructos o dimensiones que se midieron**

- Percepción
- Conocimientos
- Comportamiento
- Actitudes

### **VARIABLES COMPUESTAS (DEPENDIENTES)**

Construidas con escalas de diferentes rangos :

Cada una de las siguientes variables fue utilizada como *variable dependiente* dentro de los diferentes modelos de regresión y sometida al efecto de las variables independientes detalladas mas adelante.

\*drgindv: Uso individual de drogas o actitud individual frente al uso de drogas (corresponden a esta variable compuesta las preguntas 29<sup>a</sup>, 29b, 29c, 29d, 29e del cuestionario anexo).

\*drgcauses: conocimientos general sobre uso de drogas (corresponden a esta variable compuesta las preguntas 36<sup>a</sup>, 36b, 36c, 36e,36f, 36h, 36i del cuestionario anexo).

\*drgcauses1: conocimiento de cambios de comportamiento y actitudes frente al uso de drogas (corresponden a esta variable compuesta las preguntas 36a, 36b, 36c, 36e,36f, 36h, 36i, 36j del cuestionario anexo).

.

\*drgcauses2: conocimiento de efectos colaterales y enfermedades causados por las drogas (corresponden a esta variable compuesta las preguntas 38a, 38b, 38c, 38d,36f, 38e, 38f, 38g del cuestionario anexo).

.

\*sexknow1: conocimientos generales sobre practicas sexuales (corresponden a esta variable compuesta las preguntas 54a, 54b, 54d, 54f, 54g, 54h del cuestionario anexo).

.

\*sexHIV: conocimientos y actitudes sobre prevención y transmisión del HIV/SIDA (corresponden a esta variable compuesta las preguntas 57b, 57d, 59a, 59b del cuestionario anexo).

.

\*AFOclass: Valoración de beneficios del programa en cuanto a conocimientos globales adquiridos (corresponden a esta variable compuesta las preguntas 80a, 80b, 80c, 80d, 80e del cuestionario anexo).

Cada una de estas variables corresponde a alguno de los constructos previamente mencionados:

Percepción	Conocimientos	Comportamiento	Actitudes
AFOclass	drgcauses drgcauses1 drgcauses2 sexknow1 sexHIV	drgindv	drgcauses1 sexHIV

### **Variables independientes del modelo**

\*schlcode. Grupo de estudio: control o intervención.

\*age: edad del individuo entrevistado (corresponde a la pregunta 3 del cuestionario anexo).

\*male: sexo del entrevistado (corresponde a la pregunta 2 del cuestionario anexo).

\*grade: grado en el que se encuentra el entrevistado: 7º grado y 8º grado.

\*livMa: vive con la madre (corresponde a la pregunta 8 del cuestionario anexo).

\*livFa: vive con el padre (corresponde a la pregunta 8 del cuestionario anexo).

\*livoth: vive con alguien más (corresponde a la pregunta 8 del cuestionario anexo).

\*misschF: indicador del nivel socio-económico (corresponde a la pregunta 7 del cuestionario anexo).

\*sch1-sch16: escuelas en estudio (representa la variable dummy que se creó para medir el efecto individual de cada escuela; para el grupo del 2000 la escuela control fue la escuela 6 y para las del 99 la escuela control fue la número 16).

Estas últimas son similares a las variables independientes que se utilizaron en Tulane, a excepción de “schlcode” que es utilizada en la Metaevaluación, pues se considera que representa mejor el impacto del programa diferenciándolo entre las escuelas control y las de intervención.

Para hacer la comparación de los grupos a través del tiempo, se utilizó el reporte de la línea de base que se elaboró con la información de las escuelas participantes antes de que iniciara la implementación del programa.

Brindar una contribución metodológica basado en los principios de la Teoría Psicométrica (Jum Nunnally e Ira Bernstein , 1995 y Figueredo, 2000) amplía las posibilidades en el área

de Evaluación y Metaevaluación de aplicar otras técnicas de análisis, obtención e interpretación de resultados.

Para finalizar el proyecto, se plantean algunas lecciones aprendidas de esta Evaluación de Impacto basándome en las diferentes etapas, pasos y criterios de una Metaevaluación y fundamentándolas en algunos de los estudios realizados por Scriven y Stufflebeam (ambos padres de las Teorías de evaluación de Programas y Proyectos Sociales como se mencionó en los referentes teóricos) y que buscan servir como base para futuros Evaluadores u otras organizaciones interesadas, así como facilitar la planeación y la implementación de programas de Salud con resultados de impacto satisfactorios, tanto para los patrocinadores como para los beneficiarios.

## **VII. Limitaciones del estudio**

Dentro de las grandes limitaciones con las que se contó durante el proceso de la Metaevaluación, se pueden mencionar:

### **a. Durante la Metaevaluación:**

- La poca factibilidad de comunicación con el personal involucrado en la planeación e implementación del programa, redujo la relación de cooperación que se debe tener entre patrocinadores y equipo evaluador. Además que me limitó el acceso a otro tipo de información que no fuese la base de datos con la que ya contaba.

### **b. Durante la Evaluación:**

#### **En cuanto al diseño del programa**

- **Ausencia de Valoración de necesidades antes del inicio de la implementación del programa.** Hallazgo de vital importancia puesto que para tener mayores posibilidades de lograr una intervención exitosa en la población meta, el programa debe dirigirse a las necesidades de esa población y debe responder a las circunstancias y al contexto social y cultural de aquellos en necesidad.

Recordemos que la efectividad de un programa que pretende incrementar las condiciones sociales, estará fuertemente determinada por la medida en que el mismo este dirigido

realmente a las necesidades sociales y que lo haga de una manera que responda a las circunstancias de aquellos en necesidad (Rossi et al, 1999).

Para esto se requiere primeramente una descripción del problema que el programa intenta aminorar. Este diagnóstico del problema y de la población objetivo debe ser buena y convincentemente establecida y en él debe desarrollarse un proceso de comparación y priorización de las necesidades, de acuerdo a la seriedad e importancia de las mismas.

