

Evaluación del Proceso, de Resultados e Impacto

Silvia Ubillos Landa y Eider Goiburu Moreno



Universidad de Burgos



Lahia-Nahia Sexologia Elkarte

CAPACITANDO A LAS ONG/SIDA: MEJORANDO LA PREVENCIÓN DEL VIH

(Barcelona 3, 4 y 5 de Marzo del 2010)

Evaluación del Proceso, de Resultados e Impacto

Silvia Ubillos Landa¹ y Eider Goiburu Moreno²

¹ Universidad de Burgos; ² Asociación de Sexología Lahia-Nahia

Promocionar la evaluación es esencial para mejorar las intervenciones, para determinar si los programas llegan a las poblaciones destinatarias y cumplen sus objetivos, para comparar los resultados de diversos métodos o estrategias de prevención o para evaluar los efectos de la programación de la prevención en el estigma. La evaluación permite recopilar y analizar información a medida que se aplican y se amplían los programas, de forma que ayude a todas las personas implicadas en la prevención a sacar el máximo provecho de ella.

Como dicen Vega y Ventosa (1993), a pesar de la diversidad conceptual sobre el término de evaluación, todas tienen en común que evaluar es “atribuir cierto valor al objeto de la evaluación, es decir a los programas y estrategias preventivas”. A continuación, vamos a referirnos a tres definiciones por subrayar tres dimensiones diferentes de la evaluación.

- Cronbach (1981) define la evaluación de programas como la provisión de información para la toma de decisiones sobre una intervención.
- Fernández-Ballesteros (1992) afirma que consiste en una sistemática investigación del mérito, valor o éxito de intervenciones realizadas para producir cambios comportamentales en el contexto social.
- Fernández del Valle (1996) considera que la evaluación de programas es una

exigencia metodológica para todas las intervenciones que cubre una triple exigencia: en primer lugar, es una exigencia profesional ya que es responsabilidad de todos los que intervienen en los programas el conocer y estudiar sus propios procesos y resultados de intervención a fin de mejorar su ejercicio; en segundo lugar, es una exigencia científica en cuanto que la intervención psicosocial es un campo en construcción que demanda un desarrollo teórico que necesariamente pasan por analizar rigurosamente las prácticas desarrolladas, contrastando y comprobando las nuevas aportaciones técnicas y teóricas; y finalmente, la evaluación de programas es una exigencia ética ya que permite el control de los efectos causados con una determinada intervención, y por tanto, garantiza la adecuación a ciertos principios prescriptivos como la igualdad de oportunidades, el respeto a las minorías, etc.

Se considera que los tres objetivos básicos de la evaluación son: responsabilidad, mejora del programa y promoción de conocimientos básicos (Anguera, Chacon y Blanco, 2008).

La finalidad de la **evaluación de la responsabilidad** es emitir un juicio sobre el valor relativo del programa o de sus componentes, enfatizando la medición de resultados, el logro de objetivos, y el cálculo de la efectividad y eficacia de las actividades en comparación con otras alternativas. Para el análisis de las relaciones causales, el método preferido por muchos sería el método experimental controlado y aleatorizado (ver apartado evaluación de resultados).

Desde el punto de vista de la responsabilidad, la evaluación de programas ofrece una información valiosa para la toma de decisiones sobre: a) la posibilidad de generalizar o ampliar la aplicación de las acciones que componen el programa a otros contextos sociales; b) la necesidad de continuar invirtiendo recursos y esfuerzos en la dirección señalada por el programa evaluado; c) la conveniencia de reajustar los objetivos, las acciones o los recursos para un mejor logro de las metas; d) la posibilidad de reducir la intervención o e) darla por finalizada.

Es difícil llegar a respuestas fiables a partir de procedimientos informales, no sistemáticos; los planificadores y los técnicos de programas tienden al optimismo, los agentes que desarrollan las acciones pueden tener un interés particular en concluir que el programa ha resultado exitoso, los usuarios suelen valorar positivamente los programas en los que reciben atención (sin ningún esfuerzo o compromiso por su parte), independientemente de la permanencia en el tiempo de los efectos o del logro de los

verdaderos objetivos que se buscaban. Los estados, los gobiernos locales, las instituciones y empresas (públicas o privadas), dado que disponen de unos recursos limitados necesitan información para decidir lo más acertadamente posible la asignación de sus inversiones; esta información debería permitir, por un lado, priorizar las necesidades que se deben satisfacer y por otro determinar la eficiencia de las acciones emprendidas. Cada vez es más aceptada la idea de que la utilización de los conocimientos y la metodología científica mejorará la toma de decisiones, dará lugar a que se planeen mejores programas y, así, rendirán un servicio a los participantes de los programas de un modo más beneficioso y eficiente.

Desde *la perspectiva de la mejora*, la evaluación aspira orientar un mejoramiento y perfeccionamiento del programa. Persigue la mejora de las actividades, el fortalecimiento de la organización y el desarrollo de las capacidades de las personas que participan de distinto modo en los programas y en las organizaciones. Se considera que el programa continuará funcionando en un futuro previsible y por tanto la finalidad es lograr que el programa resulte más dinámico, eficiente, orientado al servicio y adaptado a las preocupaciones del cliente.

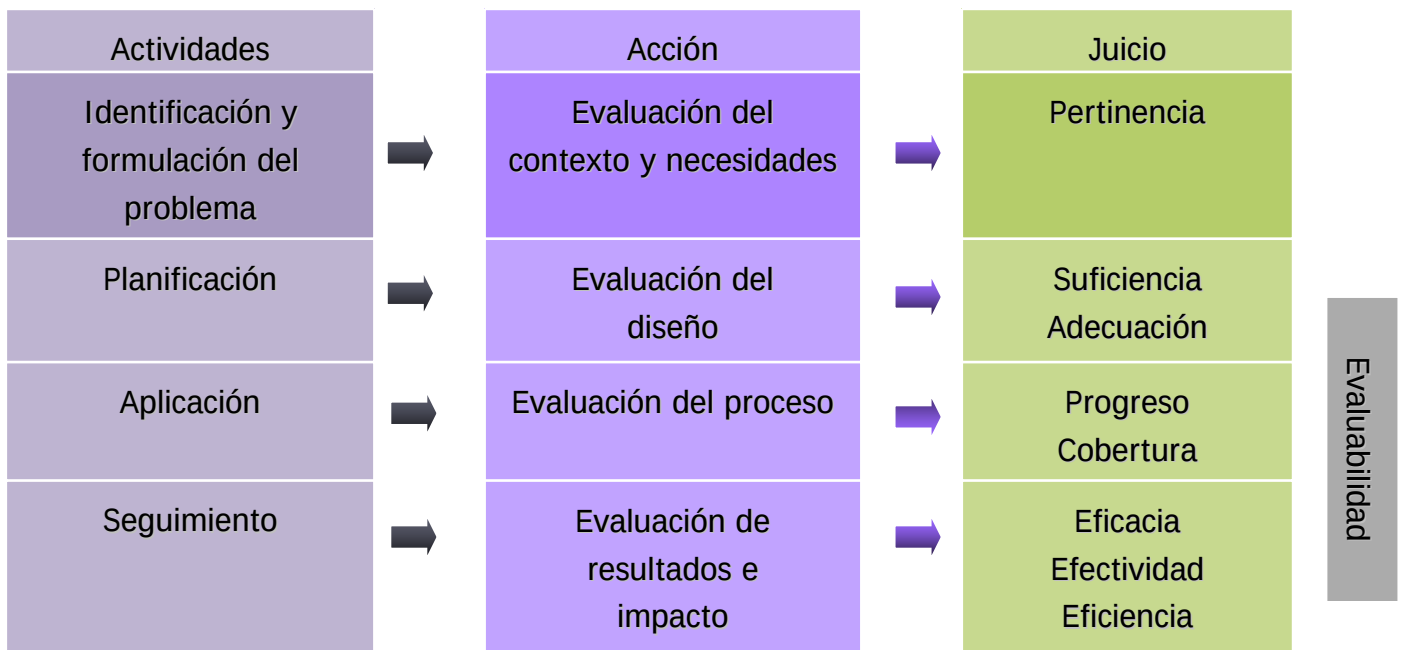
El tercer propósito sustantivo es el del *conocimiento básico*. La evaluación persigue comprender ampliamente los fenómenos bajo estudio, analizando las causas de los problemas y los porqués del funcionamiento de los programas con el objetivo de aumentar o refinar el conocimiento disponible sobre los problemas y las posibilidades de intervención. Cabe por tanto considerar que uno de los objetivos fundamentales de todo proceso evaluativo sea generar la comprensión o explicación de los problemas sociales, lo cual significa resaltar la dimensión investigadora implícita en toda evaluación (Chelimsky, 1997).

Una última consideración a tener en cuenta es no podemos hablar de auténtica evaluación mientras los “implicados” no participen activamente en la misma, de forma que se llegue al conocimiento a través de una construcción social en la que participen todos los sectores directamente afectados. Esta es la razón por la cual actualmente la investigación evaluativa utiliza soportes teóricos y metodológicos coherentes con este planteamiento, lo que ha dado lugar a una gran variedad y heterogeneidad de enfoques (evaluación pluralista, participativa, respondente, democrática, *cluster evaluation*, *empowerment evaluation*, etc-). El denominador común de estos enfoques se podría resumir en el hecho de que toda evaluación debe suponer, ante todo, un compromiso del

evaluador con la educación cívica, el desarrollo humano y la emancipación social (Fetterman, 1994; Smith, 1998).

A continuación pasaremos a presentar los tipos de evaluación. Como se refleja en el cuadro, existen diferentes tipos de evaluación, pero nosotros vamos a hacer referencia a cuatro de ellos: evaluación de proceso, de resultados y de impacto, junto a la evaluación de la evaluabilidad (Anguera, Chacon y Blanco, 2008).

Tipos de la Evaluación



La **evaluación de proceso** permite asegurarse de que el programa funciona como se quiere y al nivel que se quería, antes de atribuir efectos o consecuencias a un programa. Sus objetivos básicos son: a) averiguar y describir cómo funciona el programa y, b) analizar hasta qué punto hay diferencias entre el funcionamiento real y el previsto. Esta evaluación permite mejorar el programa sobre la marcha y establecer hasta qué punto los distintos elementos constituyentes del programa se están dispensando correctamente. Se tratará de analizar el progreso del programa, y es una medida que se realiza durante la aplicación del mismo o nada más finalizarlo, pretendiendo determinar cómo éste está actuando.

La **evaluación de resultados** nos permite conocer los efectos del programa a corto plazo, mientras que la **evaluación de impacto** se refiere a los efectos que el programa provoca a largo plazo sobre la población no destinataria y sobre el contexto del programa. Estos dos tipos de evaluación son sinónimos de evaluación de la efectividad, eficacia o eficiencia de un programa:

a) **Eficacia**. La evaluación de la eficacia trata de establecer si tras la aplicación del programa, los objetivos establecidos han sido alcanzados. Es una medida que nos permite conocer los resultados obtenidos en la población beneficiaria, en un periodo temporal

determinado, sin considerar los costos en los que incurre para obtenerlos. En caso de que el programa logre los objetivos marcados previamente, se considera que el programa es eficaz.

b) *Efectividad*. La evaluación de la efectividad tiene por objetivo establecer si independientemente de los objetivos establecidos por el planificador de la intervención, el programa ha conseguido otros determinados efectos. Existen consecuencias no intencionadas de las intervenciones que en ocasiones van en la dirección deseada, pero en otras no. Cuando los efectos son positivos se puede considerar que el programa evaluado ha sido efectivo.

c) *Eficiencia*. La evaluación de la eficiencia trata de indagar si los resultados obtenidos están en relación con los recursos y/o medios utilizados para conseguir unos determinados objetivos. Por tanto, implica su relación con los recursos invertidos y un análisis de costes-beneficios.

Como las evaluaciones suelen ser costosas en tiempo y en recursos, de forma que antes de embarcarse en una evaluación es conveniente analizar si es posible realizarla de forma satisfactoria. Para ello se debe considerar la adecuación de la planificación del programa y las posibles barreras con que se puede topar el evaluador. Para establecer si merece la pena comprometerse en un estudio evaluativo, Wholey (1987) sugiere efectuar previamente una ***evaluación de la evaluabilidad*** que busca responder a dos cuestiones: a) ¿es posible llevar a cabo la evaluación?; b) ¿en qué medida es posible responder a las preguntas evaluativas con la información que hay sobre el programa y en la situación en que se encuentra?.

1.- Evaluación de la Evaluabilidad

Comenzaremos por la evaluación de la evaluabilidad ya que debe ser previa a cualquiera de las evaluaciones mencionadas. Definiremos evaluabilidad como el análisis previo de la conveniencia de realizar la evaluación posterior de una intervención pública, valorando la existencia de las condiciones óptimas para llevarla a cabo y proponiendo las mejoras necesarias que preparen a dicha intervención para ser evaluada. Se estaría entonces en condiciones de poder determinar, en una fase posterior, el tipo de evaluación realizar.

Se usa deliberadamente el término análisis de evaluabilidad frente al concepto más estricto de evaluación de evaluabilidad para hacer referencia a un examen más rápido y sencillo de la intervención que no implique un trabajo de mucho calado ni consuma muchos recursos.

El objetivo principal que persigue el análisis de evaluabilidad es reducir incertidumbres sobre la decisión de poner en marcha la evaluación, siendo aconsejable, sobre todo, cuando se piensa que la evaluación a realizar requerirá una inversión significativa de recursos. Con ello se evitarán gastos y esfuerzos innecesarios (Wholey, 1979). También sienta las bases para que las evaluaciones a realizar tengan más probabilidades de cumplir las expectativas de las personas promotoras, usuarias y/o destinatarias finales, pues acota el alcance de lo que puede llegar a ofrecerse con la evaluación, mejorando por tanto la utilización y utilidad de los resultados evaluativos.

El análisis de evaluabilidad, como cualquier tipo de ejercicio evaluativo, tiene entre sus cometidos servir de herramienta al proceso de toma de decisiones públicas y a la mejora de las intervenciones ya que va a permitir determinar en qué ámbitos hay margen para mejorar, conocer qué circunstancias del evaluando y su contexto debe ser clarificado antes de evaluar y establecer una línea de base que pueda utilizarse a fin de que los resultados esperados de dicha intervención sean verificables de manera fiable y creíble.

Ahora bien, la valoración de la evaluabilidad no debe concluir, exclusivamente, con la respuesta positiva o negativa sobre la conveniencia de evaluar la intervención, puesto que las recomendaciones del análisis de evaluabilidad deben permitir preparar a la intervención para ser evaluada, en caso de que se considere necesaria o sea obligatorio realizar la evaluación. Entre las preguntas que la evaluabilidad permite responder se encuentran las siguientes:

- ¿Qué puede aportar la evaluación a esta intervención pública?
- ¿Merece la pena evaluar esta intervención pública?
- ¿Existen recursos suficientes para llevar a cabo el tipo de evaluación que se ha predeterminado?
- ¿Qué elementos de la intervención pública deberían considerarse en la evaluación?
- ¿Qué elementos de la intervención pública necesitan mejorarse para que ésta sea evaluable?

Aquí nosotros vamos a proponer el protocolo de evaluabilidad diseñado por el Instituto Navarro de Administración Pública del Gobierno de Navarra (2008) por proponer una herramienta que nos facilita la realización de este tipo de evaluación de intervenciones sociales. Para llevar a cabo este protocolo se propone recabar información sobre cinco criterios o dimensiones que deben cumplirse para garantizar la evaluabilidad de una intervención pública: calidad de la planificación, calidad del sistema de información, claridad en los propósitos de la evaluación, recursos destinados a la evaluación, implicación del personal relacionado con la intervención en el seguimiento y evaluación.

Con el fin de facilitar la aplicación de estos criterios, se detallan las preguntas de evaluación relevantes que deben ser formuladas y que caracterizan las cinco dimensiones analizadas: calidad de la planificación (0% cuando no existe cumplimiento alguno, 25% si considera que se cumple de forma insuficiente y por tanto, necesariamente mejorable; 50% si considera que se cumple pero que es mejorable; 75% si se cumple pero es parcialmente mejorable; y 100% si se cumple de forma muy satisfactoria). Se describen en un formato de tabla, a modo de fichas y en base a una escala de evaluación, fundamentada en la valoración del cumplimiento de cada una de las cuestiones a través de la generación de evidencias.

La metodología específica a aplicar se basa en la adaptación del denominado “Enfoque del Marco Lógico”¹.

Antes de pasar a describir esta herramienta, es importante señalar que sus creadores parten del convencimiento de que será más provechoso el proceso de aplicación del Protocolo que los resultados que se obtengan. Esto es así puesto que las personas implicadas en su puesta en marcha aprenderán de la intervención, de su teoría, de los mecanismos articulados para su gestión y aplicación y de la adecuación de los sistemas de información disponibles para recopilar información sobre la evolución de lo realizado y sus efectos.

Este Protocolo se ha diseñado para ser utilizado, fundamentalmente, de manera interna por el personal implicado en la elaboración del diseño, gestión y ejecución de las

¹ El enfoque de marco lógico (EML) es una herramienta analítica, desarrollada en los años 1970, para la planificación de proyectos con una orientación hacia el cumplimiento de objetivos y es utilizado con frecuencia por organismos de cooperación internacional. En el EML se considera que la ejecución de un proyecto es consecuencia de un conjunto de acontecimientos con una relación casual interna: insumos, actividades, resultados, objetivo específico y global.

intervenciones públicas, y especialmente por las personas encargadas de gestionar y coordinar su seguimiento y evaluación.

