



**MÁSTER EN EVALUACIÓN DE PROGRAMAS Y POLÍTICAS PÚBLICAS**

***XIX PROMOCIÓN (2020 – 2021)***

## TRABAJO FIN DE MÁSTER

DISEÑO DE EVALUACIÓN DE IMPACTO DEL PROGRAMA  
DE SENSIBILIZACIÓN DE BUENAS PRÁCTICAS HIGIÉNICAS  
Y DE ALIMENTACIÓN EN COLEGIOS EN BURKINA FASO

**Realizado por: Emilio Carrero Muñoz**  
**Tutor: Carlos Mediano**  
**Octubre 2022**

*Tenemos que comprender el cosmos tal como es y no confundir lo que es con lo que queremos que sea. **Lo obvio es a veces falso, lo inesperado es a veces cierto.**" Sobre la importancia de la alfabetización científica. Carl Sagan*

## Índice del documento

|                                                                       |    |
|-----------------------------------------------------------------------|----|
| Agradecimientos .....                                                 | 1  |
| Introducción .....                                                    | 2  |
| Antecedentes .....                                                    | 2  |
| Bloque I. Planteamiento.....                                          | 3  |
| 1. Encargo.....                                                       | 3  |
| 2. Contexto y funcionamiento del programa .....                       | 4  |
| 3. Definición de la unidad de evaluación .....                        | 5  |
| 4. Aproximación evaluativa.....                                       | 6  |
| Bloque II. Diseño metodológico .....                                  | 6  |
| 1. Identificación de teoría del cambio.....                           | 7  |
| 2. Identificación del modelo lógico .....                             | 7  |
| 3. Identificación de las necesidades información .....                | 9  |
| 4. Identificación de indicadores .....                                | 10 |
| 5. Técnicas de recolección de información .....                       | 11 |
| 6. Estrategia de muestreo .....                                       | 13 |
| 7. Cálculo del tamaño de la muestra. Efecto mínimo detectable .....   | 15 |
| 8. Metodología para la medición del impacto. Operacionalización ..... | 18 |
| Bloque III. Ejecución .....                                           | 22 |
| 1. Trabajo de Campo .....                                             | 22 |
| 2. Análisis e interpretación de datos.....                            | 25 |
| 3. Enjuiciamiento.....                                                | 29 |
| 4. Comunicación de resultados .....                                   | 30 |
| 5. Recomendaciones .....                                              | 30 |
| Bloque IV. ....                                                       | 31 |
| 1. Calendario .....                                                   | 31 |
| 2. Presupuesto.....                                                   | 32 |
| 3. Protocolos éticos .....                                            | 33 |
| 4.Utilización de los Resultados.....                                  | 33 |
| 5.Limitaciones de la evaluación .....                                 | 34 |
| Reflexiones personales.....                                           | 35 |
| Referencias bibliográficas .....                                      | 37 |
| ANEXOS .....                                                          | 39 |
| Anexo 1. Matriz de evaluación .....                                   | 39 |

|                                                  |    |
|--------------------------------------------------|----|
| Anexo 2. Instrumentos:.....                      | 39 |
| Anexo 3. Encuesta de línea de base alumnado..... | 42 |
| Anexo 4. Encuesta de pareamiento .....           | 42 |
| Anexo 5. Encuesta monitoreo .....                | 42 |
| Anexo 6. Encuesta de padres .....                | 42 |
| Anexo 7. Presupuesto .....                       | 42 |

## Índice de figuras

|                                                                                     |    |
|-------------------------------------------------------------------------------------|----|
| Figura 1 Modelo Lógico del programa. Elaboración propia .....                       | 8  |
| Figura 2 Distribución de la muestra. ....                                           | 14 |
| Figura 5 Diferencia de medias. Tamaño del efecto y Potencia (fuente: G*Power) ..... | 16 |
| Figura 6 EMD, Potencia y tamaño de la muestra .....                                 | 16 |
| Figura 7 Método de diferencia en diferencias. Fuente BID .....                      | 19 |
| Figura 8 Pareamiento por puntajes de propensión y rango común. Fuede BID.....       | 20 |

## Índice de tablas

|                                                                                       |    |
|---------------------------------------------------------------------------------------|----|
| Tabla 1 Cálculo efecto mínimo detectable. Fuente BID.....                             | 18 |
| Tabla 2 Ejemplo de puntaje de propensión. ....                                        | 21 |
| Tabla 3 Ejemplo de prueba de equilibrio del emparejamiento de la variable Manos ..... | 21 |
| Tabla 4 Variables posibles y asociaciones para análisis .....                         | 27 |
| Tabla 5 Calendario aproximado del trabajo de campo .....                              | 31 |
| Tabla 6 Presupuesto estimado de la evaluación.....                                    | 33 |

## Agradecimientos

En primer lugar, quiero agradecer a mis amigos y amigas del máster por sus ánimos continuos en este largo viaje del máster y del TFM. También a mis compañeros de trabajo que me cambiaron muchas guardias para poder asistir a las clases.