En este contexto se concibe la noción de que una “necesidad” es una construcción social negociada entre un conjunto de agentes con responsabilidad para programas sociales y políticas, así como un conjunto de personas que manifiestan la condición y que certifican que el problema existe y requiere intervención. Sin embargo el hecho de que los individuos constituyan una población necesitada, en base a un criterio que representa una construcción social de necesidad usada por un programa en un contexto político, no significa necesariamente que ellos requieran el servicio. El deseo de servicios o la voluntad de participar, es lo que se llama demanda de servicios y usualmente la población de demanda es solamente una porción de la población necesitada (Rossi et al, 1999).

Al conjunto de procedimientos para describir y diagnosticar necesidades sociales es a lo que se llama “**valoración de necesidades**”, y es uno de los elementos que este trabajo pretende resaltar.

Si recordamos como se dijo en la justificación de este proyecto, otro aspecto importante es la presencia de una conceptualización o “teoría del programa”. Entendida ésta como el establecimiento de las relaciones entre los insumos, actividades y resultados intermedios del eventual programa y la identificación de una concepción de lo que debe hacerse o que debe intervenir para lograr los beneficios sociales que se intentan, es decir, como se supone que el programa logrará los resultados deseados (el impacto).

Y en nuestro caso en particular esta fue una de las grandes limitaciones.

- **Ausencia de una Teoría del Programa** de acuerdo con las personas que lo realizaron o las metas de la Organización que lo implementó. Esta teoría que tendría que haberse formulado explícitamente antes de la implementación del programa estableciendo las relaciones entre los insumos, actividades y resultados intermedios esperados del eventual programa, la identificación de

una concepción de lo que debería hacerse o en que aspectos debería intervenir para lograr los beneficios que se pretenden alcanzar, es decir, saber cómo se supone que el programa logrará los resultados deseados, y la definición de énfasis específicos en determinados temas relacionados con el uso y abuso de drogas y salud sexual y reproductiva, para este caso en particular.

Sin esta teoría se hace más difícil emitir juicios e interpretaciones de resultados en una Evaluación de Impacto, así como igualmente difícil es hacer un análisis sustantivo de los resultados obtenidos a través de las metodologías estadísticas (utilizadas tanto en Tulane como en Costa Rica) sin caer en la posibilidad de generar conjeturas o suposiciones no muy bien fundamentadas.

El otro aspecto importante es la articulación de una conceptualización o “teoría del programa”, el cual consta de dos partes, a saber: el establecimiento de las relaciones entre los insumos, actividades y resultados intermedios del eventual programa (Hipótesis de acción), y la identificación de una concepción de lo que debe hacerse o que debe intervenir para lograr los beneficios sociales que se intentan, es decir, saber cómo se supone que el programa logrará los resultados deseados (Hipótesis conceptual) (Rossi et al, 1999).

**c. En cuanto el diseño del Instrumento.**

- No existió validación empírica piloto del instrumento una vez este fue elaborado, es decir que faltaron procedimientos psicométricos para evidenciar la validez y confiabilidad de los ítems que representan cada una de las dimensiones o constructos en el cuestionario. Específicamente me refiero a la discriminación de ítems (correlación entre puntaje del ítem y el puntaje total en la escala, 0.30 o más) y alpha de Cronbach (estimación de confiabilidad para toda la escala, 0.70 para efectos de investigación y 0.9 para toma de decisiones), ambas consideradas herramientas útiles dada la naturaleza del programa y las dimensiones que se pretendían medir en el cuestionario.

## VIII. Resultados y discusión

Considero prudente iniciar la presentación de los resultados, ilustrando la construcción de las variables compuestas después de haber elegido los ítems que correspondían a los constructos definidos.

En los siguientes cuadros se muestran los ítems que conformaron las escalas, su dirección modificada, sus respectivos índices de discriminación y valores de alpha de Cronbach (sólo se presentan los ítems significativos después de depurados).

### Importante notar el valor de Alpha de Cronbach para la variable compuesta.

- ❖ Variables dicotómicas consideradas significativas en la construcción de la medida compuesta (**drvindv**) que se usó como indicador del constructo “uso de drogas”.

Item	Obs	Signo	correlación ítem-test	correlación ítem con el resto	alpha
Mganjtl	3983	+	0.6460	0.3987	0.6380
Mganjal	3982	+	0.7445	0.5462	0.5659
Mcigar1	3963	+	0.7170	0.5011	0.5925
Malcoll	3969	+	0.6612	0.4268	0.6222
Mcrack1	3977	+	0.5897	0.3058	0.6806
alpha total de la escala					<b>0.6729</b>

- ❖ Variables dicotómicas consideradas significativas en la construcción de la medida compuesta (**sexknow1**) que se usó como indicador del constructo “conocimientos generales sobre practicas sexuales”.

Item	Obs	Signo	correlación ítem-test	correlación ítem con el resto	alpha
np_water	3909	+	0.6222	0.3933	0.6409
np_first	3937	+	0.5925	0.3567	0.6540
np_jump	3876	+	0.6950	0.5019	0.6053
np_wee	3858	+	0.6835	0.4880	0.6105
np_stdup	3871	+	0.6435	0.4309	0.6297
bcnt_HIV	3815	+	0.4989	0.2495	0.6903
alpha total de la escala					<b>0.6807</b>

- ❖ Variables dicotómicas consideradas significativas en la construcción de la medida compuesta (**sexHIV**) que se usó como indicador del constructo “conocimientos y actitudes sobre prevención y transmisión del HIV/SIDA”.

Item	Obs	Signo	correlación ítem-test	correlación ítem con el resto	alpha
hiv_know	3790	+	0.6989	0.3864	0.3790
hiv_hlth	3761	+	0.6937	0.3791	0.3873
myageHIV	3853	+	0.6045	0.2362	0.5098
virgHIV	3763	+	0.5985	0.2373	0.5097
alpha de la escala					<b>0.5219</b>

A pesar que en la construcción de esta escala los índices de discriminación y el valor de alpha para la variable compuesta están por debajo de los definidos en la metodología, se consideró la mejor escala para representar este constructo dentro del cuestionario.

Estos valores bajos se pueden deber en gran medida al escaso número de ítems que componen esta escala.