1.1.- Calidad de la planificación

Uno de los factores que pueden determinar la evaluabilidad de una intervención es la calidad de la planificación. Cuanto más imperfecta sea la planificación (problemas no identificados, objetivos deficientemente definidos y relaciones de causalidad entre actuaciones y objetivos mal establecidas), más dificultades habrá para realizar la evaluación. Los elementos a considerar en la valoración de la calidad de la planificación son: calidad de la gestión, definición de objetivos y examen de la coherencia interna de la estrategia propuesta en la intervención. Las preguntas a responder para determinar la evaluabilidad de la intervención en relación con la calidad del diagnóstico (problemas, población objetivo y contexto operativo) son las siguientes:

Calidad del Diagnóstico. Definición de problemas						
	0%	25%	50%	75%	100%	Observaciones
¿Existe un diagnóstico de los problemas que se pretenden resolver con la intervención?						
¿Se ha recogido el diagnóstico en algún documento programático de la intervención?						
¿Se identifican claramente los problemas y necesidades en el diagnóstico?						
¿Están cuantificados o son medibles?						
¿Se han definido indicadores que permitan establecer la situación de partida (línea de base) para la posterior valoración de la intervención en relación con la solución de los problemas identificados?						
¿Se cuenta con teorías de base que tratan de explicar los problemas que se quieren resolver?						
¿Se han identificado las causas de los problemas?						
¿Se ha acotado la cobertura de la intervención? Es decir, ¿se han definido las dimensiones institucionales, temporales, sectoriales y geográficas de la intervención?						
¿Se ha hecho un análisis de los elementos relevantes del contexto operativo en el que se desarrolla la intervención?						
¿Se ha explicitado a quiénes afectan los problemas detectados de manera adecuada? Es decir, ¿se ha definido la población objetivo?						
PROMEDIO²						

² El promedio de cada subdimensión se calcula sumando el valor obtenido en cada pregunta y se divide por el número de preguntas

Las preguntas a responder para determinar la evaluabilidad de la intervención en relación con la definición de los objetivos son las siguientes:

Definición de objetivos (para los diferentes niveles de objetivos)						
	0%	25%	50%	75%	100%	Observaciones
¿Se han definido explícitamente los objetivos perseguidos con la intervención?						
¿Son claros?						
¿Son concisos?						
¿Están cuantificados o son medibles o valorables?						
¿Responden los objetivos de la intervención a las necesidades y problemas detectados?						
¿Se ha definido el horizonte temporal en el que se pretenden alcanzar los objetivos?						
PROMEDIO						

Las preguntas a responder para determinar la evaluabilidad de la intervención en relación con la coherencia interna de la intervención son las siguientes:

Coherencia interna						
	0%	25%	50%	75%	100%	Observaciones
¿Se han establecido distintos niveles de objetivos?						
¿La jerarquía de objetivos responde a relaciones causa-efecto lógicas?						
¿Son coherentes los objetivos entre sí? Es decir, ¿la consecución de los objetivos de un nivel (por ejemplo, objetivos específicos) implica o puede implicar con bastante probabilidad el logro del objetivo del nivel superior de planificación (por ejemplo, objetivo final)?						
¿Se han concretado los objetivos en propuestas de actuación (actividades, acciones)?						
¿Son estas propuestas de actuación claras?						
¿Son concisas?						
¿Concretan un compromiso de acción?						
¿Son constatables, es decir, se pueden medir o valorar?						
¿Existe coherencia entre las actuaciones propuestas y los objetivos? Es decir, la puesta en marcha de las actividades ¿implica o puede implicar con bastante probabilidad el logro de los objetivos planteados?						
¿Se han asignado responsabilidades y presupuesto en relación con la aplicación de las actuaciones?						
PROMEDIO						

1.2.- Calidad del sistema de información

Uno de los principales condicionantes de la viabilidad y alcance de la evaluación es la disponibilidad de información sobre la intervención a evaluar. La cuestión radica fundamentalmente en la ausencia de datos de calidad, organizados y estructurados adecuadamente, lo que dificulta conocer quién hace qué, cómo y cuándo; y sobre todo, limita conocer los elementos clave en el desarrollo de la intervención que pueden motivar sus resultados y efectos.

Para evitar que esto ocurra debe haberse definido, desde el momento de planificar la intervención, un buen sistema de información que posibilite realizar el seguimiento periódico y sistemático de toda aquella información necesaria que facilite la posterior o posteriores evaluaciones.

Las preguntas a responder para determinar la evaluabilidad de la intervención en relación con la calidad del sistema de información son las siguientes:

Sistema de información						
	0%	25%	50%	75%	100%	Observaciones
¿Existe información suficiente sobre los aspectos de la intervención a evaluar que pueda suministrarse para la evaluación?						
¿Se ha definido algún procedimiento para la recogida de información sobre la intervención?						
¿Es fiable, eficaz, ágil y no muy costoso?						
¿Se han definido indicadores para recopilar sistemáticamente la información sobre la evolución de la intervención y sus actividades?						
¿Son relevantes los indicadores para la intervención?						
¿Son relevantes los indicadores para responder a las preguntas de evaluación que se han pensado?						
¿Hacen referencia a los objetivos definidos?						
¿Hacen referencia a las actividades definidas?						
¿Son los indicadores fiables?						
¿Son los indicadores fáciles de obtener?						
¿Son fáciles de interpretar?						
¿Se ha definido cuándo se debe recoger la información de cada indicador?						
¿Se han asignado recursos suficientes (financieros, materiales y humanos) para llevar a cabo el seguimiento?						
PROMEDIO						

1.3.- Claridad en los propósitos de la evaluación

La conveniencia de realizar una evaluación viene condicionada por la utilidad de sus resultados para la toma de decisiones y por su capacidad para influir en la intervención. Pues bien, la ausencia de un entendimiento claro sobre la utilidad de los procesos de evaluación es precisamente uno de los factores que dificultan y debilitan la práctica evaluativa.

Las cuestiones a responder para determinar la evaluabilidad de la intervención en relación con la definición de los propósitos de la evaluación son las siguientes:

Propósitos de la evaluación posterior						
	0%	25%	50%	75%	100%	Observaciones
¿Se han definido los propósitos de la evaluación?						
¿Se han definido las preguntas a las que la evaluación debe dar respuesta?						
¿Se ha identificado a los actores implicados en la intervención?						
¿Se ha realizado un análisis previo de los principales intereses y posibilidades de participación en el proceso de evaluación?						
En la formulación de preguntas de evaluación ¿se ha tenido en cuenta o se tendrán en cuenta los intereses de los grupos más representativos relacionados con la intervención?						
¿Se han acotado las dimensiones (temporales, geográficas, institucionales, etc.) de la intervención que van a ser consideradas en la evaluación?						
¿Perciben los actores implicados en la intervención la utilidad de la evaluación?						
¿Se utilizará la información que suministre la evaluación?						
¿Podría la evaluación, si así se concluye, influir en la intervención?						
¿Se difundirán los resultados de la evaluación, especialmente entre las organizaciones y entidades implicadas?						
PROMEDIO						

1.4.- Recursos destinados a la evaluación: coste, tiempo y recursos humanos

Otro de los principales condicionantes del alcance de una evaluación, y por tanto, determinante de la conveniencia o no de llevarla a cabo, es la adecuación de los recursos económico-financieros, materiales y humanos, asignados a esta tarea en relación con las necesidades reales detectadas.

Teniendo en cuenta las preguntas de evaluación formuladas y las dimensiones de la intervención que deben ser evaluadas, las cuestiones a responder para determinar la evaluación de la intervención en relación con los recursos destinados a la evaluación son las siguientes:

Recursos destinados a la evaluación						
	0%	25%	50%	75%	100%	Observaciones
¿Es el presupuesto adecuado para llevar a cabo la evaluación? Es decir, ¿es el presupuesto disponible un factor que puede limitar la evaluación?						
¿El tiempo asignado para realizar la evaluación es suficiente para llevarla a cabo? En definitiva, ¿es el tiempo disponible un factor que puede limitar la evaluación?						
En caso de optar por realizar una evaluación interna, ¿son los recursos humanos disponibles suficientes para llevar a cabo la evaluación?						
¿Cuando se espera recibir los productos de la evaluación? ¿Estarán a tiempo para que puedan mejorar la intervención?						
¿Se va a establecer algún tipo de procedimiento para dar seguimiento a la incorporación de las recomendaciones de la evaluación?						
PROMEDIO						

1.5.- Implicación del personal relacionado con la intervención en el seguimiento y evaluación

Ya se han señalado los aspectos fundamentales que determinan la evaluabilidad de una intervención. Existen, sin embargo, otros factores relacionados con la actitud ante la evaluación y con las condiciones técnicas e institucionales de la organización responsable de la intervención y de la encargada de la evaluación. Entre ellos:

Las preguntas a responder para determinar la evaluabilidad de la intervención en relación con la implicación del personal relacionado con la intervención en el seguimiento y la evaluación son las siguientes:

Implicación en la evaluación						
	0%	25%	50%	75%	100%	Observaciones
¿Hasta qué punto estarán implicados los agentes clave?						
¿Existe una actitud favorable hacia la evaluación en el personal técnico, gestores y políticos?						
¿Existen grupos de presión que puedan interferir en la independencia de la evaluación?						
¿Se facilitará al equipo de evaluación el acceso a toda la información y documentación relevante de la intervención, y a los agentes e informantes clave que deban participar en entrevistas, grupos de discusión o cualquier otra técnica de recopilación de información?						
¿Se facilitarán los canales de comunicación más adecuados para que el equipo de evaluación recopile la información necesaria?						
¿Se facilitará la participación de los “actores implicados” en el desarrollo de la evaluación?						
¿Existe o existirá un clima favorable de cooperación entre personas y entidades evaluadoras y evaluadas?						
¿Se tiene experiencia previa en el diseño, gestión o coordinación de una evaluación?						
¿Tiene el equipo encargado de gestionar la intervención, formación específica en evaluación?, ¿y en la materia objeto de la evaluación?						
¿Se va a establecer algún tipo de procedimiento para dar seguimiento al desarrollo de la evaluación?						
¿Se van a asignar responsabilidades en relación con la ejecución, validación y control de calidad de las evaluaciones a realizar?						
PROMEDIO						

1.6.- Orientaciones para la aplicación e interpretación del protocolo de evaluabilidad

En primer lugar, se propone evaluar cada una de las dimensiones y subdimensiones de forma independiente. De forma que si una de las subdimensiones no supera el promedio de un 50%, se debe intentar mejorar aquellos aspectos que requieran atención antes de comenzar el proceso de evaluación.

En segundo lugar, también se puede hacer una valoración global de todas las dimensiones, dándole a cada una de ellas un peso relativo. En las II Jornadas de Sensibilización y Formación sobre Evaluación de Políticas Públicas celebrada el mes de noviembre de 2008 por el INAP, se desarrolló un Taller de Trabajo en el que se propuso

a los asistentes la especificación y ponderación de cada uno de los ítemes presentados. Finalmente, el peso relativo asignado a cada una de las cinco dimensiones del Protocolo de Evaluabilidad se especifica en la siguiente tabla. Para una mejor comprensión, se ha definido de modo que la suma de todos ellos sea igual a 100. Se recomienda que sean los propios equipos de trabajo quienes establezcan el peso relativo de las dimensiones, ya que la ponderación que se presenta a continuación sólo es a modo orientativo.

Ponderación Dimensiones Protocolo de Evaluabilidad	%
■ Calidad del Sistema de Información	50
■ Calidad de la Planificación ³	20
■ Recursos destinados a la evaluación	15
■ Implicación del personal relacionado con la intervención en el seguimiento y evaluación	10
■ Claridad en los propósitos de evaluación	5

La recomendación final del Protocolo es que para garantizar la evaluabilidad de una intervención pública las cinco dimensiones propuestas deben de estar evidenciadas de acuerdo al peso relativo determinado y con un nivel de cumplimiento, al menos, por encima del 50%. En caso contrario, se recomienda realizar acciones de mejora previas a abordar una evaluación.

2.- Evaluación de Proceso

La evaluación de proceso se dirige a la comprobación de las actividades que se están desarrollando en la intervención y su congruencia con la planificación inicial, con vistas a la mejora continua y el apoyo a los responsables de la implementación. Esta evaluación es una actividad de evaluación interna que tiene un carácter continuo y suele ser realizada por los propios implicados en la marcha del programa, recogiendo información constante con el objetivo de gestionar y dirigir correctamente las acciones.

La evaluación de proceso pretende conocer si el programa se está dispensando convenientemente, e implica dos objetivos:

³ Calidad de la Planificación. Para calcular el promedio conjunto de las tres subdimensiones que forman esta dimensión se aconseja sumar los tres promedios y dividirlos entre tres.

- Establecer hasta qué punto los distintos elementos constituyentes del programa se han aplicado como se había previsto.
- Evaluar la opinión de los participantes sobre los distintos aspectos estructurales del programa.

Entre los métodos de investigación útiles para este tipo de evaluación se pueden utilizar todo tipo de métodos y técnicas de investigación, incluyendo métodos cuantitativos y cualitativos, datos de archivo, estadísticas de indicadores clave, etc. Algunos prefieren las técnicas de tipo cualitativo porque ofrecen una información más rica para el seguimiento de la aplicación, recogiendo directamente las opiniones, percepciones o valoraciones que los implicados tienen sobre las actividades que se realizan, la dotación de recursos y los primeros efectos que se empiezan a producir. En general, se utilizan técnicas de observación o entrevista, aunque no se descarta, por ejemplo, el uso de cuestionarios o las hojas de registro, útiles para una recogida rápida de información entre el personal o los usuarios del programa.

En este apartado vamos hacer referencia al segundo objetivo de la evaluación de proceso, es decir la evaluación de la opinión de los participantes sobre los distintos aspectos estructurales del programa. Para ello vamos a presentar un ejemplo de la evaluación de proceso de un programa de prevención del VIH/SIDA dirigido a jóvenes.

Vamos a hacer una síntesis lo más sencilla posible de los pasos que se deben dar para llevar a cabo una evaluación de proceso (Ubillos, 2005):

- 1) Listado de Aspectos a Evaluar.
- 2) Selección o diseño de instrumento de medida
- 3) Administración del instrumento de medida.
- 4) Análisis e interpretación de datos.
- 5) Redacción de un informe de evaluación

2.1.- Listado de aspectos a evaluar

En primer lugar, se debe realizar una descripción de qué es lo esencial del programa. Para ello, se elabora una lista del conjunto de rasgos y actividades que lo definen y constituyen. Por ejemplo, en un programa de prevención sexual del VIH dirigido a jóvenes o a usuarios de drogas serían el horario, el lugar físico, la asistencia y participación, los contenidos, etc. Todos ellos podrían ser aspectos del proceso que se

podrían evaluar. Muchos de los aspectos mencionados son genéricos, por tanto pueden ser seleccionados para evaluar muchos tipos de programas o talleres.

Evaluación de Proceso: Listado de Aspectos a Evaluar

- Valoración general del Programa
- Grado de Satisfacción con el Programa
- Utilidad del Programa
- Tiempo destinado al Programa
- Interés suscitado por el Programa
- Implicación provocada por el Programa
- Participación de los asistentes
- Grado de comprensibilidad de la información proporcionada
- Adecuación de los contenidos
- Adecuación de la metodología
- Aportaciones del programa

2.2.- Selección o diseño del instrumento de medida

En segundo lugar, hay que diseñar o seleccionar un instrumento de medida que nos permita recoger información sobre los aspectos que se quieren evaluar. Lo que se muestra a continuación es un ejemplo de instrumento de medida en el que se han incluido preguntas genéricas que pueden ser utilizadas en la mayoría de los programas o intervenciones. Se podrían incluir más preguntas sobre datos de identificación, como el curso o el centro donde se imparte el programa en el caso de que nuestros participantes sean estudiantes. Por tanto, además del sexo y la edad, el tipo de preguntas que consideremos relevantes para nuestra evaluación dependerá del tipo de usuarios de nuestro programa (ejemplo, homosexuales, transexuales, lesbianas, usuarios de drogas, etc.). Así mismo también se pueden incluir otras preguntas sobre aspectos más específicos del programa, como por ejemplo cuestiones sobre los distintos temas abordados en el programa (por ejemplo, normas de uso del preservativo, habilidades de comunicación y negociación, etc.), o sobre las distintas dinámicas aplicadas (por ejemplo, modelaje y entrenamiento para el uso del preservativo, role playing para las habilidades de comunicación y negociación). También se suelen introducir preguntas abiertas a fin de que los usuarios del programa puedan expresar de forma cualitativa los aspectos positivos y negativos de la intervención (Ubillos, 2005).



Lafia-Nahia Sexologia Elkartea

ESCALA PARA EVALUAR EL PROCESO DEL PROGRAMA

Atención:

Para responder a las preguntas del cuestionario debes elegir tu respuesta y marcar con un círculo o marcar con una cruz el número que corresponde a dicha respuesta. Antes de nada no olvides indicar tu sexo y edad:

Sexo: 1 Hombre 2 Mujer

Edad: _____

1.- Haber realizado este programa me ha parecido:

1 = Totalmente negativo

4 = Positivo

2 = Muy negativo

5 = Muy positivo

3 = Negativo

6 = Totalmente positivo

2.- Señala el grado de satisfacción con el programa que has realizado:

1 = Totalmente insatisfecho

4 = Satisfecho

2 = Muy insatisfecho

5 = Muy satisfecho

3 = Insatisfecho

6 = Totalmente satisfecho

3.- Señala el grado de utilidad que ha tenido para ti el realizar este programa:

1 = Nada útil

4 = Útil

2 = Poco útil

5 = Muy útil

3 = Algo útil

6 = Totalmente útil

4.- El tiempo empleado en la realización del programa, te ha parecido:

1 = Totalmente corto

5 = Largo

2 = Muy corto

6 = Muy largo

3 = Corto

7 = Totalmente largo

4 = Normal

5.- Señala tu grado de interés en el programa que has realizado:

1 = Nada interesado

4 = Interesado

2 = Poco interesado

5 = Muy interesado

3 = Algo interesado

6 = Totalmente interesado

6.- Señala tu grado de implicación en el programa:

1 = Nada implicado

4 = Implicado

2 = Poco implicado

5 = Muy implicado

3 = Algo implicado

6 = Totalmente implicado

7.- Señala tu grado de participación en el programa:	
1 = Nada participativo	4 = Participativo
2 = Poco participativo	5 = Muy participativo
3 = Algo participativo	6 = Totalmente participativo
8.- La información te ha parecido:	
1 = Totalmente incomprensible	4 = Comprensible
2 = Muy incomprensible	5 = Muy comprensible
3 = Incomprensible	6 = Totalmente comprensible
9.- Valora en qué medida consideras adecuados contenidos abordados en los talleres:	
1 = Nada adecuados	4 = Adecuados
2 = Poco adecuados	5 = Muy adecuados
3 = Algo adecuados	6 = Totalmente adecuados
10.- Valora en qué medida consideras adecuada la metodología empleada:	
1 = Nada adecuada	4 = Adecuada
2 = Poco adecuada	5 = Muy adecuada
3 = Algo adecuada	6 = Totalmente adecuada
11.- ¿En qué medida este programa te ha aportado conocimientos?:	
1 = Muy deficiente	4 = Bien
2 = Insuficiente	5 = Notable
3 = Suficiente	6 = Sobresaliente
12.- ¿En qué medida este programa te ha hecho reflexionar?:	
1 = Muy deficiente	4 = Bien
2 = Insuficiente	5 = Notable
3 = Suficiente	6 = Sobresaliente
13.- ¿En qué medida este programa te ha influido o va a influir en tu forma de actuar?:	
1 = Muy deficiente	4 = Bien
2 = Insuficiente	5 = Notable
3 = Suficiente	6 = Sobresaliente
14.- A continuación valora del 1 al 6 la actuación de la profesional responsable de impartir el programa:	
1 = Muy deficiente	4 = Bien
2 = Insuficiente	5 = Notable
3 = Suficiente	6 = Sobresaliente

2.3.- Administración del Instrumento de Medida

El tercer paso es decidir cuándo y cómo se administra el instrumento de medida a los asistentes al programa.