Agradezco a los profesores Lucila Finkel y Oscar Franco su ayuda y disposición a mis preguntas para este TFM y a mi tutor por sus comentarios.

A mi compañera Rosa María Flores, que me acompañó en el camino del TFM y a quién consulté muchos asuntos que me ayudaron de forma muy importante en este trabajo.

A Lucía, por su apoyo y ánimo continuo y su ayuda en la lectura y corrección de este trabajo.

A mi madre y mi padre, que me enseñaron a seguir adelante a pesar de las dificultades.

A todas vosotras, muchas gracias. No hubiese sido posible sin vuestro apoyo y ayuda.

## Introducción

El presente documento detalla la propuesta metodológica para llevar a cabo la evaluación de impacto del programa de «Sensibilización de buenas prácticas higiénicas y de alimentación» que la ONG SocialClown y su contraparte la asociación Reflexe implementarán en Burkina Faso en el curso escolar 2023-2024 con el acuerdo de las autoridades en materia de Salud y Educación.

Para plantear la evaluación que se aborda en este trabajo se ha elegido un proyecto que ya existe como base, pues eso permite plantear un escenario tangible y realista. Los consensos y comunicaciones que se mencionan en el texto son reales, fruto de conversaciones realizadas a través del correo electrónico y el teléfono con la asociación SocialClown. Sin embargo, eso no quiere decir que la evaluación vaya a llevarse al terreno, aunque pueda parecerlo al leer el texto. Solamente se optó por una redacción que situara al lector en un escenario real para facilitar la lectura.

Para llevar a cabo el diseño se ha tomado como base la metodología propuesta en la Rayuela (Rayuela, 2019) y se utilizarán tanto métodos cuantitativos como cualitativos. Debido a que este programa aún no cuenta con datos, y sabiendo que los impactos del programa son resultados a medio y largo plazo, se ha optado por realizar una evaluación de impacto orientada por la teoría del cambio. Esto permitirá analizar la lógica y la congruencia del programa, un requisito indispensable para valorar si los resultados que se observen pueden explicarse a partir de la intervención y así saber el impacto real de ésta.

Cómo se verá más adelante es un enfoque mixto con la intención de no valorar solo el impacto del programa sino también entender cómo funciona para mejorar su implementación. Es por ello por lo que durante la lectura se puede tener la sensación de que hay dos evaluaciones imbricadas que persiguen estos dos objetivos principales; saber cuál es el impacto y cómo funciona el programa para conseguirlo.

## Antecedentes

En el año 2015, la Asociación SocialClown empezó a trabajar en la región de Hauts Bassins junto con la asociación local Reflexe con un objetivo claro: acabar con la malnutrición en la zona. Comenzaron con los casos más urgentes: los niños/as que ya sufrían esta enfermedad. Aplicando la estimulación psicomotriz y sensorial se obtuvieron excelentes resultados en la recuperación de los niños/as malnutridos atendidos en el Centro de Recuperación Nutricional del Hospital Souro Sanou de Bobo Dioulasso (SocialClown, 2015), disminuyendo las tasas de mortalidad y de abandonos del programa nutricional y aumentando el peso medio ganado por los niños/as ingresadas. Sin embargo, aunque existieron buenos resultados, se observó que la malnutrición seguía afectando a toda la región sin existir una causa directa.

Buscando entender cómo era posible que la región de Hauts Bassins, donde no faltan recursos nutricionales, se vea azotada por la malnutrición, dichas organizaciones realizaron un trabajo etnográfico holístico cuyos resultados analizaron también con médicos locales, llegando a la conclusión de que **la principal causa de malnutrición la podría estar generando una falta de conocimientos sobre buenas prácticas alimentarias y de higiene.**

Para revertir esta situación, en 2016 se puso en marcha un proyecto de cuatro años para dotar a la población de la zona de los conocimientos prácticos oportunos y la actitud necesaria para

reducir y prevenir la incidencia de la malnutrición «Estimulación y sensibilización de la población malnutrida región de Hauts-Bassins», en Burkina Faso (SocialClown, 2016).

Antes de su implementación, se realizó una investigación para saber qué actividades eran las más adecuadas para conseguir el fin buscado y con qué frecuencia debían realizarse. Para ello, se realizó una encuesta CAP proporcionada por la FAO y adaptada al contexto por el Dr. Alain Hien, responsable de nutrición en la región. Los resultados obtenidos apuntaron a que la población incrementaba sus conocimientos más y mejor a través de talleres teatralizados.