- ❖ Variables dicotómicas consideradas significativas en la construcción de la medida compuesta (**drgcauses**) que se usó como indicador del constructo “conocimientos generales sobre uso de drogas”.

Item	Obs	Signo	correlación ítem-test	correlación ítem con el resto	alpha
ganjhall	3858	+	0.5401	0.3291	0.6970
smoke2nd	3933	+	0.6354	0.4500	0.6665
ganjsch	3931	+	0.6441	0.4627	0.6628
crackadd	3851	+	0.6669	0.4960	0.6546
beeradd	3913	+	0.6302	0.4464	0.6681
prescadd	3905	+	0.5820	0.3848	0.6840
drughiv	3920	+	0.5437	0.3360	0.6956
alpha de la escala					<b>0.7088</b>

- ❖ Variables dicotómicas consideradas significativas en la construcción de la medida compuesta (**drgcauses1**) que se usó como indicador del constructo “conocimientos de cambios de comportamiento y actitudes frente al uso de drogas”.

Item	Obs	Signo	correlación ítem-test	correlación ítem con el resto	alpha
ganjhall	3858	+	0.5145	0.3245	0.7287
smoke2nd	3933	+	0.6368	0.4766	0.6992
ganjsch	3931	+	0.6566	0.5032	0.6935
crackadd	3851	+	0.6533	0.5010	0.6941
beeradd	3913	+	0.6172	0.4544	0.7042
prescadd	3905	+	0.5705	0.3965	0.7156
drughiv	3920	+	0.5136	0.3269	0.7285
rightfrn	3933	+	0.6000	0.4304	0.7088
alpha de la escala					<b>0.7362</b>

- ❖ Variables dicotómicas consideradas significativas en la construcción de la medida compuesta (**drgcauses2**) que se usó como indicador del constructo “conocimiento de efectos colaterales y enfermedades causados por las drogas”.

Item	Obs	Signo	correlación ítem-test	correlación ítem con el resto	alpha
harm_cig	4007	+	0.8063	0.7198	0.8896
harm_alc	3964	+	0.8031	0.7210	0.8895
harm_gan	3886	+	0.7988	0.7138	0.8903
harm_gnt	3870	+	0.6172	0.4826	0.9149
harm_spf	3947	+	0.8215	0.7470	0.8864
harm_cok	3955	+	0.8687	0.8108	0.8791
harm_crk	3944	+	0.8736	0.8182	0.8783
alpha de la escala					<b>0.9044</b>

- ❖ Variables dicotómicas consideradas significativas en la construcción de la medida compuesta (**AFOclass**) que se usó como indicador del constructo “valoración de beneficios del programa en cuanto a conocimientos adquiridos”.

Item	Obs	Signo	correlación ítem-test	correlación ítem con el resto	alpha
AFO_drg1	687	+	0.8276	0.7164	0.8498
AFO_viol	680	+	0.8117	0.6940	0.8557
AFO_alcl	689	+	0.8381	0.7322	0.8451
AFO_sex1	688	+	0.8205	0.7065	0.8517
AFO_hiv1	692	+	0.8074	0.6842	0.8575
alpha de la escala					<b>0.8780</b>

Como se había mencionado este paso se obvió durante la fase de evaluación de impacto, en donde se utilizaron como variables dependientes cada una de las variables dicotómicas del cuestionario.

Para la construcción del indicador (variable compuesta), primero se procedió a estandarizar cada una de las variables antes mencionadas, y luego se calculó el promedio de estos valores estandarizados para cada individuo. La estandarización de las variables elimina el efecto de las diferentes unidades de medida y diferentes variancias en el cálculo de la medida compuesta, de tal manera que todas tienen la misma ponderación o peso. (Figueredo et. al.; 2000)

Este indicador, así construido, tiene las ventajas de no ser una variable dicotómica, utiliza otros ítems del cuestionario que miden el mismo constructo que se esta evaluando e involucra a todos los casos en la muestra. Se soluciona así el problemas de los valores faltantes. (Figueredo et. al.; 2000).

Utilizando este mismo razonamiento, se fueron construyendo cada una de los indicadores. Antes de iniciar propiamente con la descripción de los resultados, es necesario hacer una breve recapitulación de lo que fue la línea de base, con el fin de sustentar que tanto las escuelas control e intervención se encontraban en condiciones similares en lo referente a conocimientos, actitudes y comportamiento respecto a la salud sexual y reproductiva y al uso/abuso de drogas. Y que inclusive por diferenciación de grado escolar no se observan diferencias sustancialmente significativas, en la mayoría de las dimensiones evaluadas. Lo que nos lleva a pensar con mayor certeza que los resultados obtenidos con esta nueva metodología, son debidos al impacto generado por el programa. Los cuadros que se presentan a continuación son indicativos de la equivalencia inicial de los grupos en estudio (control e intervención) y de esta forma se da garantía de que las diferencias observadas en los resultados no son debidos a diferencias iniciales en los grupos, sino a los efectos propios del programa.

Cuadro 1. Uso de drogas por tipo de escuela reportado en línea de base

	Escuelas Intervención %	Escuelas control %
Alcohol	35.4	34.7
Cigarros	15.2	15.4
Té de Marihuana	14.5	9.9
Cigarros de Marihuana	11.2	11.4
Inhalantes	5.8	2.8
Crack/cocaína	3.7	1.7

Fuente: AAAP evaluation: Baseline report, Nov. 1999

Al analizar los resultados del reporte base según el tipo de escuela, el uso individual de drogas como el alcohol, cigarros y marihuana fue similar en las escuelas de intervención y las de control, pero significativamente diferentes para el uso de té de marihuana, inhalantes y crack / cocaína, en cada uno de los casos las diferencias significativas fueron mayores para las escuelas de intervención que las escuelas de control.