Si el programa tiene una duración corta, es decir entre 2 y 20 horas aproximadamente, y se hace de forma intensiva, lo más conveniente es pasarlo al finalizar la última sesión. Si, por el contrario, el programa tiene una duración amplia y se extiende mucho en el tiempo, lo aconsejable es hacer evaluaciones intermedias, teniendo en cuenta algún criterio de tiempo (una vez al mes) o de bloques (si el programa está dividido en módulos de temas diferenciados).

Si los asistentes tienen un nivel cultural medio, es decir que son capaces de leer, comprender y escribir sin problemas, entonces lo ideal es pasar el instrumento de medida de forma auto-aplicada. Sin embargo, si los participantes tienen dificultades de lectura y comprensión, entonces lo más aconsejable es administrarlos de forma hetero-aplicada, es decir los participantes lo van cumplimentando con ayuda externa, por ejemplo de la persona responsable de pasar las escalas. Dentro de esta opción, la administración hetero-aplicada puede hacerse con todo el grupo a la vez, de forma que el responsable de la evaluación va dando las instrucciones mientras que los participantes van respondiendo. También puede hacerse en forma de entrevista estructurada, es decir los participantes van respondiendo uno a uno la encuesta, con la ayuda de un profesional. El problema de esta última modalidad es que se requiere invertir mucho tiempo y recursos.

2.4.- Análisis e Interpretación de Datos

Una vez cumplimentadas las escalas, pasamos al análisis e interpretación de los datos. Para ello debemos de seguir una serie de pautas. La primera de ellas es la introducción de los datos de los cuestionarios en una base que nos permite analizarlos. El paquete estadístico más utilizado en el campo de las ciencias sociales es el SPSS, y es al que vamos a hacer referencia. Se recomienda el libro de Martín, Cabero y del Rosario (2007).

Antes de nada a cada cuestionario cumplimentado asígnele un número correlativo (1, 2, 3, 4,..). A continuación para introducir los datos primero hay que preparar lo que se denomina “la matriz de datos”. Para ello abra el SPSS, y le aparecerá una pantalla en blanco, igual a la que se indica en la siguiente figura. Fíjese que en el ángulo inferior

izquierdo hay dos pestañas, usted debe hacer clic en la pestaña que indica **vista de variables** y parece la ventana que se muestra en la figura 1. En realidad son dos ventanas que están conectadas: una la de vista de variables y otra la de vista de datos.

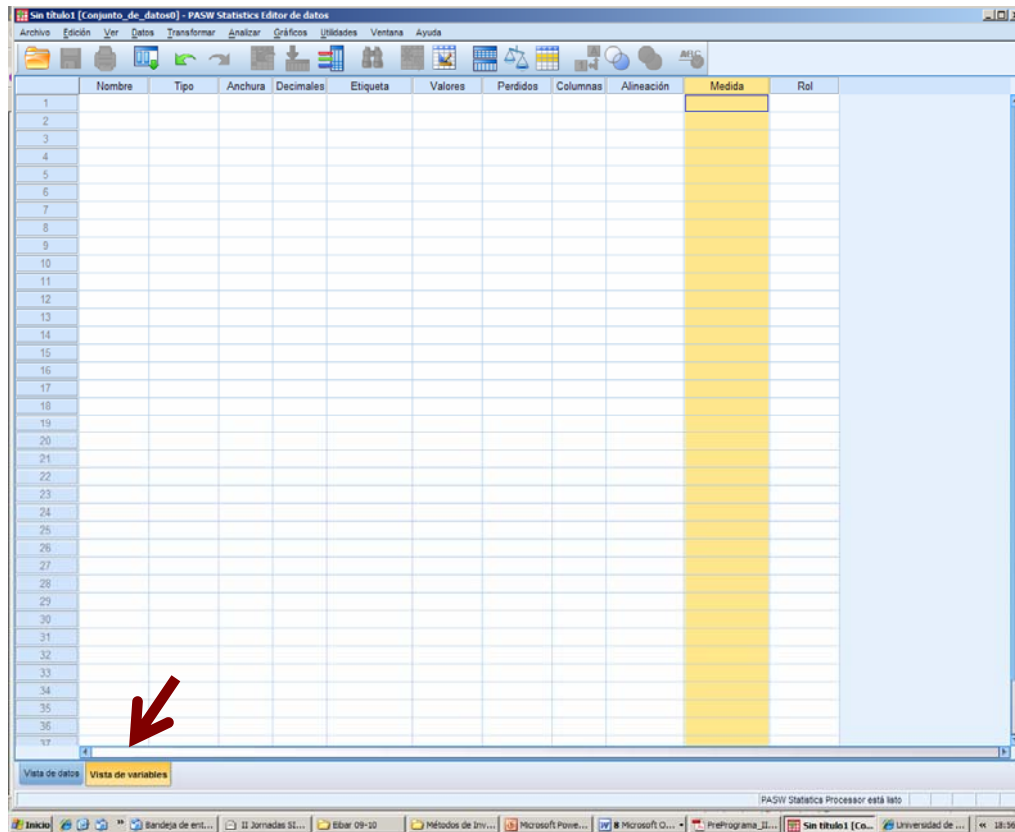


Figura 1

A continuación debemos diseñar la matriz de datos que consiste en introducir las preguntas del cuestionario en esta ventana de variables. Cada fila corresponde a la definición de un aspecto que a partir de ahora lo denominaremos variable. En la primera columna, denominado “*nombre*”, se indica el nombre de la variable y no debe contener espacios en blanco. Tomemos de ejemplo la primera variable de nuestro estudio que le vamos a denominar sujeto, se refiere al número que le hemos asignado a cada cuestionario. La segunda fila la reservamos para el sexo, de forma que el nombre será sexo. La columna denominada “*etiqueta*” sirve para explicar el acrónimo del nombre de la variable, que en este caso es obvio, sexo del sujeto. Y en la columna de “*valores*”, se indica la codificación de la variable. De forma que hacemos clic en valores y aparece un cuadro de diálogo, donde en valor indicamos 1, etiqueta “hombre”, y damos añadir, y volvemos a indicar en valor 2, etiqueta “mujer”, damos a añadir, y aceptar. En la tercera fila hemos indicado la variable edad. Y en la cuarta la variable que le hemos dado el nombre de

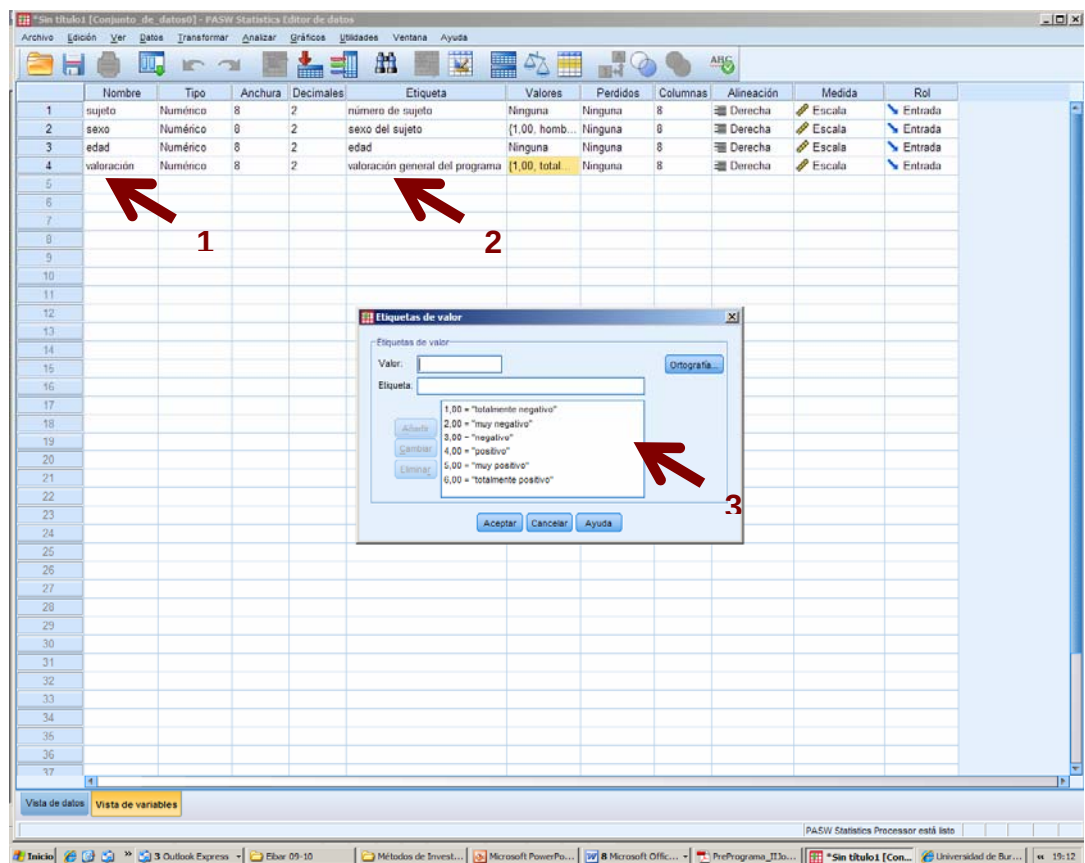


Figura 2

The screenshot shows the PASW Statistics Editor de datos interface. The main window displays a table with columns: Nombre, Tipo, Anchura, Decimales, Etiqueta, Valores, Perdidos, Columnas, Alineación, Medida, and Rol. The table contains variables 1 through 17. The table is shown in the 'Vista de variables' view.

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
1	sujeto	Númerico	8	2	número de sujeto	Ninguna	Ninguna	8	Derecha	Escala	Entrada
2	sexo	Númerico	8	2	sexo del sujeto	{1,00, homb...	Ninguna	8	Derecha	Escala	Entrada
3	edad	Númerico	8	2	edad	Ninguna	Ninguna	8	Derecha	Escala	Entrada
4	valoración	Númerico	8	2	valoración general del programa	{1,00, total...	Ninguna	8	Derecha	Escala	Entrada
5	satisfacción	Númerico	1	0	satisfacción	{1, totalmen...	9	12	Derecha	Escala	Entrada
6	utilidad	Númerico	1	0	utilidad del programa	{1, totalmen...	9	10	Derecha	Escala	Entrada
7	duración	Númerico	1	0	duración del programa	{1, totalmen...	9	12	Derecha	Escala	Entrada
8	interés	Númerico	1	0	interés de la persona	{1, nada int...	9	14	Derecha	Escala	Entrada
9	implicación	Númerico	8	0	implicación de la persona	{1, nada im...	Ninguna	13	Derecha	Escala	Entrada
10	participación	Númerico	1	0	participación de la persona	{1, nada par...	9	12	Derecha	Escala	Entrada
11	comprensibilidad	Númerico	1	0	comprensibilidad de la inform...	{1, totalmen...	9	15	Derecha	Escala	Entrada
12	información	Númerico	8	0	contenidos adecuados	{1, totalmen...	Ninguna	8	Derecha	Escala	Entrada
13	metodología	Númerico	1	0	metodología adecuada	{1, totalmen...	9	14	Derecha	Escala	Entrada
14	contenidos	Númerico	1	0	el programa me ha enseñado	{1, muy def...	9	10	Derecha	Escala	Entrada
15	actitudes	Númerico	1	0	he reflexionado	{1, muy def...	9	9	Derecha	Escala	Entrada
16	conductas	Númerico	1	0	el programa ha influido	{1, muy def...	9	11	Derecha	Escala	Entrada
17	profesionales	Númerico	1	0	actuación de las profesionales	{1, muy def...	9	9	Derecha	Escala	Entrada

Figura 3

“valoración” (paso 1), la etiqueta la hemos definido como valoración general del programa (paso 2), y en la columna de valores hemos introducido los 6 posibles rangos de respuesta (paso 3). En la figura 2, se ve como a modo de ejemplo hemos ido introduciendo las variables hasta la variable denominada valoración. Y seguimos este proceso con cada una de las variables del cuestionario (ver figura 3).

Una vez construida la “matriz de datos”, introducir los datos de los cuestionarios, para ello se debe hacer clic en la pestaña del ángulo inferior izquierdo que pone “vista de datos” (paso 1), y se comienza a introducir en la primera fila todos los datos del sujeto o cuestionario 1, en la segunda los datos del sujeto 2, etc. (paso 2). Si alguna persona no ha respondido a una de las preguntas se deja en blanco (paso 3) (ver figura 4).

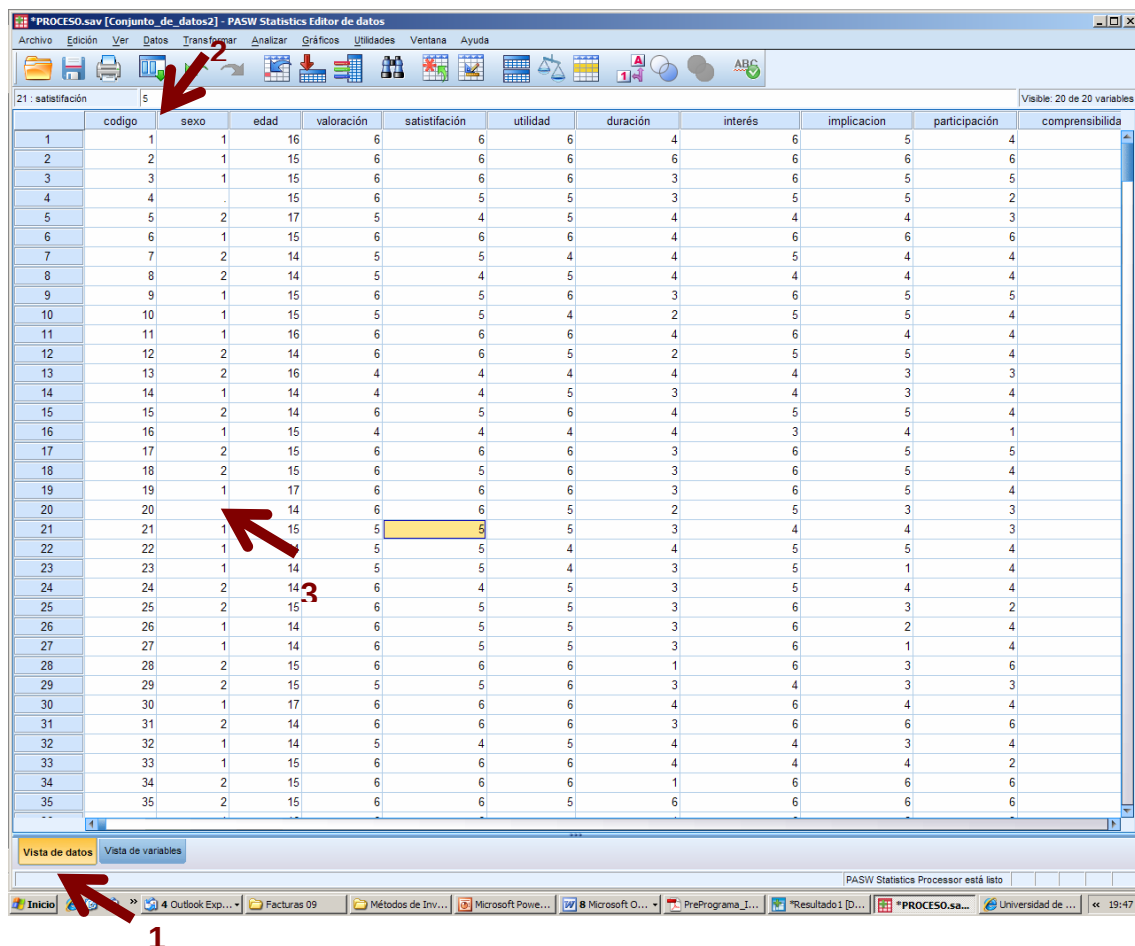


Figura 4

A continuación, hay que depurar la matriz de datos, ya que en su introducción puede haber errores que deben ser detectados antes de pasar a efectuar los análisis pertinentes. Los errores que se pueden detectar se refieren a aquellos valores que hemos introducido por error y que no se sitúan dentro del posible rango de respuesta de las variables. O que exceden el valor mínimo o máximo.

Para llevar a cabo la depuración, como se refleja en la figura 5, hay que hacer clic en analizar (paso 1), de forma que se despliega una serie de opciones y se debe hacer clic en la segunda opción, es decir en “*estadísticos descriptivos*” (paso 2), lo que a su vez da lugar a otra serie de opciones y se hace clic en la que indica “*frecuencias*” (paso 3) y aparece un cuadro de diálogo. En este cuadro en la ventana izquierda se encuentran todas las variables de nuestra matriz, se seleccionan todas las variables (paso 4), se hace clic en la flecha y automáticamente todas las variables se sitúan en la ventana de la derecha (paso 5). En este mismo cuadro se hace clic en la opción de “*estadísticos*” (paso 6), y aparece otro cuadro de diálogo en el que señalamos valor mínimo y máximo (paso 7), damos a continuar (paso 8) y a aceptar (paso 9).

Al hacer clic en aceptar aparece otro archivo denominado “archivo de resultados” que está conectado con la matriz de datos. Fíjese en la figura 6 que en el ejemplo que le mostramos, en la variable “valoración global” aparece como valor máximo 66, dato que es imposible, ya que el rango de respuesta posible de esta variable oscila entre 1 (totalmente negativo) y 6 (totalmente positivo), por lo que ninguna persona ha podido señalar el valor 66. Si nos fijamos en la tabla de frecuencias podemos observar que el valor 66 aparece en un caso (podría ser el caso de que el valor 66 por error lo hubiésemos introducido en vez de en una sola ocasión en varias ocasiones). Asimismo pueden aparecer errores en diversas variables de nuestra matriz.