Posteriormente se realizó un programa piloto en 113 escuelas públicas situadas en los barrios periféricos de la ciudad de Bobo Dioulasso cuyos estudiantes provienen, en su gran mayoría, de barrios «non loti» (no confortables), ya que carecen de acceso a agua corriente y electricidad. Cabe señalar que la población que habita en estos barrios, asentados recientemente, proviene del mundo rural y, en general, no ha disfrutado de un servicio de educación para la salud de calidad y aún conserva ciertas prácticas culturales nefastas, tales como la mutilación genital femenina, ciertas prohibiciones alimentarias y varios tabús.

A pesar de que no se tenían datos de impacto del programa, ya que no se pudo realizar ninguna evaluación por falta de fondos, la buena acogida que tuvo la campaña de sensibilización por parte de los colegios, padres y alumnos, llevó a las organizaciones Reflexe y SocialClown a implantar el programa al extrarradio de Bobo Dioulasso, con la intención de extender la campaña de sensibilización a toda la provincia en los años siguientes si se lograban observar buenos resultados en la disminución de la malnutrición.

Una vez realizado un primer trabajo documental y un acercamiento a los responsables de la campaña, se observó que ésta no contaba con datos suficientes que proporcionaran evidencia sobre la lógica de intervención y de resultados que avalaran su extensión a toda la provincia. Con el propósito **recoger los datos necesarios**, se propone el diseño de una evaluación de impacto orientada bajo la teoría del cambio.

## Bloque I. Planteamiento

### 1. Encargo

La intención del presente diseño de evaluación de impacto responde a las necesidades de información de las organizaciones SocialClown y Reflexe. El objetivo es conocer qué impacto ha tenido este programa en la mejora de la nutrición de la población infantil y ayudar a la consolidación del programa para que se pueda incluir en el currículo educativo. La evidencia de buenos resultados les permitirá solicitar más fondos y, de esta forma, extenderlo al resto de los colegios de la provincia, así como que se repita todos los años en los cursos de ciclo medio 1 (CM1) y que las actividades de sensibilización figuren como obligatorias en el currículo de los colegios, de forma que estos conocimientos sigan mejorando las condiciones de salud de la población.

De acuerdo con el manual del Banco Interamericano de Desarrollo (en adelante, BID), el propósito de una evaluación del impacto «es aprender y cambiar aquello que hace que la vida sea mejor para ellos (la población que recibe el programa). Para lograrlo necesitamos métodos mixtos y pluralismo. Para la valoración del impacto pueden y deberían emplearse muchas aproximaciones y herramientas» (Paul J. Gertler, 2017). Considerando lo anterior, los principales propósitos que de la evaluación planteada son:

- Conocer el impacto del programa de sensibilización sobre buenas prácticas higiénicas y de alimentación.

- Comprender y conocer el funcionamiento del programa para mejorar la implantación y la toma de decisiones por parte de los actores correspondientes.

## 2. Contexto y funcionamiento del programa

El programa de sensibilización quiere proporcionar información y conocimientos básicos sobre higiene y alimentación que permita a la población la incorporación de hábitos más saludables en su vida y en la aquellos de su entorno. Para ello se utiliza el lenguaje emocional mediante una serie de actividades creadas con el método del Teatro-Fórum. Éste, se convierte en un instrumento para comprender la situación y buscar alternativas al problema social de la malnutrición. ***El objetivo es estimular al alumnado que son no-actores, a ser protagonista de la acción dramática y expresar sus vivencias cotidianas, ayudándolo a reflexionar sobre su situación para poder cambiar la realidad y mejorar el futuro.***

El público ***objetivo de la intervención son los escolares de 9 y 10 años de enseñanza primaria de ciclo medio 1 (CM1).*** ¿Por qué se han elegido estos/as niños/as? Porque es en esta edad, también llamada «edad de la razón» según Jean Piaget, cuando se empieza a desarrollar el pensamiento crítico y se empieza a ser capaz de entrar en mayor contacto con la realidad. Además, se encuentran a la entrada en la pubertad, que es cuando se afianzan las prácticas con respecto a la salud.

Los/as niños/as son los mayores vectores de propagación de enfermedades contagiosas y, en muchas ocasiones, se encargan de ayudar en las tareas domésticas (acarrear agua, preparar comidas, cuidar de hermanos menores, limpieza...). Según consultas a maestros y médicos locales, y por la experiencia anterior en el proyecto de sensibilización y el programa del hospital, los/as menores de 11 años son los/as que tienen más probabilidad de practicar lo aprendido y los mayores vectores de propagación de la información, por lo que actuando sobre este sector de la población se busca que el mensaje llegue con mayor efectividad a sus hogares, convirtiéndolos en difusores de información dentro de la familia y dentro de la comunidad, en verdaderos agentes de cambio social.