Cuadro2. Conocimientos adecuados sobre prácticas sexuales según tipo de escuelas y grado

	Escuelas intervención (7°) %	Escuelas Control (7°) %
Si una chica salta de arriba abajo o baila después de tener relaciones sexuales, ella no puede quedar embarazada (falso)	59.6	53.9
Una chica no puede quedar embarazada la primera vez que tenga relaciones sexuales (falso)	54.0	57.4
Si la chica orina después de tener relaciones sexuales no quedará embarazada (falsa)	52.6	5.9
Una chica puede quedar embarazada cuando tiene relaciones sexuales en el agua (cierto)	57.3	53.3
Una chica puede quedar embarazada cuando tiene relaciones sexuales de pie (cierto)	54.9	44.5
Cuando un chico eyacula fuera de la chica, ésta no puede quedar embarazada (falso)	47.0	41.0
No todos los métodos de planificación familiar previenen el HIV (cierto)	44.9	39.1
Usar Pepsi o Coca-Cola después de tener relaciones sexuales no previenen de un embarazo (cierto)	34.2	30.5
	Escuelas intervención (8°)	Escuelas Control (8°)
Si una chica salta de arriba abajo o baila después de tener relaciones sexuales, ella no puede quedar embarazada (falso)	72.3	70.1
Una chica no puede quedar embarazada la primera vez que tenga relaciones sexuales (falso)	65.9	72.4
Si la chica orina después de tener relaciones sexuales no quedará embarazada (falsa)	70.6	62.6
Una chica puede quedar embarazada cuando tiene relaciones sexuales en el agua (cierto)	62.9	53.9
Una chica puede quedar embarazada cuando tiene relaciones sexuales de pie (cierto)	61.2	52.8
Cuando un chico eyacula fuera de la chica, ésta no puede quedar embarazada (falso)	38.2	32.3
No todos los métodos de planificación familiar previenen el HIV (cierto)	43.5	41.2
Usar Pepsi o Coca-Cola después de tener relaciones sexuales no previenen de un embarazo (cierto)	30.9	27.4

Fuente: AAAP evaluation: Baseline report, Nov. 1999

El conocimiento correcto en algunos de los indicadores sobre prácticas sexuales fue generalmente mayor en los estudiantes de 8° grado que en los de 7° tanto por tipo de escuela (control o intervención) como por sexo del entrevistado. En forma global el nivel de conocimiento adecuado en los muchachos fue significativamente mayor que el nivel que poseen las muchachas.

Cuadro3. Conocimientos sobre prevención y transmisión del HIV/SIDA por tipo de escuela y grado

	Escuelas intervención	Escuelas control
	7° grado	7° grado
	%	%
Alguien de mi edad no puede contraer HIV/SIDA (falso)	64.8	65.5
Tener relaciones sexuales con una persona virgen previene contraer HIV/SIDA (falso)	67.3	69.8
El HIV se previene utilizando condom siempre (si)	79.1	78.8
El HIV se previene solo teniendo relaciones sexuales con personas conocidas (si)	41.5	57.4
El HIV se previene con la abstinencia sexual (si)	50.5	53.8
El HIV se previene teniendo relaciones sexuales con personas que lucen saludables (no)	53.7	55.6

  

	Escuelas intervención	Escuelas control
	8° grado	8° grado
	%	%
Alguien de mi edad no puede contraer HIV/SIDA (falso)	73.2	76.6
Tener relaciones sexuales con una persona virgen previene contraer HIV/SIDA (falso)	82.0	82.0
El HIV se previene utilizando condón siempre (si)	85.9	84.2
El HIV se previene solo teniendo relaciones sexuales con personas conocidas (si)	37.4	33.6
El HIV se previene con la abstinencia sexual (si)	62.3	58.7
El HIV se previene teniendo relaciones sexuales con personas que lucen saludables (no)	56.0	61.2

Fuente: AAAP evaluation: Baseline report, Nov. 1999

Cuadro4. Actitudes frente a la prevención y transmisión del HIV/SIDA por tipo de escuela y grado

	Escuelas intervención	Escuelas control
	7° grado	7° grado
	%	%
Alguna vez han escuchado sobre HIV/SIDA	74.7	82.5
Rechazo de ser amigo de alguien que tenga HIV/SIDA	47.6	49.0

  

	Escuelas intervención	Escuelas control
	8° grado	8° grado
	%	%
Alguna vez han escuchado sobre HIV/SIDA	85.7	89.8
Rechazo de ser amigo de alguien que tenga HIV/SIDA	60.0	62.8

Fuente: AAAP evaluation: Baseline report, Nov. 1999

A pesar que la mayoría de muchachos y muchachas reportaron haber escuchado sobre la enfermedad del SIDA, los conocimientos fueron significativamente mayores en los estudiantes de 8° grado. Y de los de 7° grado, los estudiantes de las escuelas control dijeron haber escuchado sobre SIDA.

Sin embargo estudiantes de ambos grados mostraron adecuado conocimiento, independiente del tipo de escuela (intervención o control).

La mayoría mostró un entendimiento apropiado respecto a la forma de transmisión y prevención del SIDA. Parecería que el conocimiento de reducción de riesgos iba en aumento mientras mayor es el nivel académico del entrevistado.

El temor de contraer la enfermedad se presenta bastante similar independientemente del grado y del tipo de escuela.

Cuadro5. Conocimientos generales sobre el uso de drogas según tipo de escuela y grado

Drogas dañinas para la salud	Escuelas intervención	Escuelas control
	7° grado %	7° grado %
Inhalantes	63.9	80.3
Cocaína	58.0	66.7
Crack	56.0	63.9
Alcohol	56.5	60.7
Cigarros	54.8	62.9
Marihuana	53.4	64.6
Mezcla sazónada	49.3	59.0
Té de marihuana	39.7	44.2

  

Drogas dañinas para la salud	Escuelas intervención	Escuelas control
	8° grado %	8° grado %
Inhalantes	79.4	80.5
Cocaína	79.6	80.3
Crack	77.2	78.6
Alcohol	74.4	76.8
Cigarros	73.7	72.0
Marihuana	72.9	76.3
Mezcla sazónada	68.9	71.4
Té de marihuana	53.3	52.2

Fuente: AAAP evaluation: Baseline report, Nov. 1999

El conocimiento acerca de lo dañino que pueden ser algunas drogas específicas aumenta significativamente entre los grados, lo cual se demuestra al observar los porcentajes para los estudiantes de 8° grado en comparación con los de 7° grado. Sin embargo para este

mismo grado (8° grado) no hubieron diferencias significativas entre las escuelas intervención y las escuelas control.