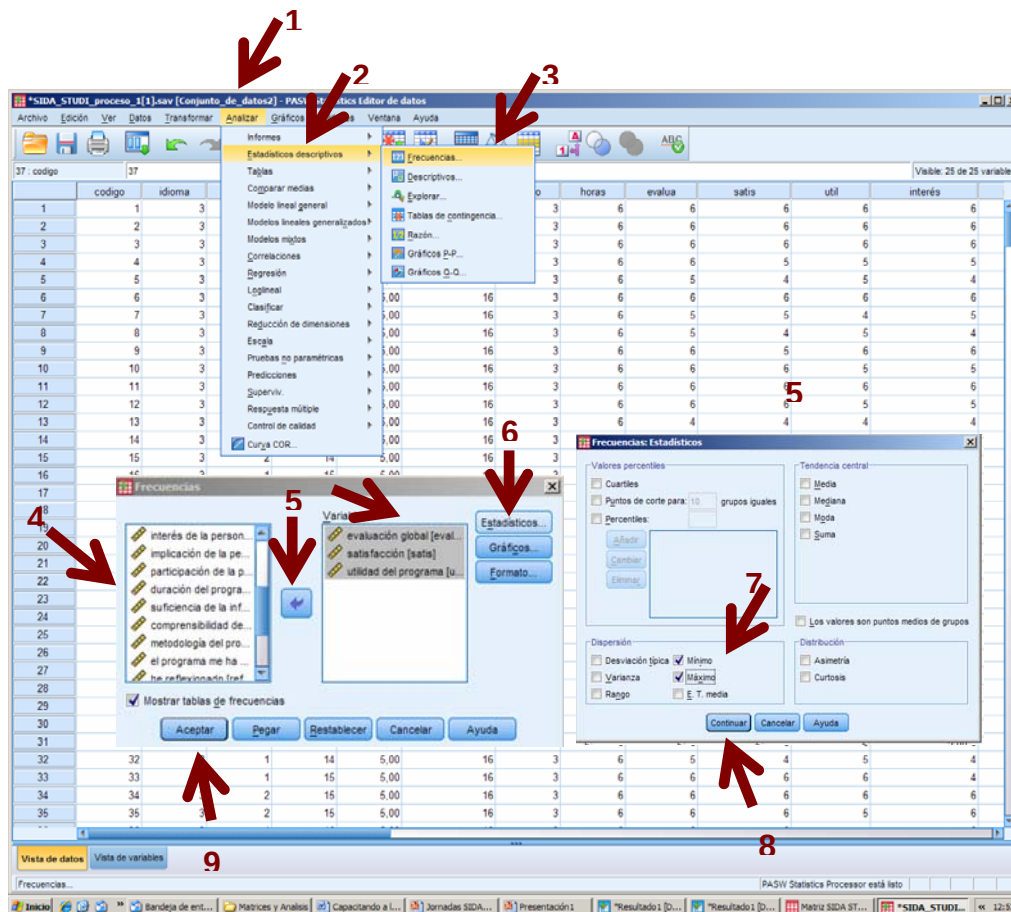


Figura 5

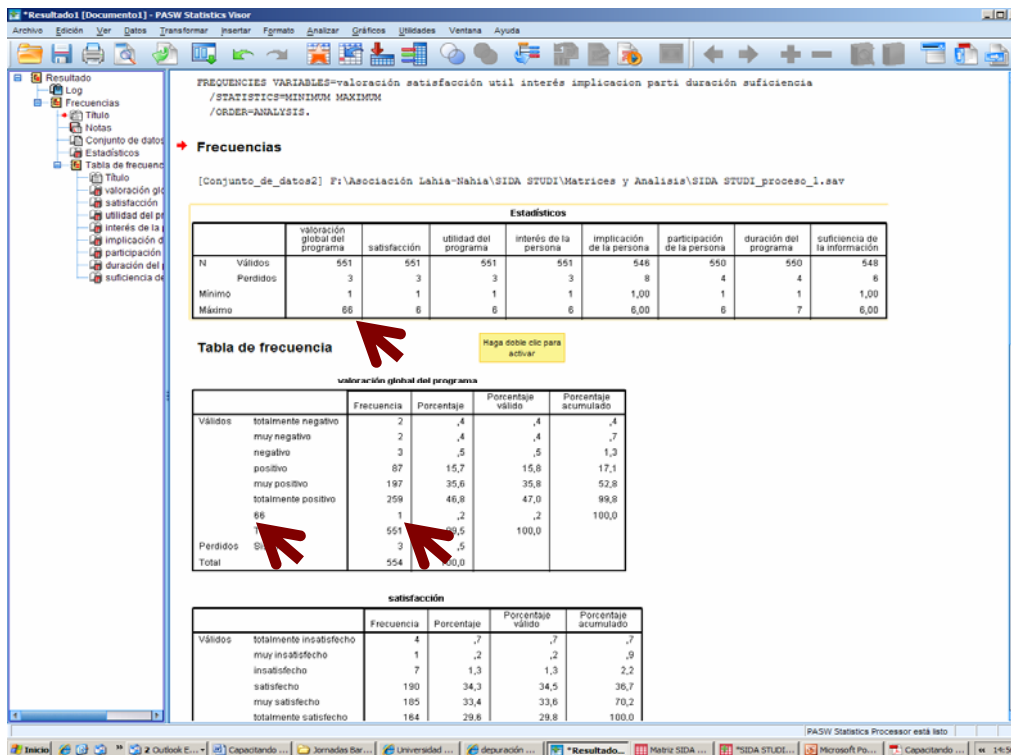


Figura 6

En la figura 7 describimos cuál es el procedimiento que hay que llevar a cabo para poder corregir el error. Uno se sitúa en la matriz de datos, en vista de datos, selecciona la variable donde se encuentra el error, en nuestro caso en la variable “valoración global del programa”, cuyo acrónimo es “evalua” (paso 1), se hace clic en editar (paso 2), se abren una serie de opciones y se hace clic en buscar (paso 3), de forma que aparece un cuadro de diálogo. En la ventana que indica buscar ponemos el valor que queremos hallar, en este caso 66 (paso 4) y hacemos clic en buscar siguiente (paso 5) de forma que automáticamente el cursor se sitúa encima del valor 66 (paso 6). Comprobamos a que número de cuestionario se refiere, buscamos el cuestionario, revisamos el valor correcto y lo introducimos (paso 7).

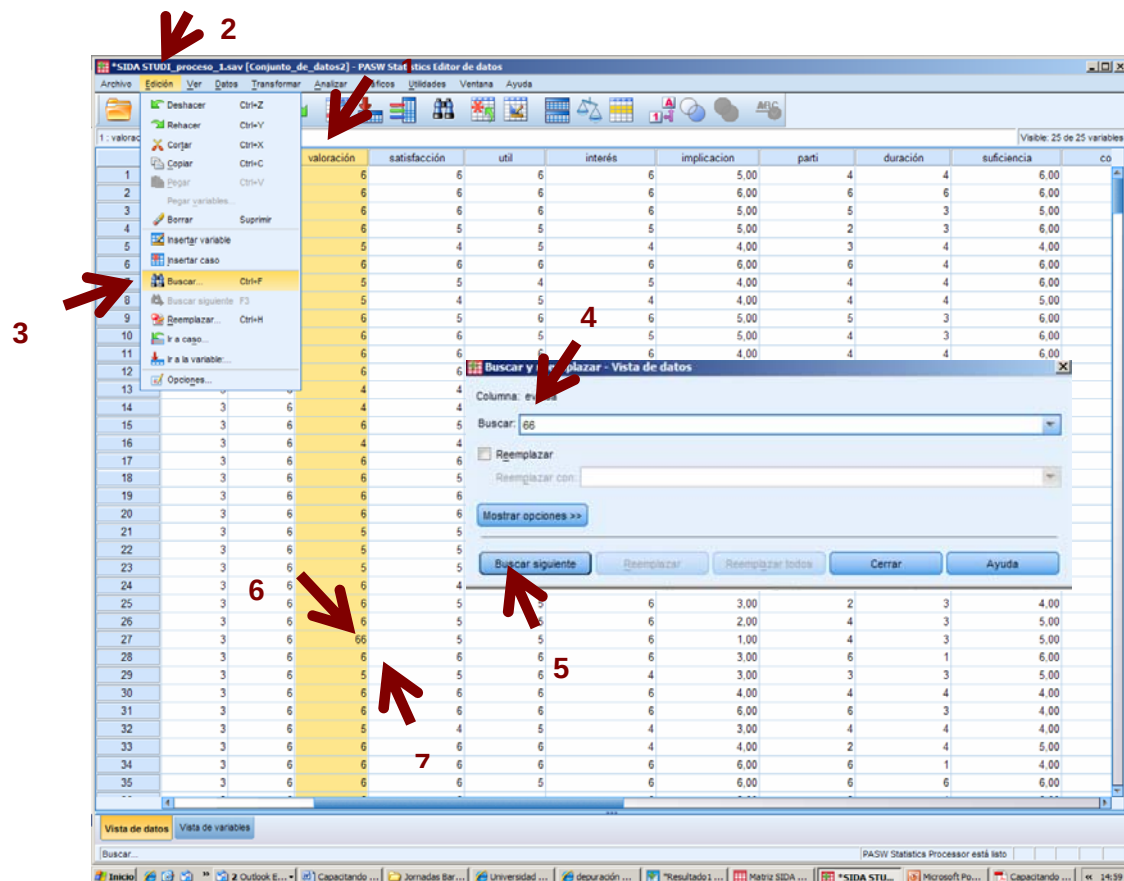


Figura 7

Ahora pasamos a lanzar los análisis e interpretarlos. En este tipo de evaluación vamos a aplicar los estadísticos descriptivos. Antes de explicar cómo se lanzan, describiremos brevemente dos de ellos.

- **Tendencia Central:** Son valores que intentan representar a toda una muestra con un valor numérico hacia el cual tienden a agruparse los datos. Aquí vamos a hacer referencia a la Media ($\bar{X}$ o M): Es el punto en la escala de valores que corresponde a la suma de los valores dividida entre el número total de ellos. Es el índice de tendencia central que suele referirse como promedio de los valores de la variable e indica el valor en el que tienden a agruparse los valores de la misma. Es la medida que se utiliza con mayor frecuencia. Muchas de las pruebas importantes de significación estadística se basan en la media.
- **De Dispersión:** Son valores que pretenden indicar el grado de variabilidad de los datos de la muestra. Aquí sólo vamos a hacer referencia a la desviación estándar o típica (S o DE). Este valor acompaña siempre a la media, y nos indica hasta qué punto las puntuaciones de los participantes se desvían de la media. Por ejemplo: un grupo de dos personas que tienen que valorar el programa de 0 a 10, uno puntúa 5 y el otro también, la media del grupo es 5, y la desviación típica 0, ya que todos los miembros del grupo califican el programa de la misma manera. Sin embargo otro grupo de dos personas, uno puntúa el programa con un 0 y otro con un 10, la media del grupo también es 5, pero la desviación típica es 5, ya que es un grupo muy heterogéneo, a una persona le ha parecido muy positivo y a otro muy negativo. Por ello, junto a la media siempre se debe indicar la desviación típica para saber si se trata de un grupo homogéneo o heterogéneo.

Tanto las medidas de tendencia central como las medidas de dispersión únicamente son aplicables a variables cuantitativas. La única forma de sintetizar los resultados de una variable cualitativa (por ejemplo el sexo) es mediante la frecuencia de sus categorías o en forma de porcentajes.

A fin de analizar los datos, se debe hacer clic en la pestaña superior que indica *analizar* (paso 1), otra vez haga clic en la segunda opción que indica *estadísticos* (paso 2) y en la primera opción que indica *frecuencias* (paso 3) (ver figura 8).

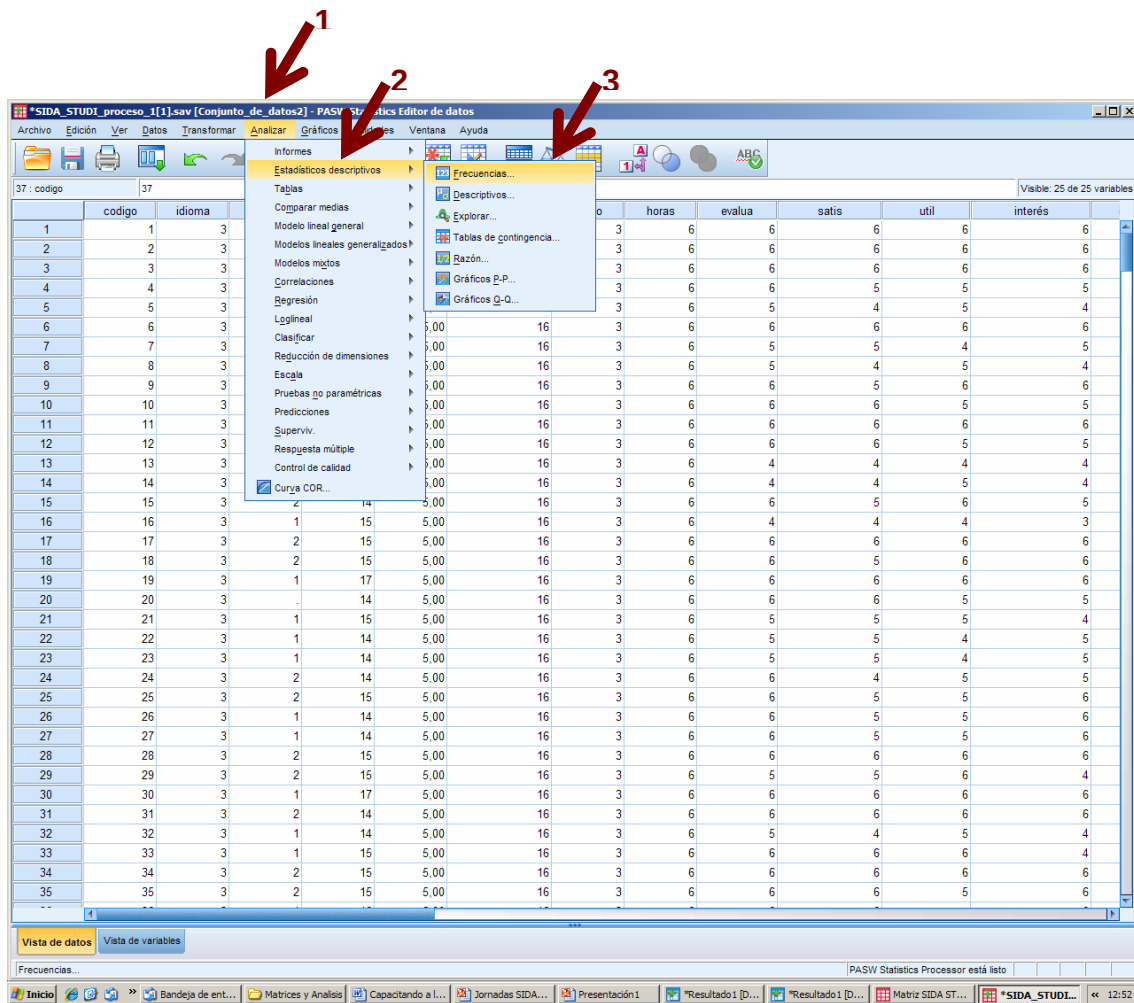


Figura 8

Como se puede observar en la figura 9, al hacer clic en frecuencias aparece un cuadro de diálogo, con dos ventanas, en la de la izquierda vienen indicadas todas las variables que hemos introducido en la matriz, se seleccionan todas aquellas que queramos analizar (paso 1), se hace clic en la flecha que hay entre las dos ventanas (paso 2) y automáticamente las variables seleccionadas se sitúan en ventana de la derecha (paso 3). Una vez efectuado esto, se hace clic en *estadísticos* (paso 4), y aparece otro cuadro de diálogo donde se hace clic en *media y desviación típica* (paso 5), y se da a continuar (paso 6). Finalmente se da a aceptar en el cuadro de frecuencias (paso 7) y aparece otro archivo que está conectado con la matriz de datos, y que es el archivo de resultados. En este archivo se van anexando todos los resultados (ver figura 10).

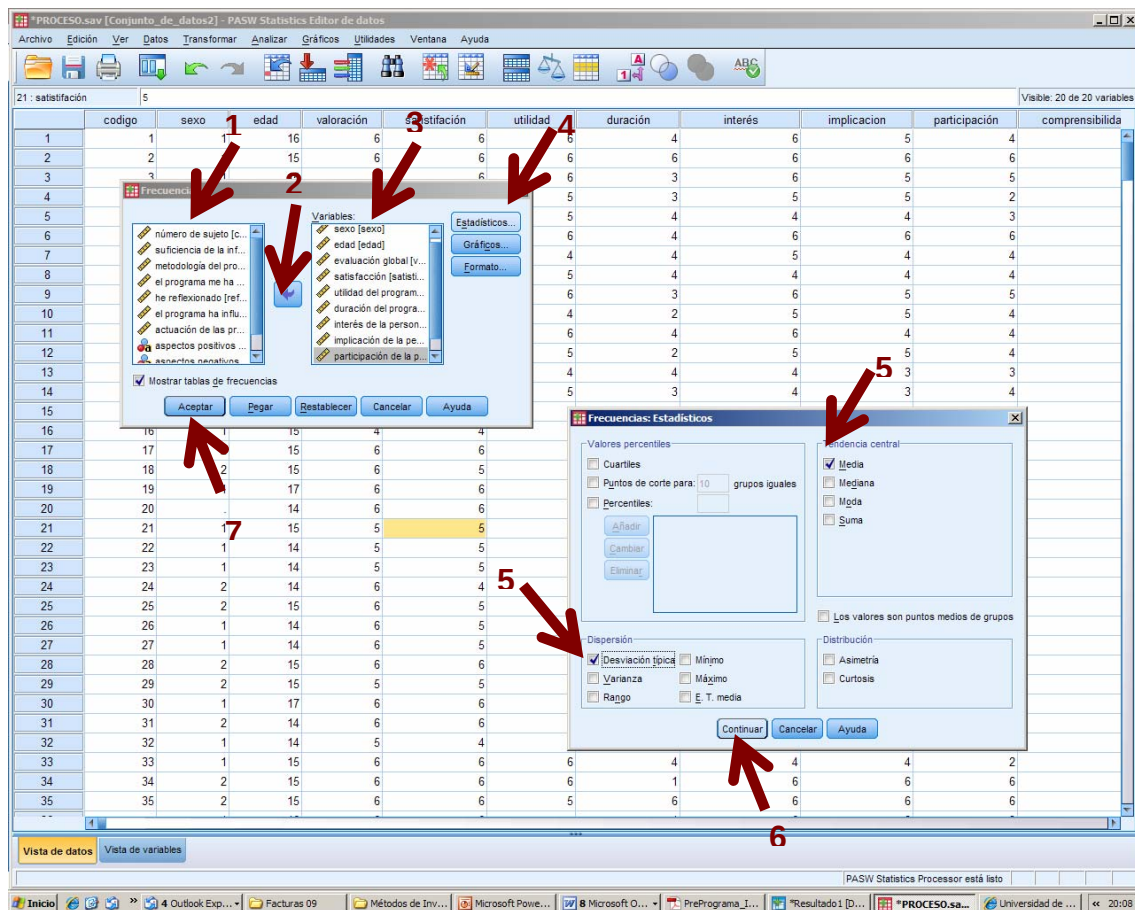


Figura 9

Para interpretar los resultados, a modo de ejemplo se observa que en la primera tabla que pone estadísticos, se indica la media y desviación típica de cada variable cuantitativa (ver figura 10). De forma que en la valoración global, los participantes han otorgado una media de 5,28 y una desviación típica de 0,81 (paso 1). Si tenemos en cuenta que el rango de respuestas posibles era de 1=totalmente negativo a 6= totalmente positivo, se deduce que los participantes han hecho una valoración muy positiva del programa. En las siguientes tablas vienen las frecuencias y porcentajes de personas que han respondido a cada categoría o rango de respuestas. Si nos fijamos en la variable de valoración global vemos que un 98,2% lo ha valorado de forma positiva (paso 2), por tanto el porcentaje de personas que lo ha valorado negativamente es mínimo. Y así se deberían interpretar cada uno de los aspectos o variables evaluadas. Hay que tener en cuenta que en variables de tipo cualitativo como el sexo, los estadísticos descriptivos que se utilizan es la frecuencia o el porcentaje de hombres y mujeres (la media no es el estadístico que se debe aplicar).