La provincia de Houet es grande y la población está muy dispersa. El programa se implementaría en un futuro en todos los colegios públicos de la provincia. Por problemas de presupuesto, de momento para el curso 2023-24 sólo se realizará en 3 pueblos. Entre los tres abarcan 95 escuelas públicas y se llegará a unos 10500 alumnos y alumnas de las clases CM1. Como sucedía en el programa anterior el contexto socioeconómico es muy similar y precario. Carecen de acceso a agua corriente y electricidad, no han tenido una educación para la salud y por estar ahora en un entorno más rural las prácticas alimentarias y de higiene son tan malas o peores que en la ciudad.

Las escuelas se presentan como el marco idóneo donde los escolares pueden adquirir buenos hábitos en alimentación e higiene, gracias a las actividades planteadas en la campaña de sensibilización planteadas en este programa. Gracias a la asimilación de estos hábitos y su aplicación en los propios centros educativos y en sus hogares, junto a toda su familia, se podrá combatir la malnutrición. Además, este proyecto dotará a los docentes con una guía que les permita continuar con esta labor.

Este programa es una apuesta de futuro para combatir la malnutrición a través del aprendizaje de buenos hábitos de higiene y de alimentación. Con la presente intervención, se pretende

ampliar el rango de acción realizando una campaña de sensibilización a través de actividades participativas para los niños/as de, dejando, tras la intervención, una guía para que los docentes puedan continuar con esta labor. Las escuelas se presentan como el marco idóneo donde los escolares pueden adquirir buenos hábitos en alimentación e higiene para aplicarlos en la escuela y luego, en sus hogares junto a toda su familia. Los menores son uno de los colectivos más vulnerables de Burkina, así como los mayores vectores de propagación de información y la esperanza de un futuro mejor.

### 3. Definición de la unidad de evaluación

El objeto de esta evaluación es el programa de buenas prácticas de higiene y de alimentación que se realizará en tres pueblos de la región de Hauts Bassins durante el curso escolar 2023-2024. Esto supone que se aplicará en 95 colegios y afectará a 10.500 niños/as de entre 9 y 10 años de las clases de Ciclo Medio 1.

La evaluación estará realizada por un equipo externo al programa y a los organizadores formado por un equipo de personas de España y otro de Burkina Faso. Estos dos equipos tendrán comunicación permanente y una total coordinación. Además, estos dos equipos estarán coordinados con el personal que organiza el programa. El proceso de evaluación tiene varios momentos de toma de datos durante todo el año escolar.

Los actores clave de este programa son: la asociación SocialClown, principal organizador del programa y quien encarga la evaluación; la asociación local Reflexe, que es la implementadora de las actividades; los representantes regionales de Educación y de Salud; los/las alumnos/as; los/as profesores/as y los padres y madres que estarán en contacto con el programa.

|                                                                                                                                                                                                                         |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Nombre del programa: Sensibilización de buenas prácticas de higiene y de alimentación                                                                                                                                   |
| Año de creación: 2023-2024                                                                                                                                                                                              |
| Años que lleva implementándose: desde 2016 un proyecto similar.                                                                                                                                                         |
| Lugar en los que se implementa: Bobo Dioulasso y región.                                                                                                                                                                |
| Tipo de beneficios que otorga: conocimientos sobre hábitos de higiene y alimentación                                                                                                                                    |
| Población objetivo: escolares de enseñanza primaria de ciclo medio 1 (CM1) de 9 y 10 años.                                                                                                                              |
| <ul style="list-style-type: none"> <li>- Alumnado atendido: 10500</li> <li>- Alumnado participa en la evaluación: 2300</li> </ul>                                                                                       |
| Objetivo General:                                                                                                                                                                                                       |
| Reducción de la malnutrición en la provincia de Houet                                                                                                                                                                   |
| Probable teoría de cambio:                                                                                                                                                                                              |
| Mediante las actividades teatralizadas los alumnos/as adquieren conocimientos sobre alimentación e higiene que aplican en su día a día y que les reporta beneficios saludables y poco a poco se reduce la malnutrición. |
| Funcionamiento:                                                                                                                                                                                                         |

A través de las actividades sobre todo del teatro fórum, los niños y niñas adquieren los nuevos conocimientos que practicarán en adelante y esto repercutirá en su salud y las de sus familiares si les pasan estos nuevos hábitos y prácticas en higiene y alimentación. Estos nuevos conocimientos reducirán la malnutrición en la zona en el futuro.