Cuadro 6. Conocimientos acerca de cambios en comportamientos por uso de drogas y actitudes frente a dicho uso, según tipo de escuela y grado

	Escuelas intervención 7° grado %	Escuelas control 7° grado %
Escoger un buen grupo de amigos ayuda a estar lejos del consumo de drogas (cierto)	63.2	66.1
Fumar marihuana puede hacerte perder interés por la escuela (cierto)	66.6	66.7
Fumar marihuana puede volverte loco (cierto)	62.8	66.8
Respirar el humo de otro que fuma cigarrillos puede causarte cáncer de Pulmón (cierto)	54.5	55.9
El crack/cocaína son adictivos (cierto)	46.7	51.0
La cerveza y el vino son tan adictivos como el alcohol puro (cierto)	46.0	47.4
Puedes convertirte en un adicto con medicinas prescritas por un médico (cierto)	42.9	37.1
La adicción a las drogas aumenta el riesgo de contraer HIV/SIDA (cierto)	30.0	31.3
Las personas manejan mejor sus problemas cuando usan drogas/alcohol (falso)	39.0	20.1
Fumar marihuana no hace a las personas ver o escuchar cosas que no son reales (falso)	31.1	29.1
La mezcla sazónada es marihuana combinada con tomillo (falso)	14.4	14.6
	Escuelas intervención 8° grado %	Escuelas control 8° grado %
Escoger un buen grupo de amigos ayuda a estar lejos del consumo de drogas (cierto)	78.8	77.7
Fumar marihuana puede hacerte perder interés por la escuela (cierto)	74.2	69.1
Fumar marihuana puede volverte loco (cierto)	66.3	61.7
<b>Respirar el humo de otro que fuma cigarrillos puede causarte cáncer de Pulmón (cierto)</b>	<b>67.0</b>	<b>60.0</b>
El crack/cocaína son adictivos (cierto)	63.2	57.2
La cerveza y el vino son tan adictivos como el alcohol puro (cierto)	55.4	51.4
Puedes convertirte en un adicto con medicinas prescritas por un médico (cierto)	42.2	38.1
La adicción a las drogas aumenta el riesgo de contraer HIV/SIDA (cierto)	27.4	22.1
<b>Las personas manejan mejor sus problemas cuando usan drogas/alcohol (falso)</b>	<b>48.1</b>	<b>14.2</b>
Fumar marihuana no hace a las personas ver o escuchar cosas que no son reales (falso)	30.6	29.1
<b>La mezcla sazónada es marihuana combinada con tomillo (falso)</b>	<b>23.7</b>	<b>14.4</b>

Fuente: AAAP evaluation: Baseline report, Nov. 1999

Los muchachos de 7° grado de las escuelas de intervención tienden a mostrar el mismo nivel de conocimientos apropiados que los de las escuelas control, y en algunos casos

menor; caso contrario con lo que sucede para los muchachos de 8° grado de las escuelas de intervención los cuales demuestran una cierta tendencia a mayores conocimientos en algunos de los indicadores (representados en “negritas” en el cuadro anterior) que los de las escuelas control correspondientes a su mismo nivel.

En las actitudes hacia el uso de drogas, se observan diferencias significativas entre grupos de edades, independientemente del tipo de escuelas. Las diferencias en actitudes por grupos de edad fueron significativas en casi todas las preguntas. Lo cual sugiere que las actitudes frente al uso de drogas varían naturalmente de acuerdo con la edad del individuo (representado por el grado escolar cursado).

Las diferencias significativas observadas entre el grupo intervención y grupo control para 8° grado son:

- Inhalar el humo de cigarrillo de otra persona puede causar cáncer en el pulmón.
- Fumar marihuana puede hacerte perder el interés por la escuela.
- La mezcla sazónada es marihuana mezclada con tomillo.
- La adicción por la droga aumenta el riesgo de contraer HIV/SIDA.

#### META-EVALUACION

Los resultados obtenidos en la Metaevaluación se presentarán de acuerdo a cada uno de los constructos definidos previamente y en algunos de ellos se harán las comparaciones con los resultados obtenidos durante la Evaluación, con el fin de demostrar diferencias significativas cuando se utilizan diferentes métodos para analizar datos e interpretar resultados.

Tomando en cuenta que al utilizar *variables dicotómicas* como *variables dependientes* en los modelos de regresión, se da pie a que se realicen análisis con poca potencia estadística, ósea, altos errores de medición; y además se obtienen rangos restringidos de variabilidad (ejem: variable dicotómica por su naturaleza presenta una tendencia hacia la baja variabilidad pues sólo toma dos posibles valores), durante la Metaevaluación se cambió la metodología estadística (regresiones lineales con análisis multinivel) utilizando *escalas* (medidas compuestas) como *variables dependientes* en los modelos y se obtuvo mayor poder estadístico en el análisis y algunos resultados más claramente interpretables.

**Análisis Multivariable:** además de la comprobación de los cambios existentes en los indicadores a través del tiempo, se considera la posibilidad de controlar los cambios de

acuerdo a las características de la muestra, para lo cual se plantea un nuevo modelo de regresión lineal (diferente al utilizado en Tulane), pero esta vez que incluya diferentes niveles, y así se controlarán los indicadores o variables de acuerdo al sexo, la edad del participante, el grado cursado en ese momento, estructura familiar del alumno (si vive con ambos padres, solo con uno de ellos o alguien diferente a sus padres) y un indicador del estado socioeconómico del entrevistado (si presentó algún tipo de dificultad económica que lo hiciera faltar a la escuela).

Se introduce en este modelo una variable “dummy” para cada una de las escuelas ante la sospecha que existieran cambios significativos entre ellas, a pesar que presentaban características similares.

Algunos de los resultados más relevantes según los diferentes constructos definidos y correspondientes a escuelas que iniciaron el programa en el año 2000 son:

## Comportamiento en cuanto a uso de drogas

El rango de la variable compuesta es: de 0 a 6.