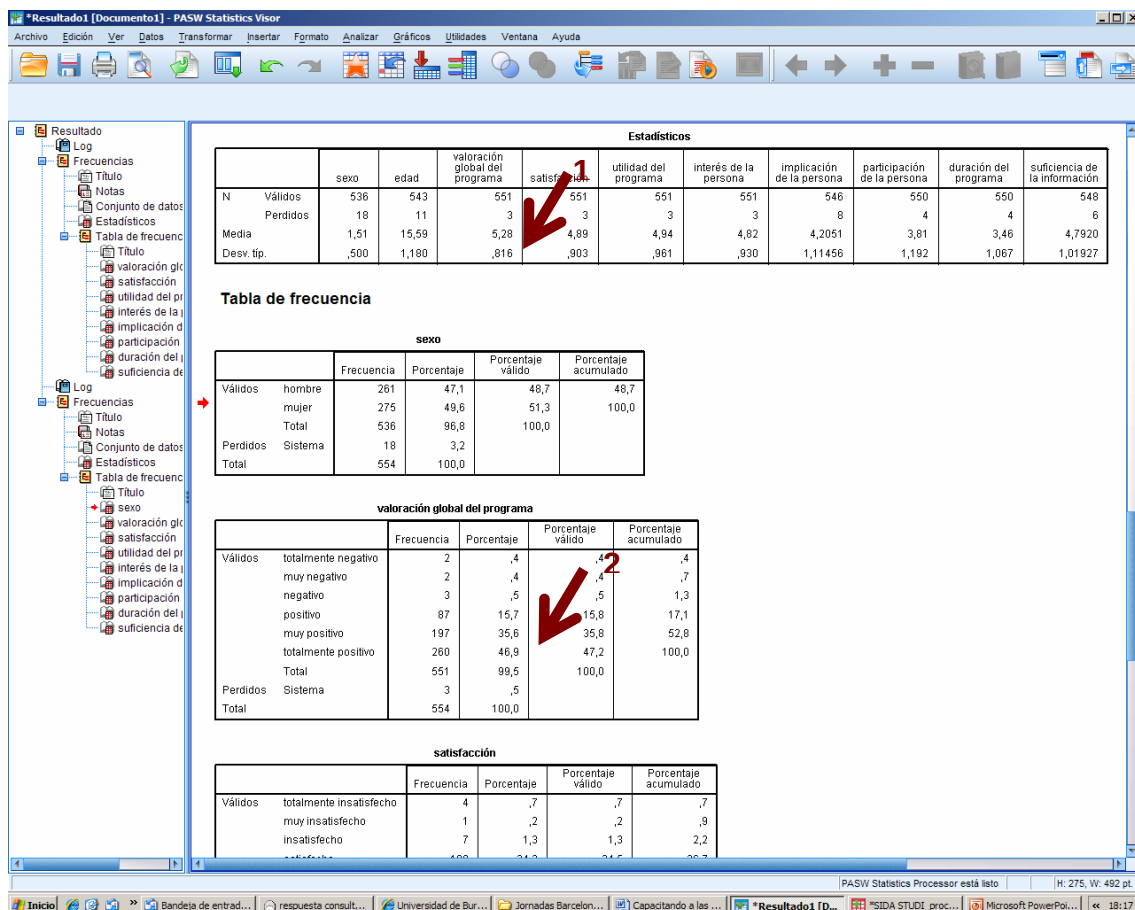


Figura 10

Finalmente, se debe elaborar un informe en el que se recoja toda la información que se desprende de la evaluación de proceso. Un posible índice de informe sería el siguiente (ver índice de informe de evaluación de resultados):

Índice del Informe

1. Descripción breve del programa.
2. Descripción de objetivos.
3. Descripción de la muestra (los datos que se hayan recogido, como el sexo, la edad...).
4. Descripción del procedimiento de la evaluación: instrumento de medida, procedimiento de administración y mención de los análisis aplicados.
5. Descripción de los resultados (redactar los resultados y ejemplificarlos con tablas o gráficos).
6. Conclusiones.

3.- Evaluación de Resultados

En general, la evaluación de resultados pretende determinar si se han alcanzado los efectos deseados y no otros, y si los resultados se pueden atribuir al programa con un grado de certidumbre razonable.

Como ya se ha indicado la evaluación de resultados se centra en el análisis valorativo a partir de tres criterios: eficacia (es decir si se logran los objetivos deseados), efectividad (si el programa tiene consecuencias directas múltiples) y eficiencia (si se logran resultados positivos a un coste razonable). Nosotros vamos hacer referencia al criterio de eficacia.

La metodología más utilizada en las evaluaciones de resultado es de carácter cuantitativo, con el fin de establecer resultados conseguidos y los estándares de ejecución iniciales. Sin rechazar la posibilidad de emplear otro tipo de métodos, se defiende el uso de la metodología experimental y cuasi-experimental, en cuyo caso, el diseño experimental debe ser contemplado y aceptado desde la fase de planificación.

Hoy por hoy para asegurarnos que los resultados que obtenemos son debidos al programa debemos aplicar los denominados diseños de evaluación. Aunque los diseños experimentales son los que presentan las mayores garantías de validez, aquí vamos a comenzar por dos diseños más sencillos: uno que hace referencia a un diseño pre-experimental y otro a un diseño cuasi-experimental. Hay que señalar que con estos diseños difícilmente se podrían obtener inferencias de causa-efecto ya que los resultados obtenidos podrían deberse a una gran cantidad de factores ajenos al programa, pero son más factibles y son una buena forma de adentrarse en el mundo de la evaluación de programas.

El diseño pre-experimental de un solo grupo con pruebas previa y posterior, en el que un único grupo de personas es medido en las variables objeto del programa antes y después de la aplicación de la intervención. Este diseño tiene muchas amenazas de validez. Este tipo de diseños tiene poca potencia debido a la gran cantidad de potenciales problemas de validez que plantea, de ahí que se use poco para generar inferencias causales. En general, este tipo de diseños usualmente sólo podría ser adecuado en aquellos casos en los que la variable medida sólo puede ser afectada por la intervención implementada, y que a la vez transcurra poco tiempo entre ambas medidas. En todo caso, las diferencias encontradas deberían ser tan evidentes que las potenciales

amenazas, por sí solas, no pudieran explicar el cambio ocurrido tras la aplicación del programa de intervención. A pesar de los inconvenientes de este diseño, nosotros vamos a exponer un ejemplo basado en este tipo de diseño ya que es el más factible de todos.

Otra posibilidad es el diseño cuasi-experimental de grupo control no equivalente que es el más utilizado en la práctica ya que, además de las ventajas que presenta frente al caso anteriormente descrito, es relativamente fácil de aplicar. Este diseño se caracteriza por:

1. La existencia de un grupo experimental (que lo forman el conjunto de personas que recibe el programa o intervención) y control (que lo forman aquellas personas que no reciben el programa, y que son lo más similares posibles al grupo experimental). Los grupos vienen asignados no se forman de forma aleatoria. Ejemplo, en un centro escolar contamos con 4 grupos de 4º de la ESO, dos los asignamos a los grupos experimentales y dos a los grupos control.
2. La recogida de información se hace antes (es decir los llamados pre-test) y después (es decir los llamados post-tests) de la realización del programa.

Para que los grupos control no se queden sin programa, se les da la intervención de forma postergada, es decir una vez que hayan cumplimentado las pruebas de evaluación (tanto las denominadas pre-test como post-test).

Con este diseño se favorece una mejor medida de las variables en cuanto las amenazas pueden ser más controlables, es decir, menos plausibles, aunque no imposibles.

Otra tercera propuesta de gran interés es el diseño de línea base no causal construida. En este caso, la media post-test del grupo experimental se compara con una muestra funcional o disfuncional. Este diseño cuasi-experimental se apoya en la comparación del cambio de medias de los grupos experimentales antes, y después de éste, en relación a una norma (generalmente los percentiles o deciles de una población comparable). De ahí su nombre de línea base (la distribución normativa del grupo de referencia) no causal (porque no se ha producido causalmente la diferencia entre el grupo experimental y el de control, sino que este último se usa como referencia pasiva, sin programa) y construida (porque generalmente la distribución normativa no existe y hay que construirla de una muestra del universo) (Ubillos 1996; Páez, Ubillos y León, 1996).

Para construir una matriz de datos de una muestra funcional o disfuncional con la que comparar nuestros resultados, se pueden aprovechar los datos obtenidos en el pre-test, cuando aplicamos el diseño pre-experimental mencionado previamente.

Este diseño plantea dos supuestos importantes: la ausencia del efecto espontáneo del paso del tiempo y la similitud entre el grupo experimental y la población de la que se extrae la distribución normativa. La primera premisa presupone que cuando no hay presencia de programa, la media en percentiles en el pre-test de un grupo control es igual a la media del post-test. El segundo postulado implica que los datos deben ser válidos para una población normativa similar a la del estudio, en cuanto a características socio-demográficas y culturales.

A continuación como en el caso de la evaluación de proceso, vamos a plantear un ejemplo de evaluación de resultados desde el punto de vista de la eficacia, y con la aplicación de un diseño pre-experimental de un solo grupo con pruebas previa y posterior. Los pasos a seguir serían los siguientes (Ubillos, 1995):

- 1) Fundamentación teórica y objetivos del programa.
- 2) Aspectos, contenidos o variables a evaluar.
- 3) Seleccionar o diseñar los instrumentos de medida.
- 4) Seleccionar un diseño de evaluación.
- 5) Administración de los instrumentos de medida.
- 6) Análisis e interpretación de datos.
- 7) Redacción del informe.

3.1.- Fundamentación teórica y objetivos del programa

El primer paso importante en este tipo de evaluaciones es que el programa se fundamente en teorías. En el caso de la prevención sexual del VIH/SIDA, las teorías que mejor explican la adopción de las conductas sexuales de riesgo son varias como (Ubillos, 2007): la Teoría de la Acción Razonada (Fishbein y Ajzen, 1975; Ajzen y Fishbein, 1980), la Teoría del Comportamiento Planificado (Ajzen, 1985, 2002), la Teoría del Aprendizaje Social (Bandura, 1986, 1997) y el Modelo de Creencias de

Salud (Becker, 1974; Janz y Becker, 1984). Existen otros modelos como son el Modelo de Información-Motivación-Habilidades Conductuales (Fisher y Fisher, 1992), el Modelo de la Motivación de la Protección (Floyd, Prentice-Dunn, y Rogers, 2000) o el Modelo Precede (Green y Kreuter, 1999) que contemplan de una manera u otra los mismos constructos. En síntesis vienen a indicar que *“la información es necesaria pero no suficiente para que las personas adopten conductas sexuales más seguras”*, planteando la importancia que juegan las creencias sobre el uso del preservativo, las actitudes hacia las medidas preventivas, la percepción de riesgo, la percepción de control y las habilidades de negociación, entre otros aspectos, a la hora de que las personas sean capaces de mantener relaciones sexuales más saludables (es decir tener relaciones coitales usando el preservativo). Si contamos con alguna teoría, de ahí se derivan los objetivos de nuestro programa.

Objetivo General
■ Disminuir las prácticas sexuales de riesgo, es decir disminuir el porcentaje de relaciones sexuales coitales sin uso del preservativo.
Objetivos Específicos
1. Aumentar los conocimientos sobre el uso del preservativo y la transmisión del VIH/SIDA. 2. Aumentar las creencias positivas hacia el uso del preservativo (beneficios) y disminuir las creencias negativas hacia el uso de esta medida preventiva (costes). 3. Mejorar la actitud positiva hacia el uso del preservativo, y reforzar la actitud negativa hacia el uso del coito interrumpido. 4. Aumentar las habilidades de negociación o de autoeficacia para que las personas posean recursos para proponer prácticas de sexo más seguro.

3.2.- Aspectos o contenidos a evaluar

De la teoría y de los objetivos se derivan los aspectos o contenidos que se van a abordar en el programa y que van a ser evaluados. Cuando estos aspectos o contenidos son operacionalizados para poder ser evaluados se denominan variables. Siguiendo con nuestro ejemplo serían los que se especifican en el siguiente cuadro.

Aspectos a trabajar y evaluar

1. Conductas sexuales de riesgo y de prevención
2. Conocimientos
3. Creencias (costes y beneficios) sobre el uso del preservativo.
4. Actitudes hacia el uso del preservativo y el coito interrumpido
5. Autoeficacia o habilidades de negociación.

3.3.- Seleccionar o diseñar los instrumentos de medida.

El siguiente paso consiste en diseñar un instrumento de medida que nos permita recoger información sobre estos aspectos. Existen multitud de escalas para evaluar estos aspectos. Las búsquedas bibliográficas a través de la red, en bases de datos especializadas, en artículos científicos, en libros... nos permiten localizar multitud de escalas que han sido previamente validadas y fiabilizadas, lo que nos da cierta garantía a la hora de evaluar nuestro programa. Se recomienda que se seleccionen estas escalas, que han sido validadas y fiabilizadas previamente, porque el diseñar instrumentos de evaluación que sean válidos y fiables requiere un manejo metodológico y estadístico muy elevado.

A modo de ejemplo, se van a incluir en este apartado una serie de escalas ampliamente utilizadas en este campo (para más información véase Ubillos, 2005; Ubillos, Insúa y De Andrés, 1999, manual accesible en <http://www.pnsd.msc.es/Categoria2/publica/publicaciones/home.htm>). En estos manuales se da información sobre las escalas que se presentan en este apartado. Aquí hemos seleccionado una versión abreviada de la mayoría de las escalas para facilitar su aplicación. Hemos incluido una escala de evaluación de conocimientos, otra sobre creencias derivadas del uso del preservativo, otra sobre percepción de control y autoeficacia (habilidades de comunicación) y, finalmente, una de conductas sexuales preventivas y/o de riesgo. Es decir, hemos seleccionado una escala por cada una de las variables que se van a abordar en la intervención y que van a ser evaluadas.



Lahiria-Nahia Sexologia Elkarteak

ESCALA DE CONOCIMIENTOS

A continuación, responde marcando con una x la respuesta que consideres correcta, es decir si cada una de las afirmaciones es verdadera, falsa o no sabes.