#### 4. Aproximación evaluativa

Como se explicaba anteriormente, para poder contestar a las necesidades informativas de quien encarga la evaluación se ha optado por plantear una evaluación de impacto orientada por la teoría del cambio. De esta forma, será posible conocer el impacto y los procesos que dan lugar a estos resultados, así como la relación causal que hay entre unos y otros. Se tendrán en cuenta todos los elementos que forman el programa (estructura, procesos y resultados), pero de acuerdo con las necesidades de información de los organizadores que son quienes encargan esta evaluación, ésta estará más enfocada en los procesos y los resultados del programa.

### Bloque II. Diseño metodológico

Para llevar a cabo esta evaluación se propone utilizar métodos mixtos, que permiten una mayor comprensión del programa, ya que se obtiene información de distintas fuentes y por procedimientos diferentes. Los datos serán obtenidos a través de entrevistas semiestructuradas representativas, entrevistas grupales, encuestas y la observación directa.

Como dice (Bamberger, 2012) «las evaluaciones con métodos mixtos procuran integrar las disciplinas de las ciencias sociales con enfoques predominantemente cuantitativos y cualitativos a la teoría, la recolección, el análisis y la interpretación de los datos. El objetivo de los métodos mixtos es aumentar la confiabilidad de los datos y la validez de los hallazgos y las recomendaciones, además de ampliar y profundizar la comprensión de los procesos a través de los cuales se logran los resultados e impactos de un programa, y cómo estos se ven afectados por el contexto en el que se implementa el programa». Es por ello por lo que esta evaluación tiene una parte de estudio cuantitativo de datos y otra de estudio cualitativo que se complementan.

#### Análisis cuantitativo

Este tipo de análisis de la información está basado en datos numéricos provenientes de encuestas, generalmente. Para ello, se utilizan técnicas de análisis estadístico que explican el comportamiento de las variables anteriormente codificadas y que forman parte del programa.

Se realizarán encuestas a gran parte de la población, afectada y no por el programa, para obtener la información necesaria para poder evaluar el impacto del programa tal y como se recoge en los manuales de evaluación de impacto. La idea es estimar causalmente el efecto sobre el alumnado beneficiado por programa basándose en datos.

Para ello, se necesitará tener un **grupo de tratamiento**, formado por los/las alumnos/as beneficiados, y un **grupo de comparación o control**, compuesto por alumnos/as que no han recibido esa formación y que conforman el **contrafactual**. Este grupo representa lo que hubiese ocurrido al de tratamiento si no hubiese participado en el programa. Y la diferencia entre ambos al final de todo el proceso sería el impacto, en caso de haberlo.

Además, se pretende realizar encuestas a parte de los padres y las madres para conocer cómo repercute indirectamente el programa en los familiares, pues el programa quiere usar a los/as alumnos/as como diáspora de buenos hábitos. Las encuestas realizadas a los padres y las madres son para confirmar datos y converger con los obtenidos en el análisis cualitativo de sus entrevistas. No tendrán significación estadística, por lo que serán menos numerosas. Se puede consultar en el anexo 6.

## **Análisis cualitativo**

En este tipo de análisis de información de carácter no numérico se extraen conclusiones, ideas, significados que aportan un conocimiento más profundo sobre las cuestiones del programa y su funcionamiento.

Este apartado es de suma importancia, pues ayudará a enfocar mejor los resultados de la evaluación. Se elegirán de manera aleatoria hogares de la muestra del grupo de tratamiento, en los que se harán entrevistas. También se harán en servicios de salud, a profesores/as y a implementadores. Este análisis aportará información importante sobre cómo funciona el programa durante su implementación y también después.

De los hogares seleccionables de los/as niños/as que participan en el programa, se realizarán entrevistas estructuradas en aquellos que aportan información en las encuestas. El objetivo es concentrar el foco en estos/as alumnos/as para intentar controlar todas las variables posibles y evitar sesgos, además de recabar toda la información posible.

Puesto que es un proceso bastante complejo, antes de llevar a cabo la recolección de datos es necesario planificarla. Para ello, se plantean a continuación una serie de etapas o pasos importantes.

### **1. Identificación de teoría del cambio**

Los programas, por lo general, están compuestos por muchos elementos que trabajan relacionados causalmente entre sí, formando un conjunto o sistema que persigue obtener unos resultados a través de unas actividades y unos productos mediante diferentes procesos. El programa tiene una «idea base» sobre la cual funciona, una teoría que pretende cambiar la realidad. Es lo que se llama **la teoría del programa**. Parte de esta evaluación está basada en la teoría del programa y en cómo éste se ha implementado. Por lo tanto, la presente evaluación se centrará en la construcción de la teoría del cambio, que se considera fundamental para revisar el sustento teórico del programa, ya que una mala construcción de ésta no permite observar los resultados de la intervención.