### Metaevaluación, Costa Rica:

drgindv	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
schlcode	<b>-.2363078</b>	.0045969	<b>-51.41</b>	<b>0.000</b>	<b>-.2481244</b>	<b>-.2244911</b>

### Evaluación de Impacto, Tulane:

Mganjt1	Coef. OR	Std. Err.	t	P> t	[95% Conf. Interval]	
schlcode	<b>1.373665</b>	.2063953	<b>2.113</b>	<b>0.088</b>	<b>.9335583</b>	<b>2.02125</b>

En este constructo, para el grupo intervenido es significativamente menor el uso de drogas (dirección esperada) y la variable que los representa es *altamente significativa* ( $p= 0.000$ ) en el modelo con la medida compuesta, a diferencia del modelo de regresión con la variable dicotómica.

Podemos observar por ejemplo el valor del coeficiente  $-.2363078$ , el cual nos indica que los alumnos de las escuelas intervenidas en el año 2000 tienen 0.23 puntos menos de consumir drogas que los que están en las escuelas control en ese mismo año (dirección o impacto esperado). Valor que se obtiene utilizando la medida compuesta. Al contrario de lo que

sucede cuando se utiliza un modelo de regresión con una de las variables dicotómicas, pertenecientes a este mismo constructo, en el cual el valor del coeficiente OR (1.373665) nos indica todo lo contrario, es decir, que los alumnos de las escuelas de intervención presentan 1.37 más chances de consumir drogas que los de las escuelas de control, y no es un resultado estadísticamente significativo ( $p= 0.088$ ) a diferencia del obtenido en la Metaevaluación ( $p= 0.000$ ), deducción no hay impacto.

Lo mismo sucede para el valor de “t” que representa la significancia estadística de la diferencia entre el grupo intervención y el grupo control y tiene un valor negativo por lo que va en la dirección esperada (disminución de uso de drogas para los intervenidos).

## **Conocimientos sobre prácticas sexuales**

El rango de la variable compuesta es: de 0 a 6.

### **Metaevaluación, Costa Rica:**

sexknow1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
schlcode	<b>.1876269</b>	.0161883	<b>11.59</b>	<b>0.000</b>	<b>.1460135</b>	<b>.2292404</b>

### **Evaluación de Impacto, Tulane:**

Np_jump	Coef. OR	Std. Err.	t	P> t	[95% Conf. Interval]	
schlcode	<b>1.695514</b>	.2055851	<b>-1.210</b>	<b>0.257</b>	<b>.3615237</b>	<b>1.361297</b>

Se demuestra que el grupo de intervención presenta una diferencia positiva y altamente significativa ( $t = 11.59$  y  $p = 0.000$  respectivamente) con respecto al grupo de control en la dimensión “Conocimientos sobre prácticas sexuales”. Al utilizar medidas compuestas se observó que se representan mejor las dimensiones que “hipotéticamente” pretendía medir el programa, ya que a través de medidas compuestas (escalas) se miden los *constructos* (formados por varias variables dicotómicas pertenecientes al mismo constructo), ejem: conocimientos, actitudes y comportamientos ante el uso y abuso de drogas así como de salud sexual y reproductiva, lo que no sucede con variables dicotómicas.

## Conocimientos generales sobre uso de drogas

El rango de la variable compuesta es: de 0 a 7.

### Metaevaluación, Costa Rica:

drgcauses	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
schlcode	<b>.3013139</b>	.274678	<b>10.97</b>	<b>0.000</b>	<b>.2307057 .3719221</b>

### Evaluación de Impacto, Tulane:

Smoke2nd	Coef. OR	Std. Err.	t	P> t	[95% Conf. Interval]
schlcode	<b>1.685938</b>	.3043497	<b>2.893</b>	<b>0.018</b>	<b>1.120699 2.536265</b>

Trabajar con variables compuestas (escalas) nos brindó mayor precisión para la estimación del efecto del programa y de las otras *variables independientes* en los modelos. Ejem: los rangos entre intervalos de confianza son mas cortos y precisos que si se trabaja con variables dicotómicas. Observemos el rango del intervalo en la variable compuesta drgcauses (.2307057, .3719221), a diferencia del mismo rango para la variable dicotómica que se utilizo en Tulane Smoke2nd (1.120699, 2.536265).

## Conocimientos de efectos colaterales del uso de drogas

El rango de la variable compuesta es: de 0 a 7.

### Metaevaluación, Costa Rica:

drgcauses2	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
schlcode	<b>-.483769</b>	.0235119	<b>-20.58</b>	<b>0.000</b>	<b>-.5442083 -.4233298</b>

Al construir la variable compuesta para este constructo, el resultado es diferente al de Tulane, mostrando que el grupo de intervención tiene menos conocimiento en esta área que los grupo control (dirección contraría a la lógica). Observemos el coeficiente = -.483769, lo cual indica que los estudiantes de las escuelas intervención tienen .48 puntos menos en la escala de conocimiento sobre los efectos colaterales del uso de drogas, que los estudiantes de las escuelas control, al igual que el valor de “t” nos resultó con signo negativo lo cual indica que hay diferencia en dirección contraria entre el grupo intervención y el control (disminución de conocimientos para el primer grupo) y un valor de “p” altamente significativo.

### **Evaluación de Impacto, Tulane:**

De las ocho variables dicotómicas (8 modelos) que se utilizaron para medir el impacto del programa en esta dimensión, ninguna resultó ser significativa y según el análisis descriptivo los valores de “t” observados tienen signos diferentes: 5 son negativos y 3 positivos.

Resultado que llama la atención, y que difícilmente le podemos encontrar explicación, sin caer en conjeturas, puesto que podríamos pensar que el hecho que los alumnos de las escuelas intervención tengan menor conocimiento en esta área es debido a que el programa no hacía mucho énfasis en este tipo de problemas o que incluso ni siquiera estaba contemplado dentro de los temas que aborda el programa. Lo cual se hubiese podido sustentar si se contara con la Teoría del Programa.