➤ Para usar bien el preservativo es necesario:

1. Colocarlo antes de la erección	Verdadero	Falso	No Sé
2. Se puede abrir rompiendo la envoltura del plástico con la boca	Verdadero	Falso	No Sé
3. Puede tener efectos secundarios	Verdadero	Falso	No Sé
4. Colocarlo cuando el pene está en erección	Verdadero	Falso	No Sé
5. Colocarlo justo en el momento de la eyaculación	Verdadero	Falso	No Sé
6. Inflarlo de aire para comprobar si está pinchado	Verdadero	Falso	No Sé
7. Dejar un espacio vacío en la punta del preservativo	Verdadero	Falso	No Sé
8. Mirar la fecha de caducidad	Verdadero	Falso	No Sé
9. No manipularlo hasta su colocación	Verdadero	Falso	No Sé
10. Sujetar el condón por los bordes y retirar el pene de la vagina, justo después de la eyaculación	Verdadero	Falso	No Sé
11. Utilizar uno nuevo cada vez que el chico introduce el pene en la vagina	Verdadero	Falso	No Sé
12. Esperar dentro de la vagina a que el pene vuelva a su estado normal y después retirar el condón sujetándolo por los bordes	Verdadero	Falso	No Sé

➤ Según tu opinión, el SIDA se puede "pillar":

13. A través de las relaciones coitales	Verdadero	Falso	No Sé
14. Esnifando cocaína compartiendo el rulo	Verdadero	Falso	No Sé
15. Esnifando cocaína sin compartir el rulo			
16. Donando sangre	Verdadero	Falso	No Sé
17. Utilizando piscinas públicas	Verdadero	Falso	No Sé
18. Recibiendo sangre (transfusión)	Verdadero	Falso	No Sé

19. A través de los cubiertos, vasos.	Verdadero	Falso	No Sé
20. Besándose en la boca	Verdadero	Falso	No Sé
21. Inyectándose heroína siempre con la propia jeringuilla.	Verdadero	Falso	No Sé
22. Inyectándose heroína compartiendo la jeringuilla	Verdadero	Falso	No Sé
23. A través de picaduras de insecto	Verdadero	Falso	No Sé
24. A través de la saliva	Verdadero	Falso	No Sé
25. A través de las lágrimas	Verdadero	Falso	No Sé
26. Compartiendo objetos personales (cepillo de dientes, cuchillas...)	Verdadero	Falso	No Sé
27. A través del coito interrumpido	Verdadero	Falso	No Sé
28. De mujer embarazada a su hijo.	Verdadero	Falso	No Sé
➤ Para no "pillar" el Sida, es necesario:			
29. Acudir al médico para realizarse controles periódicos	Verdadero	Falso	No Sé
30. Estar bien informado sobre lo que es el Sida	Verdadero	Falso	No Sé
31. Tener siempre la misma pareja sexual	Verdadero	Falso	No Sé
32. Utilizar el preservativo	Verdadero	Falso	No Sé
33. No utilizar agujas ya usadas (piercing, tatuajes, etc.)	Verdadero	Falso	No Sé
34. Tener relaciones sexuales sólo con gente conocida	Verdadero	Falso	No Sé
35. Evitar el contacto con sangre de otros.	Verdadero	Falso	No Sé
36. Tomar la píldora anticonceptiva	Verdadero	Falso	No Sé
37. Fijarse en el aspecto exterior de la pareja	Verdadero	Falso	No Sé
38. Utilizar cremas espermicidas	Verdadero	Falso	No Sé
39. Informarse el estado de salud de la pareja	Verdadero	Falso	No Sé
40. No existen medidas que nos protejan	Verdadero	Falso	No Sé



Lafia-Nafia Sexología Elkartea

ESCALA DE CREENCIAS DEL PRESERVATIVO (COSTES Y BENEFICIOS)

Desde tu punto de vista, señala tu grado de acuerdo o desacuerdo con las siguientes características relacionadas con los preservativos. Responde marcando con un círculo o con una x el número de la escala que mejor refleje tu respuesta, es decir tu grado de acuerdo o desacuerdo con las siguientes afirmaciones:

1= Totalmente en desacuerdo

4= De Acuerdo

2= Muy en desacuerdo

5= Muy de acuerdo

3= En desacuerdo

6= Totalmente de acuerdo

LOS PRESERVATIVOS:

1. Previenen el embarazo	1	2	3	4	5	6
2. Engorroso, incómodo, complicado de usar	1	2	3	4	5	6
3. Tranquilizan y dan seguridad en la relación	1	2	3	4	5	6
4. Crea dudas en el otro acerca de tu estado de salud	1	2	3	4	5	6
5. Rompen con el romanticismo de la situación	1	2	3	4	5	6
6. Interrumpe el acto sexual	1	2	3	4	5	6
7. Da corte comentar que hay que "ponerlo"	1	2	3	4	5	6
8. Previenen de enfermedades de transmisión sexual y SIDA	1	2	3	4	5	6
9. Disminuye el placer	1	2	3	4	5	6
10. Son fáciles de obtener	1	2	3	4	5	6
11. Hay que prever su uso, hay que pensar en tenerlo	1	2	3	4	5	6
12. Me preocupa que lo encuentren en mi casa	1	2	3	4	5	6
13. Son sencillos y fáciles de utilizar	1	2	3	4	5	6
14. Se rompen fácilmente, son frágiles	1	2	3	4	5	6
15. Su colocación es un juego erótico más	1	2	3	4	5	6
16. Me da vergüenza comprarlo	1	2	3	4	5	6
17. Si los llevas encima dan la sensación de que tienes relaciones con cualquier persona	1	2	3	4	5	6
18. Los chicos que los tienen dan la sensación de que buscan llegar a la relación sexual siempre	1	2	3	4	5	6



Lafia-Nafia Sexología Elkartea

ESCALA DE PERCEPCIÓN DE CONTROL Y AUTOEFICACIA

➤ Antes de contestar señala con una cruz el tipo de relación que mantienes en este momento:

- 0 Ninguna
- 1 Pareja/Novio estable
- 2 Pareja/Novio inestable
- 3 Varias Parejas estables
- 4 Varias Parejas inestables

➤ Contesta independientemente de que tengas o no experiencia sexual:

- Si tienes experiencia sexual, contesta recordando tu última relación sexual coital.
- Si no tienes experiencia sexual coital, contesta como si estuvieras en esa situación.

Responde marcando con un círculo o con una x el número de la escala que mejor refleje tu respuesta, es decir tu grado de acuerdo o desacuerdo con las siguientes afirmaciones:

1= Totalmente en desacuerdo	4= De Acuerdo					
2= Muy en desacuerdo	5= Muy de acuerdo					
3= En desacuerdo	6= Totalmente de acuerdo					
1.- Cuando estoy con una pareja, me siento capaz de controlar lo que ocurra sexualmente con él o ella.	1	2	3	4	5	6
2.- Si una pareja y yo estamos poco excitados sexualmente y yo no quiero llegar al coito, puedo decirle fácilmente que no.	1	2	3	4	5	6
3.- Si una pareja y yo estamos muy excitados sexualmente y yo no quiero llegar al coito, puedo parar las cosas fácilmente, por lo tanto, no lo haremos.	1	2	3	4	5	6
4.- Hay veces que me siento tan implicada/o personal y emocionalmente que puedo tener relaciones coitales incluso sin usar un método anticonceptivo (condón, pastillas...).	1	2	3	4	5	6
5.- A veces sigo adelante con lo que mi pareja quiere hacer sexualmente, porque siento que no puedo tomar las riendas de la situación y decir lo que yo quiero.	1	2	3	4	5	6
6.- Hay veces que me siento incapaz de hablar con mi pareja sobre la necesidad de que usemos un método anticonceptivo (condón, pastillas...).	1	2	3	4	5	6

- Recuerda la última vez que estuviste con una pareja teniendo relaciones sexuales, os encontrabais muy excitados y yendo claramente hacia el coito. ¿Has estado alguna vez en esta situación?:

0 NO

1 SI

En caso de que no, imagina que estás en esa situación y contesta.

7.- Puedo/pude preguntarle fácilmente a mi pareja si usa algún método anticonceptivo o decirle que yo no uso ninguno (condón, pastillas...).	1	2	3	4	5	6
8.- Puedo/pude parar la situación para pedirle que se ponga un preservativo o para ponérmelo yo mismo.	1	2	3	4	5	6
9.- Puedo/pude negarme a seguir con la relación sexual coital si no estamos usando un método anticonceptivo (condón, pastillas...).	1	2	3	4	5	6
10.- Puedo/pude plantearle fácilmente a mi pareja otras formas de relación sexual, distintas al coito, si no estamos usando un método anticonceptivo (condón, pastillas...).	1	2	3	4	5	6



Lafia-Nafia Sexología Elkartea

ESCALA DE COMPORTAMIENTO SEXUAL Y PREVENTIVO

Para responder a las preguntas que se te indican a continuación, marca con un círculo o con una x aquel número que mejor refleje tu respuesta.

1.- NIVEL DE EXPERIENCIA SEXUAL: <i>A continuación se describen cinco niveles de experiencia sexual. Lee atentamente cada uno de ellos y pon una X en el nivel que tú te encuentres. Al margen de tus relaciones sexuales actuales indica la experiencia máxima que has tenido a lo largo de tu vida.</i>	1 No he tenido ninguna experiencia sexual. 2 He besado en los labios y abrazado a una chica/o con caricias en las manos y por encima de la ropa. 3 He tenido intimidades sexuales, próximas al coito pero sin llegar a él. Esto incluye haber acariciado directamente y haber sido acariciado/a directamente en cualquier parte del cuerpo, incluido los genitales, con o sin orgasmos. 4 He tenido experiencia de haber llegado al coito, con una sola persona a lo largo de mi vida. 5 He tenido relaciones coitales con más de una persona a lo largo de mi vida.
2.- En LOS ÚLTIMOS 4 MESES cómo es tu comportamiento sexual:	1 No tengo relaciones sexuales. 2 Tengo relaciones sexuales que consisten en besos, caricias, caricias en los genitales, etc..., pero no tengo experiencia de coito. 3 Tengo relaciones sexuales que incluyen el coito.
3.- Sólo responde si has tenido o tienes algún tipo de relación sexual (besos, caricias, <i>petting</i>...), ¿tienes o has tenido relaciones sexuales?:	1 Sólo con personas del otro sexo. 2 Normalmente con personas del otro sexo y alguna vez con personas del mismo sexo. 3 Tanto con personas del otro sexo como con personas del mismo sexo. 4 Normalmente con personas del mismo sexo y alguna vez con personas del otro sexo. 5 Sólo con personas del mismo sexo.
4.- ¿Qué intención tienes de utilizar el preservativo en una próxima relación sexual?	1 Ninguna intención 2 Alguna intención 3 Poca intención 4 Bastante intención 5 Mucha intención 6 Total intención

Las preguntas que van desde el número 5 hasta el 10 sólo deben ser respondidas por las personas que mantienen o han mantenido relaciones sexuales que incluyen el coito (la penetración).	
5.- Cuando mantienes relaciones sexuales coitales, ¿utilizas algún método preventivo?:	1 SIEMPRE 2 A VECES 3 NUNCA
6.- ¿Qué método utilizas? (Puedes señalar más de una respuesta).	1 El coito interrumpido (La marcha atrás, Sacarlo antes de eyacular) 2 Píldoras anticonceptivas 3 Preservativo (Condomes) 4 Otros, indica cuál _____
7.- En el último mes, ¿cuántas veces has mantenido relaciones sexuales coitales? (Indica aproximadamente el número de veces):	Nº VECES _____
8.- De estas veces, ¿cuántas veces has utilizado un método preventivo? (Indica aproximadamente el número de veces):	Nº VECES _____
9.- ¿Qué método habéis utilizado tú o tu pareja en estas ocasiones? (Si es más de un método señálalo):	_____
10.- ¿Por qué has elegido ese/os método/s preventivo/s? (Indica desde tu punto de vista la razón más importante. Sólo elige una respuesta):	1 Prevenir el embarazo 2 Prevenir una enfermedad de transmisión sexual 3 Prevenir el SIDA 4 Otros _____

3.4.- Seleccione un diseño de evaluación.

Para evaluar los resultados e impacto de un programa son necesarios los diseños de investigación con el fin de garantizar la validez de nuestros resultados. Nosotros nos vamos a centrar en la validez interna, es decir, aquella que pretende demostrar de un modo inequívoco que existen o no una serie de resultados debidos al programa y no a otros factores o variables extrañas. Los diseños de investigación se dividen en 4 grupos:

1. Los diseños experimentales son los que presentan mayores garantías de validez.
2. Los diseño pre-experimentales son los que presentan mayores problemas a la hora de controlar las VV. extrañas.
3. Los diseños cuasi-experimentales presentan problemas en el control de las VV. extrañas.

4. Los diseños cualitativos-naturalistas que no nos dan información sobre las relaciones de causalidad, sino información sobre las experiencias de las personas y nos ayudan a comprender en detalle la implementación del programa.

El **diseño experimental** se caracteriza por dos condiciones:

1. La existencia de un grupo experimental (que lo forman el conjunto de personas que recibe el programa o intervención) y control (que lo forman aquellas personas que no reciben el programa, y que son lo más similares posibles al grupo experimental). Los grupos son creados por asignación aleatoria.
2. La recogida de información se hace antes (es decir los llamados pre-test) y después (es decir los llamados post-tests) de la realización del programa.

La aleatorización es una técnica para asegurar la comparabilidad de los grupos experimental y control; es la propuesta clave para garantizar que los grupos de comparación sean lo más similares o equivalentes posibles, y poder descartar las VV. extrañas relacionadas con las diferencias individuales. En estos diseños se comparan los resultados obtenidos en el post-test por ambos grupos con las puntuaciones obtenidas en el pre-test.

Los **diseños pre-experimentales** presentan los principios a partir de los cuales se han desarrollado diseños más complejos, como los cuasi-experimentales. Con estos diseños difícilmente se podrían obtener inferencias de causa-efecto ya que los resultados obtenidos podrían deberse a una gran cantidad de las potenciales amenazas a la validez. Nosotros vamos a hacer referencia al **diseño de un solo grupo con pruebas previa y posterior**, en el que un único grupo de personas es medido en las variables objeto del programa antes y después de la aplicación de la intervención. Este tipo de diseños tiene poca potencia debido a la gran cantidad de potenciales problemas de validez que plantea, de ahí que se use poco para generar inferencias causales. En general, este tipo de diseños usualmente sólo podría ser adecuado en aquellos casos en los que la variable medida sólo puede ser afectada por la intervención implementada, y que a la vez transcurra poco tiempo entre ambas medidas. En todo caso, las diferencias encontradas deberían ser tan evidentes que las potenciales amenazas, por sí solas, no pudieran explicar el cambio ocurrido tras la aplicación del programa de intervención. No obstante este diseño es uno de los más factibles cuando no se tiene formación específica en evaluación o faltan recursos.

Cuando los diseños experimentales no se consideren una buena opción, y los pre-experimentales no son muy aconsejables se puede recurrir a diversas alternativas dentro de los **diseños cuasi-experimentales**. Estos diseños son habitualmente utilizados en evaluación de programas sociales y sanitarios, ya que aportan información relevante para el estudio de relaciones causales entre variables de interés en contextos de intervención. **El diseño de grupo control no equivalente** es el más utilizado en la práctica ya que, además de las ventajas que presenta frente a los casos anteriormente descritos, es relativamente fácil de implementar. En los diseños de grupo control no equivalente la lógica es semejante a la de los diseños experimentales de grupos aleatorios, aunque mantiene con ellos una diferencia esencial: la formación de los grupos de comparación se realiza por procedimientos distintos a la asignación aleatoria; la regla de asignación no es conocida. Sin embargo, el evaluador debe aplicar los procedimientos necesarios para tratar que los grupos que se comparen sean los más similares posibles. El diseño implica realizar registros antes y después de la intervención en ambos grupos, es decir, tanto en el grupo objeto de la intervención como en el grupo control no equivalente. En este diseño se toman las medidas previas con objeto de evaluar la posible no equivalencia entre ambos grupos, de este modo se dispone de un referente más a partir del cual abordar el estudio de uno de los problemas principales de los estudios cuasi-experimentales, es decir la comparabilidad de los grupos conformados. Campbell y Stanley (1966) plantaban el uso de grupos naturales de los que se sabe que son similares entre sí, como p.e., los grupos de clase consecutivos de 3º de ESO en los colegios.

Otro diseño cuasi-experimental es el **diseño de línea base no causal construida**. En este diseño la media post-test del grupo experimental se compara con una muestra funcional o disfuncional. Este diseño se apoya en la comparación del cambio de medias de los grupos experimentales antes, y después de éste, en relación a una norma (generalmente los percentiles o deciles de una población comparable). De ahí su nombre de línea base (la distribución normativa del grupo de referencia) no causal (porque no se ha producido causalmente la diferencia entre el grupo experimental y el de control, sino que este último se usa como referencia pasiva, sin programa) y construida (porque generalmente la distribución normativa no existe y hay que construirla de una muestra del universo). Este diseño plantea dos supuestos importantes: la ausencia del efecto espontáneo del paso del tiempo y la similitud entre el grupo

experimental y la población de la que se extrae la distribución normativa. La primera premisa presupone que cuando no hay presencia de programa, la media en percentiles en el pre-test de un grupo control es igual a la media del post-test. El segundo postulado implica que los datos deben ser válidos para una población normativa similar a la del estudio, en cuanto a características socio-demográficas y culturales.

3.5.- Administración de los instrumentos de medida.

El siguiente paso consiste en la administración de los instrumentos de medida en base al diseño de evaluación previsto. Por ejemplo, si contamos con un diseño pre-experimental de grupo único con pruebas antes y después, se deberán pasar estas escalas antes del programa (pre-test) y después del programa (post-test). Por ello, es importante que tengamos previsto el diseño de evaluación, antes de empezar a aplicarlo. Las recomendaciones sobre las estrategias que se deben utilizar para administrar las escalas son las mismas que las mencionadas en la evaluación de proceso.

3.6.- Análisis e Interpretación de los datos

Una vez que los participantes han contestado a las escalas, se deben introducir en una base de datos. En el caso de utilizar el SPSS se deberán seguir los mismos pasos que los expuestos en la evaluación de proceso. En este caso lo que se debe tener en cuenta que, en base al diseño pre-experimental aplicado en nuestro ejemplo, contamos con que cada sujeto tiene datos de todas las escalas administradas en el pre-test y post-test. Por ello, a la hora de introducir los datos, como ya hemos indicado, a cada sujeto se le asigna una fila, de forma que cuando hemos introducido todos los datos del pre-test, seguimos introduciendo los del post-test. Para diferenciar los aspectos o variables del pre-test y post-test, se recomienda que el nombre utilizado para la variable sea el mismo pero le añadamos una a (pre-test) o una b (post-test). Por ejemplo, conocimiento1a (se puede referir a nuestra primera pregunta de conocimientos en el pre-test “Para usar bien el preservativo hay que colocarlo antes de la erección”) y conocimiento1b (se referirá a nuestra primera pregunta de conocimientos en el post-test).

Una vez dispongamos de la matriz de datos y hallamos introducido los datos pasamos al análisis de datos e interpretación de resultados. Voy a exponer un ejemplo con una de las escalas, en concreto con la de costes y beneficios. Primero se deben crear

los indicadores globales de costes y beneficios, para ello se hace clic en la opción superior que pone “transformar” (paso 1), y se pincha en la primera opción “calcular variables” (paso 2) (ver figura 11), de forma que aparece un cuadro de diálogo.

En ese cuadro de diálogo en el ángulo superior izquierdo indicamos el nombre que le vamos a dar al indicador, p.e. “costesa” (paso 3). Hacemos clic en donde indica “tipo y etiqueta” (paso 4), salta una ventana donde le indicamos el nombre de la variable, por ejemplo “costes en el tiempo 0 (pre-test)” (paso 5) y le damos a continuar (paso 6) (ver figura 12), de forma que volvemos al cuadro de diálogo anterior.

En la ventana inferior izquierda vienen todas las variables que hemos introducido en la matriz, se seleccionan cada una de las variables que corresponden a los costes de la escala (paso 7), se pincha en la flecha para situarlas automáticamente en la ventana de “expresión numérica” (paso 8), se las suma (paso 9), se les pone entre paréntesis (paso 10) y se las divide por el número de preguntas, en nuestro caso son 12 (paso 11) para que el rango de respuesta de este indicador vuelva a situarse entre 1 y 6 (ver figura 13). Esto se hace tanto con las variables del pre como con las del post-test.