La relación causal que subyace entre todos los elementos del programa es parte del objeto de esta evaluación. Además, se analizará la calidad de los resultados obtenidos en base al análisis de los procesos.

### **2. Identificación del modelo lógico**

Con la intención de ver de manera rápida y clara los elementos principales que conforman el programa y su cadena causal se ha elaborado un esquema gráfico con una lógica causal donde aparecen los principales elementos del programa y sus relaciones. A esto se le denomina modelo lógico del programa. Se expone a continuación de manera escrita y gráfica:

- Insumos de los que dispone el proyecto.
- Procesos generados por y para las actividades del programa.
- Actividades y acciones que se van a llevar a cabo.

- Resultados que son previsibles a corto y a medio plazo y
- Resultados finales que son el impacto.

Para la elaboración del modelo lógico es importante que el equipo que va a realizar la evaluación revise toda la información que haya al respecto. En este caso, el proyecto piloto anterior en las escuelas de la ciudad de Bobo-Dioulasso, así como el programa de estimulación sensorial y psicomotriz de los niños/as/as ingresados por desnutrición en el hospital de Sanou-Sorou de la ciudad de Bobo-Dioulasso y su evaluación son información y documentos de mucho interés para contextualizar y acercarse al objeto de evaluación. Con esta información y mediante correo electrónico se ha consensuado con la organización SocialClown el siguiente modelo lógico del programa de sensibilización de buenas prácticas en higiene y alimentación:

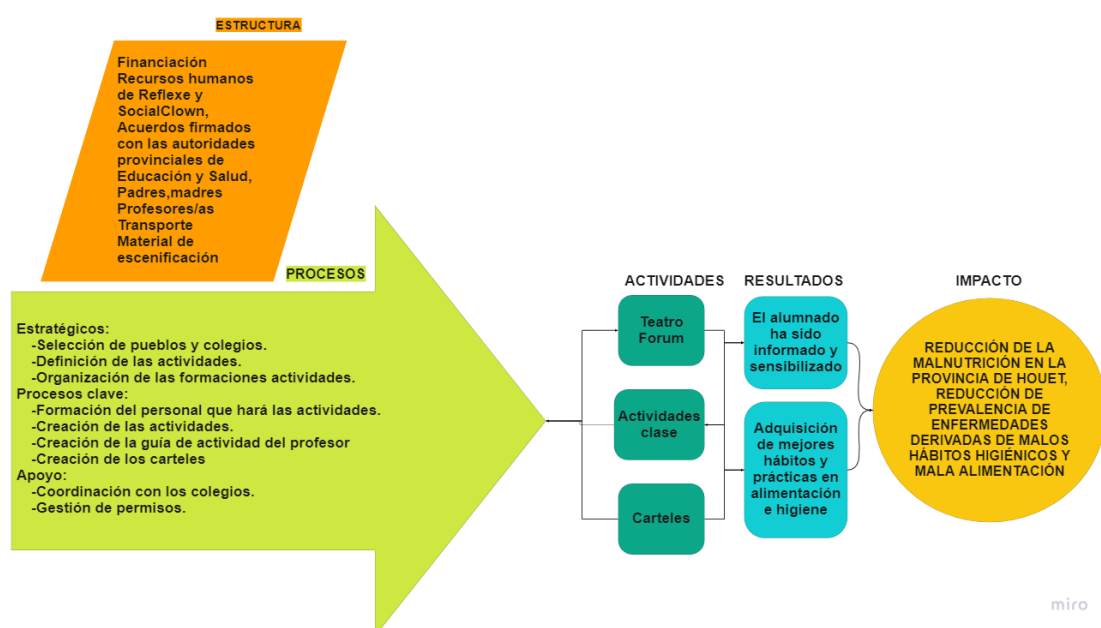


Figura -1 Modelo Lógico del programa. Elaboración propia

Los tres primeros componentes (estructura, procesos y actividades) son la parte de la implementación del programa y dependen principalmente de quienes operan el programa, los resultados están dados por la interacción entre implementadores y beneficiarios.

La elaboración de un modelo lógico que represente las principales dimensiones del programa y su vinculación con las preguntas de evaluación va a permitir no solo valorar cada dimensión de manera independiente, sino la interrelación entre ellas y la lógica que subyace.

Por lo tanto, con base al modelo lógico la formulación de **la hipótesis causal del programa sería:**

*“Si los niños/as reciben las sesiones de sensibilización de higiene (Teatro Fórum), siguen las actividades del curso en relación con la higiene y alimentación y estos **adquieren** los conocimientos básicos en buenos hábitos y prácticas en alimentación e higiene esto **debería conducir** a tener en la comunidad niños/as con buenos hábitos y conocimientos **que al final mejorará** la nutrición de estos, **contribuyendo** a la reducción de la prevalencia de la malnutrición de los niños/as en la región de Houet”*

La forma de saber si el modelo lógico funciona cómo se había supuesto y que esta hipótesis causal es así, es mediante “la medición” con los indicadores específicos seleccionados para las preguntas de evaluación.