El análisis multivariado, con nuevos modelos de regresión que incluían variables compuestas, a diferencia de las dicotómicas (Tulane) fue necesario para asegurarnos que los resultados reportados durante la Evaluación de Impacto habían sido adecuadamente obtenidos e interpretados.

Los resultados generales del análisis se presentan en la tabla #1 (Anexos), en la cual se puede observar que los resultados significativos (valor de “p”= 0.05 o menos y el valor de “t” indicando la dirección de cambio esperada) se colocaron con “negrita” y los resultados significativos pero con dirección de cambio opuesta a la esperada se colocaron con “negrita” y en “cursiva”.

### **Conocimientos generales y Actitudes frente al uso y abuso de drogas:**

El aumento en el conocimiento general en cuanto uso y abuso de drogas (alcohol, cocaína, marihuana e inhalantes) así como cambios en actitudes parecen ser consistentes y significativos en las escuelas de intervención comparados con sus similares escuelas control, independientemente del tiempo que recibieron el programa. Es decir que los resultados son más fuertes y claramente interpretables para las variables que se utilizaron como indicadores de temas relacionados con drogas y alcohol tanto en las escuelas intervención como en las de control y tanto para un año como para el otro. Y si nos basamos en los resultados de la línea de base ( en los que se refleja que los conocimientos

en estos temas eran similares para todas las escuelas que incluiría el programa) podemos afirmar que el aumento de conocimientos en esta área, son debidos precisamente al Impacto del programa implementado.

Situación que cambia cuando analizamos los resultados obtenidos para la variable de conocimientos de efectos colaterales ( enfermedades causadas por las drogas), en los cuales observamos que el impacto del programa fue negativo en las escuelas de intervención del 2000, pues los datos reflejan que hubo disminución de los conocimientos en esos estudiantes y dicha disminución es significativa (dirección opuesta a la esperada) comparada con las escuelas control para ese mismo año. Y para las escuelas que inician el programa en 1999 ni siquiera resulta ser una variable significativa.

### **Conocimientos generales sobre salud sexual y reproductiva y actitudes frente al HIV/SIDA:**

Caso contrario, ocurre en los conocimientos sobre salud sexual y reproductiva, en el cual se puede observar que para los alumnos de las escuelas de intervención del 2000 hay aumento del conocimiento y éste es significativo (dirección esperada después de implementar un programa) a diferencia de lo que se refleja en las escuelas del 99, en las que se observa que el conocimiento es menor en las escuelas de intervención comparadas con las escuelas control, y esta disminución o falta de cambio es significativo.

En lo referente a conocimiento y actitudes frente al HIV/SIDA llama la atención como para las escuelas de intervención del 99 ni siquiera resultó ser una variable significativa y para las del 2000 es significativa pero muestra una dirección opuesta a la esperada, pues se observa que las escuelas de intervención poseen menos conocimiento acerca de este tema que las control o en su defecto no han variado con respecto al inicio de la implementación del programa. Esto puede deberse a un problema de falta de confiabilidad de la escala, pues recordemos que la constituyen pocos ítemes.

### **Uso individual de drogas (Comportamiento):**

En lo referente al comportamiento ante el uso de drogas a nivel individual, el programa tuvo un gran impacto pero llama la atención que sólo para las escuelas de intervención del

año 2000, en los cuales se puede observar que hay disminución del uso de drogas y que dicha disminución es significativa, lo cual está dentro de la dirección esperada. Situación que varía para las escuelas del 99, en las que los datos reflejan aumento en el uso de drogas para los estudiantes de las escuelas intervención y extrañamente ese aumento resultó significativo.

En este caso también es importante remitirse a los resultados de la línea de base, los cuales nos sirven de comparación y de sustento para poder decir, que si ambos grupos, tanto intervención como control, parten del mismo estadio de comportamiento, para las escuelas de intervención que reflejan una disminución en su uso se puede decir que es debido a los efectos del programa, y para las que reflejan igual uso o en su defecto aumento en el mismo habría que crear hipótesis (conjeturas) puesto que no se cuenta con una Teoría del Programa ni una contextualización de la valoración de necesidades que refuerce o nos permita concluir en una ineficiencia del programa, como se planteó en Tulane.

En resumen los resultados obtenidos en el análisis multivariable reflejan un completo efecto positivo del programa para las dimensiones de conocimiento, actitudes y comportamiento positivos con respecto al uso de alcohol y drogas, en las cuales se observan resultados más fuertes y consistentes en varias de las salidas de STATA que se corrieron con el fin de la Metaevaluación, no así para las dimensiones o constructos denominados “conocimientos sobre prevención y transmisión del HIV/SIDA” y “conocimientos de efectos colaterales del uso de drogas”, en las cual se observaron resultados poco convincentes para catalogar un efecto positivo del programa.

## **Valoración del Programa**

En cuanto al último constructo, que se construyó durante la Metaevaluación, llama la atención cómo la escala (variable compuesta) que se utilizó para su medición obtiene un alto valor de alpha de Cronbach y al momento de correr el modelo de regresión podemos observar que en general los alumnos de las escuelas de intervención tienen 4.68 chances más (valor del coeficiente) de haber obtenido resultados positivos en sus conocimientos, comportamientos y actitudes con respecto al uso de drogas y salud sexual y reproductiva. Y es una variable estadísticamente significativa presentando un valor de  $P= 0.000$ .