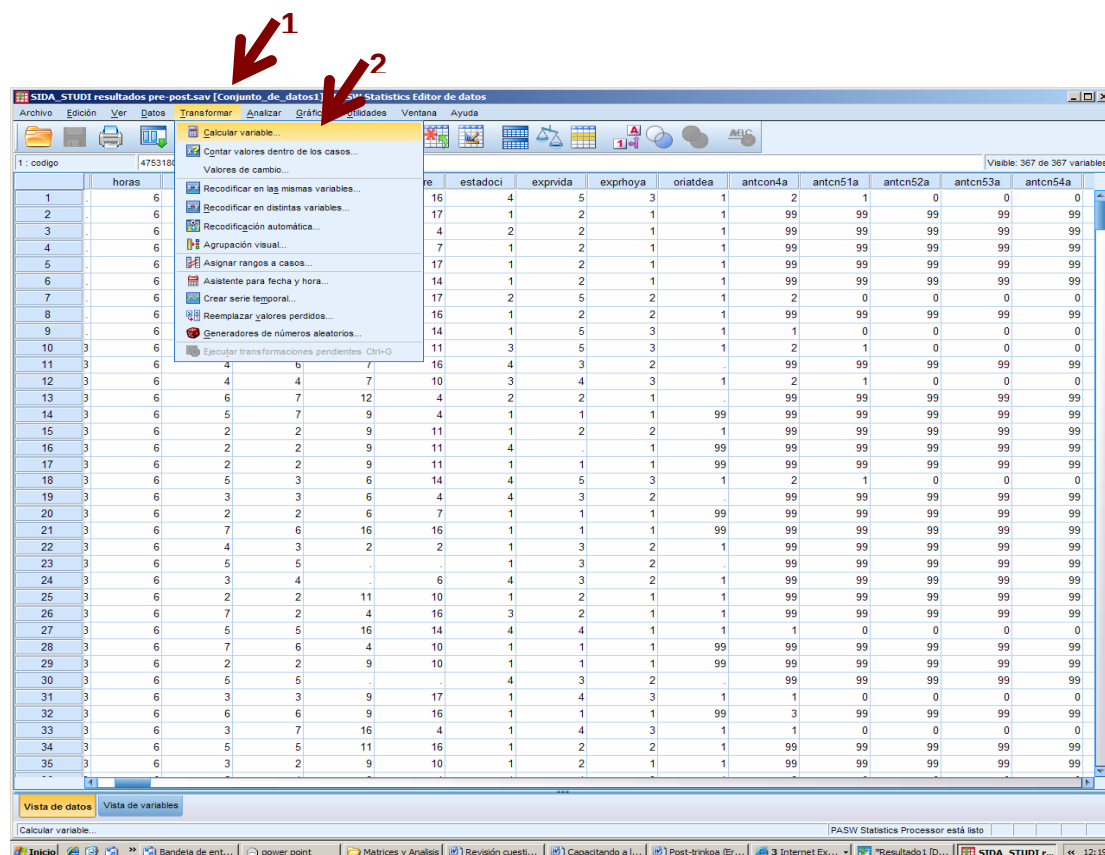


Figura 11

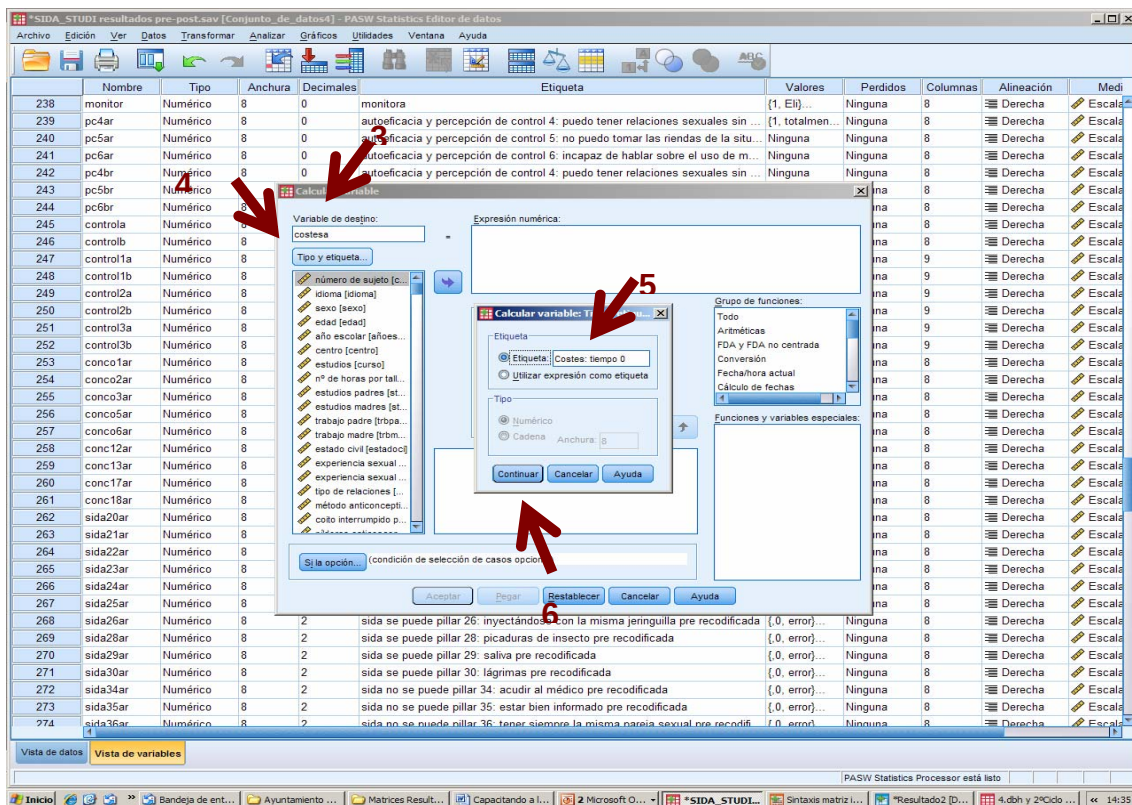


Figura 12

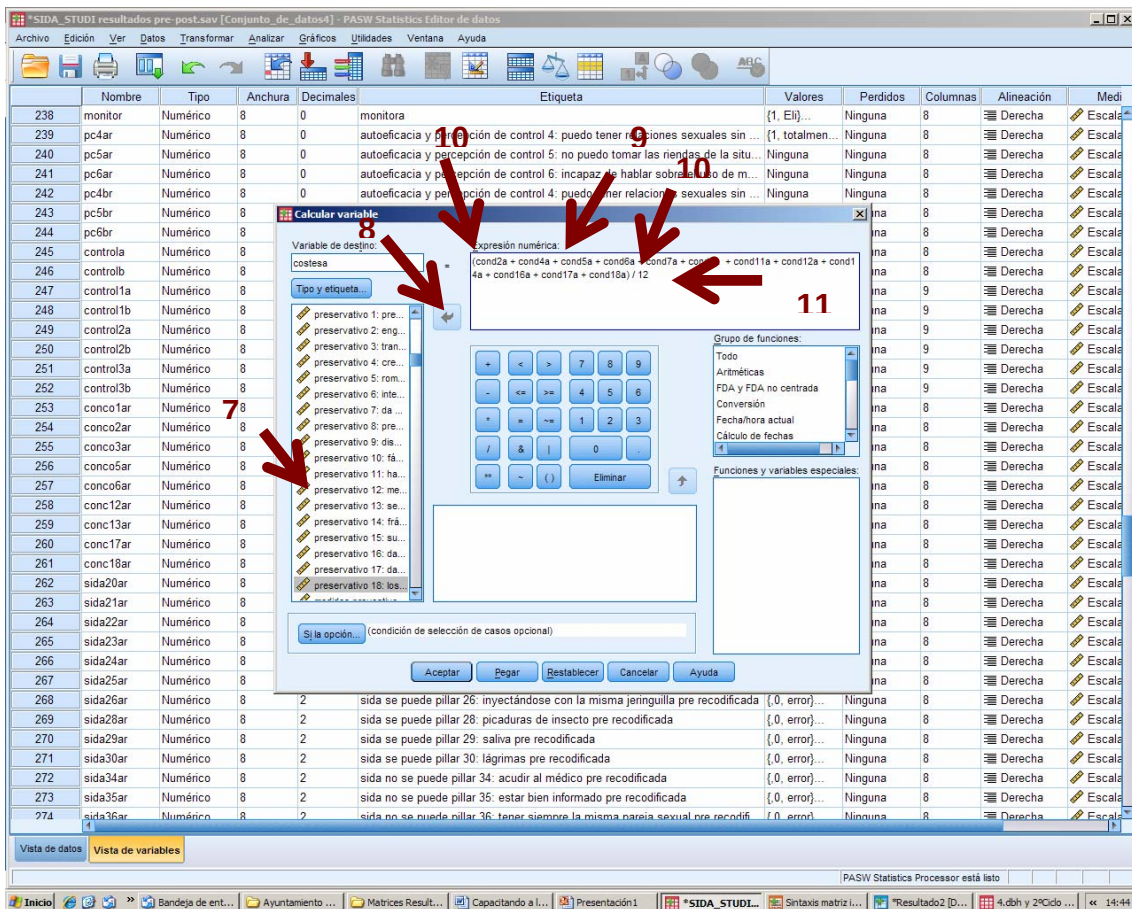


Figura 13

Una vez que tengamos creados los indicadores se pasa a lanzar el análisis. En este caso lo aconsejable es un t-test que compara si hay una diferencia entre las puntuaciones medias del pre-test y del post-test. De forma específica se aplica la prueba T para muestras relacionadas ya que compara las medias de dos variables de un solo grupo, como en nuestro ejemplo. Es decir, nosotros queremos comparar la media obtenida por el grupo en el pre-test con la puntuación media obtenida en el post-test.

Como se observa en la figura 14, se hace clic en analizar (paso 1), se hace clic en la opción de comparar medias (paso 2), lo que da lugar a que se desplieguen otras opciones y se hace clic en la de prueba test para muestras relacionadas (paso 3), de forma que aparecerá un cuadro de diálogo.

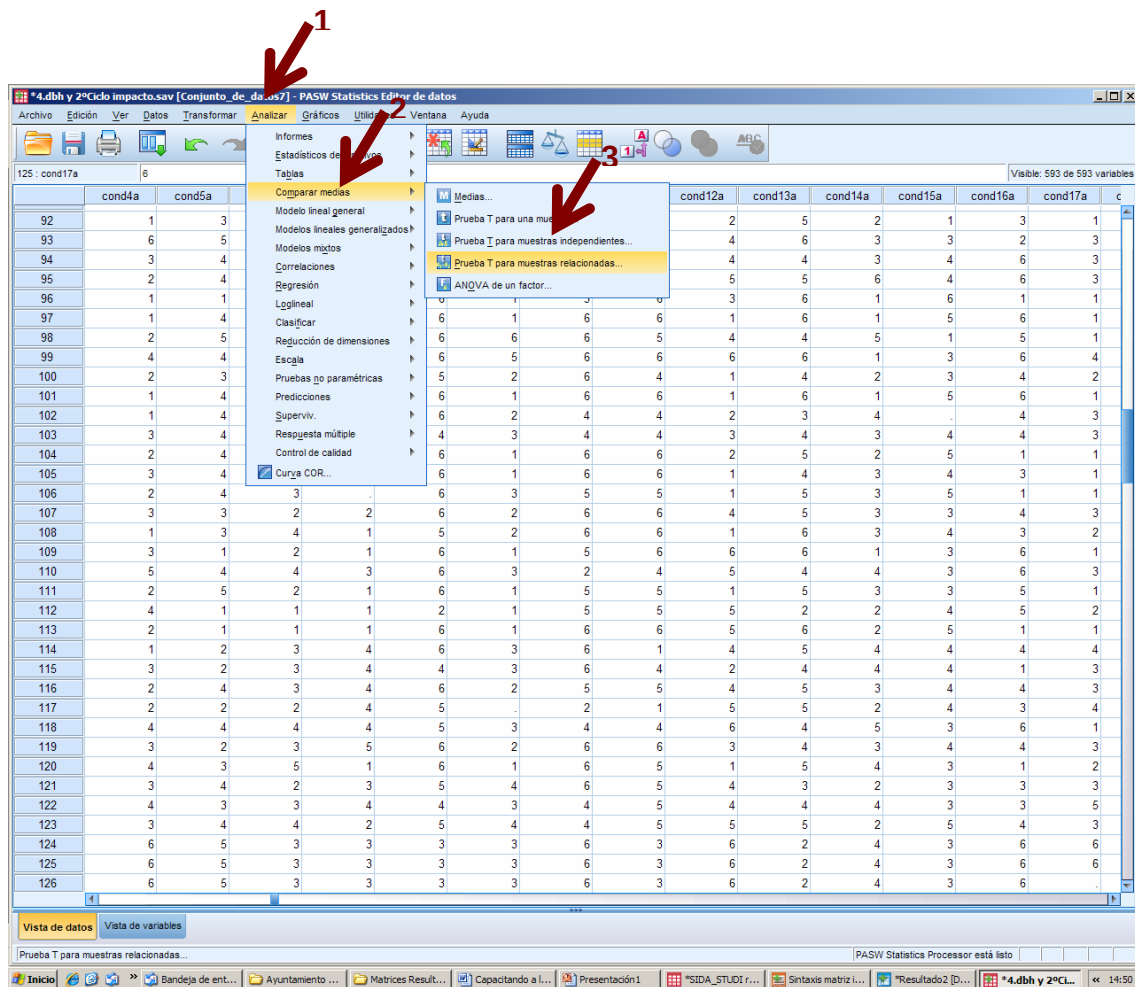


Figura 14

A continuación, en el cuadro de diálogo seleccionamos los dos indicadores, cuyas medidas queremos comparar, en nuestro ejemplo los “costesa” (costes en el pre-test) con “costesb” (costes en el post-test) (paso 1), hacemos clic en la flecha y de esta forma se sitúan en la ventana de variables emparejadas (paso 2), y le hacemos clic a aceptar (paso 3) (ver figura 15). De forma que se abrirá otro archivo con los resultados, que está conectado con nuestra matriz de datos. De forma que así podemos comparar la media obtenida por nuestro grupo o grupos en el pre-test con la media obtenida en el post-test. Esto nos dará a conocer si la puntuación media después del programa ha mejorado con respecto a la media obtenida antes del programa.

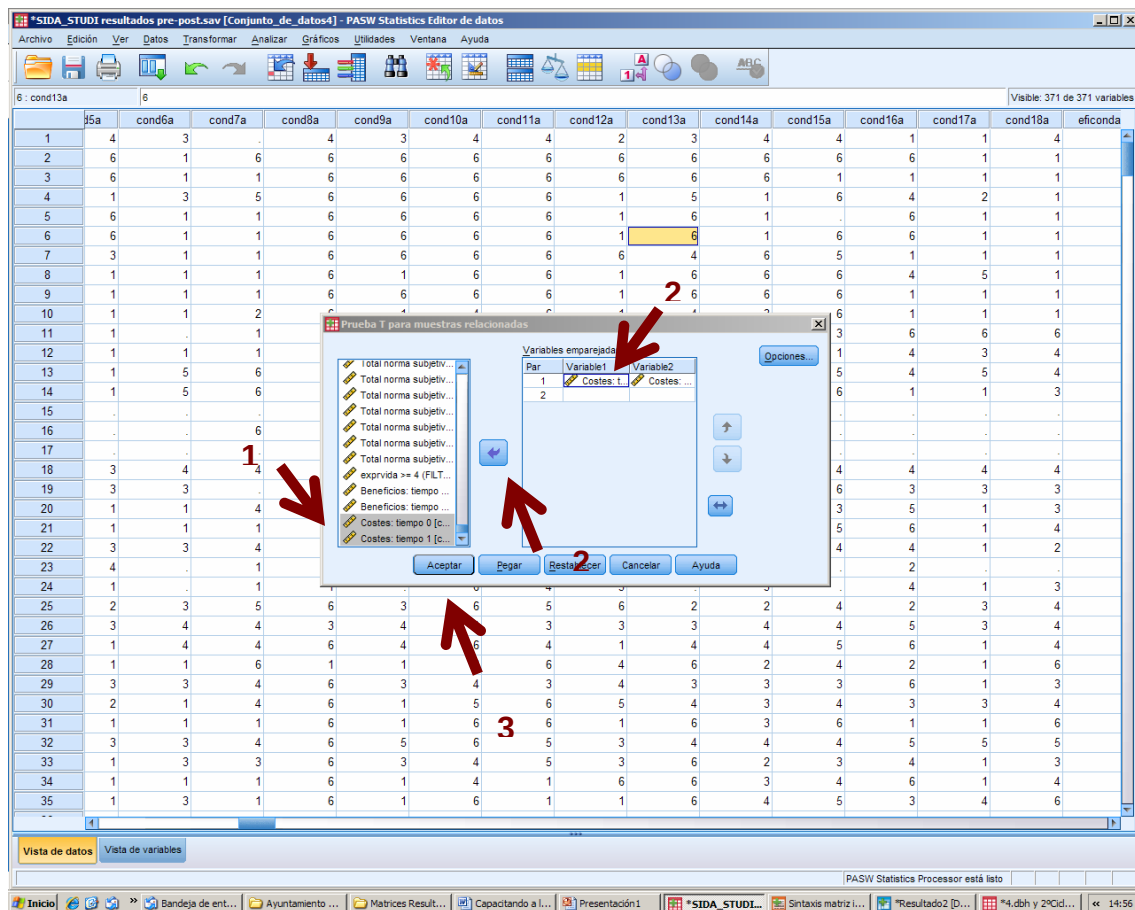


Figura 15

Los resultados se interpretan de la siguiente manera. Primero hay que mirar lo que se denomina significatividad (bilateral) o p. Si el valor de la significatividad o de la p es menor que 0,050, significa que hay una diferencia estadísticamente significativa entre ambas medias (la del pre-test y la del post-test). Si el valor de p es mayor que 0,050,

significa que no hay una diferencia estadísticamente significativa entre ambas medias. Una vez comprobado, si hay diferencias se pasa a analizar las dos medias. En nuestro caso, se observa que la media obtenida por el grupo en el pre-test es 3,12 (dt=0,63) y la del post-test 2,82 (dt=0,66). Por tanto, la puntuación media del post-test es significativamente menor que la del pre-test, es decir después del programa los participantes perciben menos costes derivados del uso del preservativo. La forma de registrar este resultado en los informes es la siguiente: $t_{(120)}=5,29$, $p<.0001$ [$t_{(gl \text{ o grados de libertad})}$ =valor de la t, p =valor de la significatividad] (paso 3) (ver figura 16). Es importante tener en cuenta que con este tipo de diseño no podemos afirmar que el programa sea el que ha producido esta mejora en la percepción de los costes, sino que debemos ser cautelosos y concluir que: “después de la intervención se observa una mejoría significativa en la percepción de costes derivados del uso del preservativo”.