Este modelo lógico y la hipótesis causal han de ser consensuados y validados con los organizadores e implementadores. Las preguntas de evaluación, los indicadores y las necesidades de información de parte de los actores implicados son o están referidos a este modelo lógico.

### 3. Identificación de las necesidades información

Se realizarán reuniones con diferentes actores implicados en el programa, profesores, implementadores y organizadores y el equipo de evaluación implicado con la intención de saber cuáles son para cada uno de ellos estas necesidades y así enfocar mejor la evaluación. Se partirá del modelo lógico del programa como eje para estructurar la información. En este caso, al tratarse de una evaluación de impacto orientada por la teoría del cambio hay distintos tipos de necesidades y de información. Unas más bien están enfocadas a medir el impacto mientras que otras están más enfocadas en el desarrollo e implementación del programa.

Desde el punto de vista de una evaluación de impacto la pregunta fundamental es según BID ¿cuál es el impacto o el efecto casual de un programa en un resultado de interés?, es decir, ¿cuáles son los cambios directamente atribuibles a un programa? (Paul J. Gertler, 2017, pág. 9) Esta pregunta fundamental tiene que estar formulada en forma de hipótesis que esté claramente definida, que sea comparable y que pueda de alguna manera cuantificar la diferencia entre los resultados obtenidos al contrastar los grupos de tratamiento y de comparación o control.

En este caso, en base a la teoría del programa, la pregunta clave de evaluación es **¿cuál es el efecto neto del programa en la prevalencia de la malnutrición?** Sería conveniente y deseable saber también si se llega al resultado general del programa **¿cuál es el efecto del programa de sensibilización en el cambio de comportamiento en los hábitos de higiene y alimentación de los niños/as y sus familias?**

**Hipótesis nula,  $H_0$ :** *el programa de sensibilización no reduce la malnutrición de los niños/as de la zona ni provee de hábitos de higiene y alimentación buenos para la salud.*

**Hipótesis alternativa,  $H_1$ :** *el programa mejora la situación de malnutrición y provee de buenos hábitos de higiene y alimentación para la salud.*

Podría suceder que los alumnos/as adquieran buenos hábitos de higiene y alimentación y, aun así, no se reduzca la malnutrición. Resulta difícil pensar que a sabiendas de que una buena higiene y alimentación es bueno para la salud y teniendo medios para ponerlo en práctica, esto no dé resultado cuando hay multitud de programas y proyectos de higiene y alimentación que tienen demostrado su impacto positivo. De suceder esto, cabría pensarse que el programa no ha identificado bien la causa de la malnutrición o que existen barreras que se desconocen que impiden poner en práctica lo aprendido y que no se han tenido en cuenta en el programa o que el programa no consigue transmitir lo planeado o quizás alguna otra explicación.

Como esta evaluación va más allá del impacto, se pretende contestar también otras preguntas:

- ¿Cuál es el efecto del programa en la adquisición de mejores hábitos de higiene y alimentación por parte de los/as alumnos/as?
- ¿Cuál es el efecto del programa en la adquisición de mejores hábitos de higiene y alimentación por parte de la familia del alumno/a que participa?
- ¿Cuál es el efecto del programa en la salud de los/as alumnos/as y de sus familiares?
- ¿Ha habido una reducción de las enfermedades derivadas de la mala higiene?
- ¿A través de qué mecanismos el programa ayuda a mejorar la calidad de vida de los beneficiarios/as y a reducir la pobreza?
- ¿Se mantienen los efectos del programa en el tiempo?
- ¿Qué factores explican los resultados obtenidos?
- ¿Todos los componentes del programa son necesarios para lograr el impacto?
- ¿Existen efectos adversos no planificados?
- ¿El programa alcanzó los objetivos para los cuales fue diseñado?
- ¿Qué factores externos han influido en el programa?

Todas estas preguntas junto con otras serán consensuadas con los actores implicados y quedarán recogidas en la matriz de evaluación del anexo 1.

#### 4. Identificación de indicadores

La manera más adecuada de poder acercarse a medir lo que está sucediendo en el programa es la utilización de unos indicadores que informan del grado de su cumplimiento en el aspecto que estudian, aunque en muchas ocasiones esto tenga que hacerse de manera indirecta debido a la dificultad de medir el concepto en cuestión. Por supuesto, estos indicadores estarán de acuerdo con el modelo lógico y las preguntas de evaluación.