## **IX. Conclusiones y lecciones aprendidas**

1. El hecho de no contar con una Valoración de necesidades previa, así como la ausencia de una Teoría del Programa, pasos importantes de cumplir durante el diseño del programa y tomando en cuenta que tanto el análisis cualitativo como los métodos estadísticos son indispensables en toda evaluación, constituyó una limitación para la Metaevaluación del programa, puesto que no se tenían claros cuáles eran los alcances que pretendía el programa y como los querían alcanzar, lo que dificultó a su vez la interpretación (sin caer en conjeturas) de algunos de los nuevos resultados.
2. La elaboración adecuada, validación previa y la prueba piloto de un instrumento, con el fin de medir su confiabilidad estadística (medir lo que pretende medir) y su calidad técnica es de suma importancia para asegurar que los resultados que se reflejan en él serán válidos para determinar cambios o impacto de un programa así como para realizar análisis sustantivos en la mayoría de las dimensiones que el programa pretende medir.
3. El contar con personal experimentado en el área de Evaluación durante la planeación e implementación del programa así como la selección adecuada del equipo de Evaluadores, permite obtener resultados favorables o de impacto y evaluaciones técnicamente adecuadas, útiles y eficaces, en lo referente al diseño del programa, elaboración de instrumentos, confiabilidad y validez de los mismos, estrategias de análisis de datos, interpretación de resultados y elaboración del reporte final.
4. El haber realizado una Metaevaluación permitió plantear un enfoque diferente al realizado en Tulane en cuanto al análisis, interpretación, explicación y la sistematización tanto del nivel teórico de la Evaluación como del práctico (metodológico), y así poder obtener resultados nuevos y diferentes a los ya existentes, y poder definir el impacto del programa con una mayor objetividad como positivo en uno de los grupos estudiados (escuelas de intervención del año 2000).

5. Utilizar una metodología estadística diferente a la de Tulane, basada en la Teoría Psicométrica, en el Principio de Agregación (regresiones lineales para variables compuestas métricas) permitió ampliar las posibilidades de aplicar otro tipo de métodos estadísticos, diferentes técnicas aplicadas de que le dieran mayor poder estadístico al análisis, obtención de resultados mas claramente interpretables.
6. Utilizar medidas compuestas dentro de los modelos de regresión lineal permitió representar mejor las dimensiones que “hipotéticamente” pretendía medir el programa, ya que a través de medidas compuestas (escalas) se midieron los *constructos* que fueron el resultado de la agregación de varias variables dicotómicas pertenecientes al mismo y nos brindó mayor precisión para la estimación del efecto del programa y de las otras *variables independientes* en los modelos, lo que no sucedió en Tulane, donde las mediciones se hicieron con variables dicotómicas.
7. El haber realizado esta Metaevaluación y aprovechar las lecciones que de ella se obtienen nos deja un gran desafío a los que pretendemos seguir en el área de la Evaluación de Proyectos y Programas Sociales, puesto que nos enseña que debemos mejorar la práctica de la Evaluación, tomando en cuenta los problemas que ésta pueda suponer, así como mejorar la calidad de los estudios evaluativos (tanto a nivel conceptual-cualitativo como a nivel cuantitativo) y nos conduce a retomar e incrementar responsabilidades como profesionales en la Evaluación

## X. Bibliografía

- Díaz, Luis Fernando. Metaevaluación: Texto para el Programa de Maestría en Evaluación de Programas y Proyectos Sociales. Ciudad Universitaria Rodrigo Facio. UCR, San José, Costa Rica; 2000; (Por editar).
- Enamorado, Ramón Antonio. Factores psicosociales asociados al riesgo de contraer una ETS o el VIH/SIDA en Chacarita, Puntarenas. Ciudad Universitaria Rodrigo Facio. UCR, San José, Costa Rica; 1995; (Tesis).
- Figueredo A.; McKnight P; McKnight K; Sidani s. Multivariate Modeling of Missing Data Within and Across Assessment Waves. Universidad de Arizona; 2000.
- Family Health Int/FOCUS; Hope Enterprises Ltd. AAAP Evaluation.Baseline Report. Kingston, Jamaica; 1999.
- González Castañón, Miguel Ángel. Estrategias Metodológicas y Criterios de Evaluación. Santa Fe de Bogotá: Fundación Alberto Merani; 1995.
- Herrera, Rafael. Pasos en la Conducción de una Evaluación. Chile: Ministerio de Educación Pública, Departamento de Evaluación, Mimeo; 1978.
- Magnani, Robert, et al. Resultados de la evaluación del Programa Antidrogas y Salud Sexual en Adolescentes de Jamaica. Universidad de Tulane, New Orleans, EEUU; 2002.
- Montero,Eiliana. Uso de la Teoría Psicométrica en la Generación de Indicadores de Impacto . San José, C.R.; 2001 (Clase magistral).
- Nilsson, N y Hogben, D. Metaevaluation en House (Ed) Philosophy of Evaluation. New Directions for Program Evaluation. San Francisco: Jossey-Bass; 1983.
- Nunnally, Jum y Bersnstein, Ira. Teoría Psicométrica. 3 ed. México, D.F.: Mc Graw - Hill; 1995; pp. 235-276.
- Rossi, P. H. ; Freeman H & Lipsey M. W. Evaluation: A Systematic Approach.: Sage Publicatios; 1999.
- Sanders, James R. Metaevaluation of The Effectiveness of Comprehensive, Case Management Interventions: Evidence From the National Evaluation of the Comprehensive Child Development Program. Stamford CT: The American Journal of Evaluation, V20, No 3; 1999.
- Scriven, Michael. The Logic of Evaluation and Evaluation Practice. En New Directions in Evaluations 68. San Francisco: Jossey-Bass; 1995.

- Stevens, Carla J y Dial, Micah (Eds). Preventing the Misuse of Evaluation. San Francisco: Jossey-Bass; 1994.
- Scott, Carol. Meta-Evaluation, What is it? How can we use it? Why should we use it?. Powerhouse Museum; 1998.
- Scott, Carol. Metaevaluación: “Evaluando otras evaluaciones”, Powerhouse Museum; 1998.
- Stockholm, W. Condensed Report on a Meta-Evaluation of The Evaluation of Diakonia and Forum Syd’s Cooperation with Cambodian NGOs; 2001.
- Stufflebeam, Daniel y Shinkfield, Anthony. Evaluación sistemática, Guía teórica y práctica. Barcelona: Paidós; 1993.
- Stufflebeam, Daniel. Metaevaluación: “Evaluación de una Evaluación de Impacto”. Universidad de Michigan; 1974.
- \_\_\_\_\_ Introduction to Meta-Evaluation: Achieving Effective Tourism Marketing Programs; 1970.
- UNAIDS; WHO; USAID; MEASURE Evaluation. National AIDS Programmes. A Guide to Monitoring and Evaluation.; 2000.