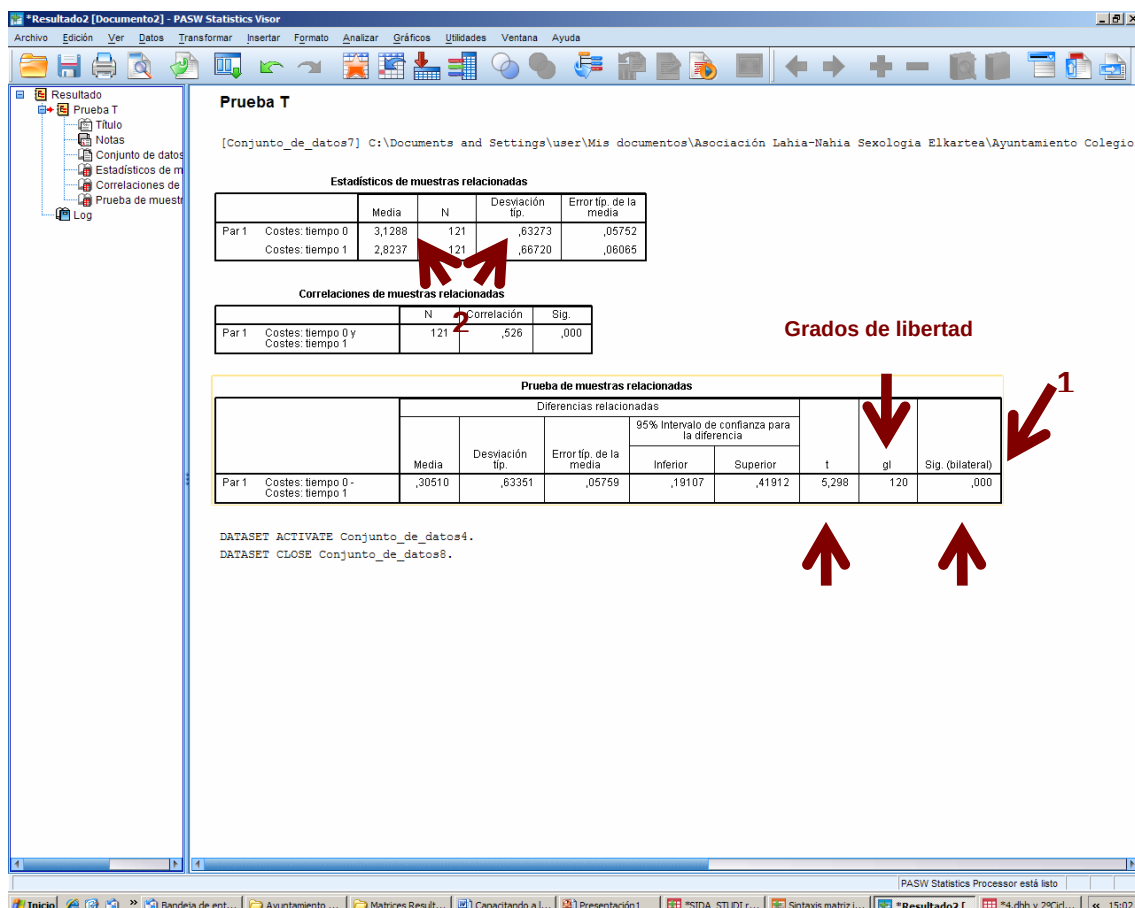


Figura 16

3.6.- Redacción del Informe

Finalmente, se debe elaborar un informe en el que se recoja toda la información que se desprende de la evaluación de resultados. Este documento puede ser de uso interno o externo y puede tener múltiples funciones, entre otras podemos mencionar las siguientes:

- Para presentar todo el proceso y los resultados obtenidos a los usuarios del programa o de los servicios.
- Para presentar la documentación que justifique el trabajo realizado a las posibles entidades que financian el programa.
- Para analizar de forma interna cuáles son los puntos débiles y fuertes del programa, a fin de mejorarlo.

En el caso de haber aplicado una evaluación de proceso y de resultados, un posible índice de informe sería el siguiente:

Índice del Informe
1. Descripción breve del programa.
2. Descripción de objetivos (proceso y resultado).
3. Descripción de la muestra (los datos que se hayan recogido, como el sexo, la edad...).
4. Descripción del procedimiento de la evaluación de proceso y de resultados: instrumento de medida, procedimiento de administración y mención de los análisis aplicados.
5. Descripción de los resultados: proceso y resultados (redactar los resultados y ejemplificarlos con tablas o gráficos).
6. Conclusiones.

4.- Condiciones básicas de una buena evaluación

El Joint Committee on Standards for Educational Evaluation (1981) se refiere a las condiciones básicas que ha de cumplir una buena evaluación. Las más relevantes son (Fernández Ballesteros y Hernández, 1995):

a) Que sea útil, es decir, que esté en relación y se corresponda con lo que necesitan las personas implicadas en el programa a evaluar (responsables, profesionales, participantes, usuarios de la evaluación o “clientes”).

- b) Que sea factible o viable, pudiéndose realizar en la práctica dentro de las posibilidades y la coyuntura que ofrezcan las circunstancias del momento.
- c) Realismo, ajustándose a los recursos disponibles.
- d) Eficiencia. La viabilidad de una evaluación también se mide en función de la relación existente entre los resultados pretendidos y los recursos o medios requeridos.
- e) Prudencia y moderación, sin que resulte “peor el remedio que la enfermedad que se pretende curar”. Antes de aplicar una evaluación conviene calcular y prever si sus resultados van a traer más ventajas (mejora del programa, de las relaciones, del clima de trabajo, etc.) que inconvenientes (reacciones negativas, conflictos, frustraciones, ansiedad, desánimo, etc.).
- g) Diplomacia. La evaluación ha de saberse llevar, tanto en su diseño como en su administración, así como en la difusión de los resultados. No se trata de falsear o escamotear cuestiones críticas, pero sí el plantearlas con delicadeza, respeto, y sobre todo con una actitud positiva y constructiva.
- h) Que sea ética, basada en la cooperación y protección de los derechos de las partes implicadas en la evaluación (institución, equipo de trabajo, participantes). Existe la obligación moral de devolver a la comunidad, de forma sistematizada (mediante el informe final) aquello que se ha extraído de ella (información), y sin ningún tipo de manipulación.
- h) Que sea precisa o exacta. Se requiere una buena pericia y preparación de los evaluadores para prever y controlar una gran cantidad de factores que intervienen. En muchos casos, la compleja realidad con la que se trabaja hace imposible el control y aislamiento de numerosas e impredecibles variables que influyen y contaminan los resultados; sin embargo, sí que deberán utilizar procedimientos con los que minimizar o prever en lo posible los factores distorsionantes.
- i) Clara y directa. Cuando se trabaja en contextos aplicados con asuntos tan sensibles, es necesario que la comunicación entre el evaluador, el cliente y los diversos involucrados en el proceso sea clara y abierta.

Existen una serie de vicios o errores usuales con los que nos encontramos frecuentemente, y que conviene desterrar de la práctica evaluativa.

- a) Pseudoevaluaciones, cuando no se lleva a cabo una auténtica evaluación por no haberse planteado previamente un objetivo ni una estrategia relativa a la recogida de información (es el que caso “vamos a hacer esto...y a ver que sale”). También se trata de pseudoevaluación cuando se realiza el proceso inverso, decidiendo de antemano “qué es lo que quiero que salga”, y dedicando después todos los esfuerzos a justificar con la evaluación lo que va a coincidir con lo que ya había prefijado (cuando al final de aplicar un programa se prepara una memoria de lo efectuado, y se presenta como evaluación) (Stufflebeam y Webster, 1980).
- b) Identificación de la medida con la evaluación. Se incurre en este error cuando se reduce la evaluación a una mera cuantificación de datos, incluso si no es reducible a cifras la realidad que se pretende evaluar. No se puede absolutizar los datos, porque en sí mismos, son insuficientes sin una referencia al contexto que les da sentido y relieve. Así, el número de jóvenes que ha participado en una determinada actividad en diversos centros escolares de una ciudad no es evaluable sin una alusión a la población de jóvenes existentes en todos los centros escolares del área, o sin saber a qué hora y en qué fecha se realizó, o si es la primera vez o se trata de un programa habitual. Se requiere completar los datos cuantitativos con referencias cualitativas y contextuales para poder valorarlos adecuadamente.
- c) Sectarismo o parcialidad. Suele aparecer debido a los prejuicios, intereses o apasionamiento que frecuentemente acompañan a las evaluaciones. Puede haber tendenciosidad tanto en la selección y obtención de informaciones como en el momento de difundirla.
- d) Subjetivismo. Afecta a la evaluación cuando los datos no están convenientemente fundamentados o contextualizados.
- e) Confusión de niveles. Es necesario iniciar la evaluación sabiendo qué, para qué y a quién interesa esa evaluación. En función de la respuesta que se dé a cada una de estas preguntas se establecen distintos niveles e intereses. Así, a la institución que patrocina una determinada actividad comunitaria le puede interesar la relación coste/beneficio por participante, mientras que si al usuario le preocupa el grado de cumplimiento de sus expectativas, y el profesional puede estar interesado en la

calidad del programa. En cada evaluación existen distintos y hasta contradictorios intereses, y el evaluador tiene que dar a cada uno lo suyo, sin confundir la perspectiva institucional con la comunitaria o la profesional.

- f) Burocratización. Es muy frecuente cuando la evaluación se realiza en instituciones muy normadas, en las que la rutina formal y los trámites administrativos pueden actuar a modo de barreras que retrasan y obstaculizan el proceso de evaluación.

La evaluación de programas, dada su cualidad de investigación aplicada, está consagrada al principio de utilidad. Si no tiene influencia en las decisiones habrá sido una tarea inútil, un esfuerzo realizado en vano. Así, de acuerdo con Weiss (1975), la evaluación no merecerá la pena realizarse bajo circunstancias como las siguientes:

- a) Cuando no haya nada que preguntar acerca del programa;
- b) Cuando el programa no está claramente especificado (se improvisa);
- c) Cuando los responsables del programa no se ponen de acuerdo sobre qué se está tratando de lograr con el programa (metas indefinidas);
- d) Cuando no hay disponibilidad de recursos (personal o materiales).

La evaluación es una estrategia más con la que debemos contar para potenciar los efectos de la prevención. La evaluación es una responsabilidad que debemos compartir todos los que trabajamos en la prevención del VIH. La evaluación se encuentra con multitud de obstáculos como son las actitudes reticentes de muchas de las personas involucradas, la falta de capacidad técnicas y recursos que sin duda alguna dificulta la acción en esta esfera fundamental, la aplicación de las estrategias evaluativas en el campo aplicado. No obstante, muchos profesionales han demostrado que estos obstáculos se pueden superar en la medida que cada vez contamos con más posibilidades que hacen factible la evaluación de programas en contextos sociales aplicados. Desde mi punto de vista, a priori el mayor obstáculo con el que nos encontramos actualmente no se traduce en un esfuerzo de trabajo adicional sino en un cambio de actitud y de perspectiva.

Recordemos para finalizar las palabras de León y Montero (2003): “*según las circunstancias, hagamos siempre lo que podamos pero, en cualquier caso, sepamos siempre lo que se hace*”. La lucha contra las ITS y el VIH/SIDA está en manos de todos, y la única lucha que se pierde es la que se abandona.

Referencias Bibliográficas

- Aguilar, M^a J. y Ander-Egg, E. (1992). *Evaluación de servicios y programas sociales*. Madrid: Siglo XXI, pp 90-124.
- Alvira, F. (1991) Metodología de la Evaluación de Programas, Madrid: Centro de Investigaciones Sociológicas. *Colección Cuadernos Metodológicos*, n° 2.
- Anguera, M.T., Chacon, S. y Blanco, A. (coords.) (2008). *Evaluación de Programas Sociales y Sanitarios*.
- Apodaca, P.M. (1999). Evaluación de los Resultados y del Impacto. *Revista de Investigación Evaluativa*, 17, 2, 363-377.
- Chelimsky, E. (1997). The political Enviroment of Evaluation and What It Means for the Development of the Field,. En E., Chelimsky y W.R. Shadish (eds.), *Evaluation for the 21st Century* (pp. 53-58). A Handbook, Sage.
- Cronbach, L.J. (1981). *Toward a reform of program evaluation*. San Francisco: Jossey-Bass.
- De Miguel Díaz, M. (1999). La Evaluación de Programas: entre el conocimiento y el compromiso. *Revista de Investigación Evaluativa*, 17, 2, 345-348.
- Fernández Ballesteros, R. (1992). *Introducción a la evaluación psicológica*, vol. 2. Madrid: Pirámide.
- Fernández del Valle, J. (1996). Roles y estrategias en evaluación de programas. *Intervención psicosocial: Revista sobre igualdad y calidad de vida*, 5 (14), pp. 9-22.
- Fernández-Ballesteros, R. (1995). *Evaluación de Programas. Una guía práctica en ámbitos sociales, educativos y de salud*. Madrid: Síntesis.
- Fernández-Ballesteros, R. y Hernández, J.M. (1995). Listado de Cuestiones Relevantes en Evaluación de Programas (L.C.R.E.P.). En R. Fernández-Ballesteros (ed.). *Evaluación de programas: Una guía práctica*. Madrid: Síntesis.
- Fetterman, D.M. (1994). Empowerment evaluation. *Evaluation Practice*, 15, 1-15.
- Gambara, H. y Vargas, E. (2007). Evaluación de programas de intervención psicosocial. En A. Blanco Y Fodríquez, J. (coord.), *Intervención Psicosocial* (pp.405-456). Madrid: Pearson. Prentice Hall.
- Gómez Serra, M. (2002). La evaluación de servicios y programas sociales: ¿un elemento de control o de transformación de la acción social?. *Revista de Servicios Sociales y Política Social*. Madrid, Consejo General de Colegios Oficiales de Diplomados en Trabajo Social, 57, 81-93.
- Gómez Serra, M. (2004). La evaluación de los servicios y programas sociales: una cuestión urgente. Barcelona: Gedisa.
- Guba, E.G. y Lincoln, Y.S. (1989). *Fourth Generation Evaluation*. California: Sage Publication.

- Instituto Navarro de Administración Pública. Gobierno de Navarra (2008). *Protocolo de Evaluabilidad de las políticas públicas en la Comunidad Foral de Navarra*. Servicio de calidad de políticas y servicios públicos del Instituto Navarro de Administración Pública (INAP).
- Joint Committee on Standards for Educational Evaluation (1981). *Standards for evaluations of educational programs, projects, and materials*. New York: McGraw-Hill.
- Joint Committee on Standards for Educational Evaluation (1988). *The personnel evaluation standards*. Thousand Oaks, CA: Corwin Press.
- Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards*. Thousand Oaks, CA: Sage Publications.
- León, O. G. y Montero, I. (2003). *Métodos de investigación en Psicología y Educación* (3ª ed.). Madrid: McGraw-Hill.
- López de la Llave, A. y Pérez-Llantada, Mª C. (2004). *Evaluación de Programas en Psicología Aplicada (Salud, Intervenciones Sociales, Deporte, Calidad...)*. Madrid: Dykinson-Psicología.
- Martín, Cabero y del Rosario (2007). *Tratamiento estadístico de datos con SPSS. Prácticas resueltas y comentadas*. Madrid: Thomson.
- Páez, D., Ubillos, S. y León, M. (1996). Evaluación del Cambio de Actitudes y Conductas ante el SIDA. En N. Basabe, D. Páez, R. Usieto, H. Paicheler, H. y J.C. Deschamps (eds.), *El Desafío Social del SIDA* (pp. 113-133). Madrid: Fundamentos.
- Rebollosa, E. Fernández-Ramírez, B. y Cantón, P. (2008). *Evaluación de Programas de Intervención Social*. Madrid: Síntesis.
- Rossi, P.H., Freeman, H.E. y Lipsey, M.W. (1999). *Evaluation. A systematic Approach (Sith Edition)*. California: Sage Publications.
- Smith, M (1998). Empowerment evaluation: Theoretical and methodological considerations, *Evaluation and Program Planning*, 21, 255-261.
- Stufflebeam, D. y Webster, W. (1980). An analysis of alternative approaches to evaluation. *Educational Evaluation and Policy Analysis*, 2 (3), 5-20.
- Ubillos, S. (1995). *Guía de Educación Sexual para Jóvenes (II): Recursos para Diseñar un Programa*. San Sebastián: Diputación Foral de Guipúzcoa. DL SS 627/95.
- Ubillos, S. (1996). Objetivos Programáticos, Modelos de Actitudes y Conductas de Salud e Intervención Psicosocial: Prevención entre Jóvenes No Escolarizados. En C. San Juan (ed.), *Intervención Psicosocial. Elementos de Programación y Evaluación Socialmente Eficaces* (pp. 206-248). Colombia: Fundamentos.
- Ubillos, S. (2007). Intervención psicosocial frente al VIH. En A. Blanco y J. Rodríguez Marín (coords), *Intervención Psicosocial* (pp. 103-134). Madrid: Pearson Educación.

- Ubillos, S., Insúa, P. y De Andrés, M. (1999). Parte I: Aspectos Teóricos y Prácticos de los Talleres de Sexo Más Seguro y de Consumo de Menos Riesgo. Evaluación de los Talleres. En P. Insúa (dir.), *Programa de Formación para la Prevención de los Programas de Salud Asociados al Consumo de Drogas. Manual de Educación Sanitaria: Recursos para Diseñar Talleres con Usuarios de Drogas* (pp. 89-142). Zarauz: Delegación del Gobierno para el Plan Nacional sobre Drogas, Plan Nacional sobre el SIDA y Universidad del País Vasco.
- Vedung, E. (1997). *Evaluación de Políticas Públicas y Programas*. Madrid: Ministerio de Trabajo y Asuntos Sociales. Instituto Nacional de Servicios Sociales (INSERSO).
- Vega, F. J. y Ventosa, J. (1993). *Programar, acompañar, evaluar*. MADRID. CCS.
- Vilà Baños, Ruth. (2006). *¿Cómo hacer un análisis cuantitativo de datos de tipo descriptivo con el paquete estadístico SPSS?*. Butlletí LaRecerca.
- Weiss, C.H. (1975). *Investigación Evaluativa. Métodos para determinar la eficacia de los programas de acción*. México: Trillas.
- Wholey, J. S., (1987). Evaluability assessment: Developing program theory. En L. Bickman (ed.), *Using program theory in evaluation*. New Directions for Program Evaluation, 33. Jossey-Bass, San Francisco.
- Wholey, J.S. (1979). *Evaluation: Promise and Performance*. Washington, DC.: The Urban Institute.