Para la mejor elección de los indicadores se consultará la documentación disponible y se consensuarán con los implementadores, organizadores y autoridades. Estos indicadores deberían ser medibles, ya sea en porcentaje, escala de juicio, etc., atribuibles y realistas, aunque no siempre es posible conseguirlo. En ese caso, se plantearán otras estrategias para completar la información necesaria. Si fuera necesario se cambiarían preguntas o técnicas o actores y se buscaría la forma de medirlo.

Se tendrá al menos un indicador por pregunta de evaluación, aunque a veces sea necesario utilizar varios para acercarse al concepto que se pretende medir.

Todo esto finalmente quedará recogido de manera resumida en una matriz de evaluación (ver anexo 1) donde estarán reflejadas las dimensiones del modelo lógico, las preguntas de evaluación, los indicadores utilizados y la/s técnica/s de recolección empleada para cada pregunta.

Para valorar los cambios en la malnutrición, se recogerá información antropométrica de los niños y niñas que participan en el programa para valorar los cambios en el estado de los alumnos/as. A falta de mejores indicadores, pues no se han podido consensuar con la asociación SocialClown, se utilizarán los siguientes que se basan en relaciones entre la talla, el peso, la edad y la circunferencia del brazo izquierdo del niño/a:

- El indicador Peso para la edad (P/E), mide la relación entre el peso de un individuo y su edad.
- Peso para la talla (P/T), mide la relación entre el peso de un individuo y su talla.
- Talla para la edad (T/E) mide la relación entre la talla de un individuo para su edad y sexo.

- Circunferencia del brazo para la edad (CB/E): expresa la relación entre la circunferencia del brazo izquierdo para su edad y sexo. Es un indicador compuesto de reserva calórica y proteica.

Una vez seleccionados estos indicadores, se procederá a la elección de las técnicas más adecuadas para la recogida de información que permitirá contestar a las preguntas de evaluación en cada dimensión del programa.

## 5. Técnicas de recolección de información

Para la recolección de información que luego será tratada se proponen diferentes instrumentos y técnicas:

- Revisión documental del programa de sensibilización.
- Entrevistas individuales con actores clave.
- Entrevistas grupales con padres, madres y profesores/as.
- Encuestas a padres y al alumnado.
- Observación directa.

Se describen a continuación los aspectos más importantes de estas técnicas. Las encuestas y los guiones de las entrevistas se pueden consultar en los anexos 2,3 y 4.

### Revisión documental

Se trata de recolectar toda la información disponible del programa, en cualquier formato, para hacer una lectura intensa y minuciosa y poder realizar un análisis de la información con la finalidad de:

1. Comprender e identificar el programa y los procesos que deben realizarse para lograr la cadena de resultados establecida en él.
2. Obtener información a partir de los documentos del programa, para revisar el grado de cumplimiento de los procesos y las acciones.
3. Búsqueda de información para confirmar la teoría del programa.

En este caso, se usará la de un proyecto piloto similar realizado en 2020 en las escuelas públicas de la ciudad de Bobo-Dioulasso del cual no se tiene ninguna evaluación pero sí toda la argumentación y fundamentación del proyecto, hay un programa de estimulación sensorial y psicomotriz de los niños/as ingresados por desnutrición en el hospital de Sanou-Sorou en la ciudad de Bobo-Dioulasso, con su respectiva evaluación, así como toda la documentación presentada por la organización SocialClown para la obtención de la financiación del programa ante las autoridades de la Junta de Castilla y León.

### Entrevistas individuales a actores clave.

Estas entrevistas tendrán lugar con los actores clave o stakeholders de este programa. Se seleccionará a responsables de SocialClown y Reflexe, a padres, madres y profesores/as de niños/as que están involucrados en todo el proceso de evaluación de los tres pueblos afectados y colegios seleccionados del grupo de tratamiento.

Las preguntas de evaluación de las entrevistas individuales saldrán de la matriz de evaluación y del modelo lógico. Estas entrevistas seguirán un guión establecido con preguntas que antes estarán consensuadas con los organizadores. Habrá suficiente flexibilidad durante la entrevista

para poder preguntar por asuntos no recogidos en el guión, pero que pueden ser interesantes para el proceso evaluativo. El guión se puede consultar en el anexo 4.

Se pedirá el consentimiento informado y firmado a todas las personas que participen en el proceso para usar sus opiniones de forma anónima, asegurándoles que no aparecerán sus nombres en ningún documento. Se grabarán las entrevistas para transcribirlas posteriormente y evitar interpretaciones y pérdida de información en la toma de notas y así después realizar el análisis del discurso con lo que las personas expresaron.

### **Entrevistas de grupo**

Este tipo de sesiones se realizarán para obtener información de los procesos más integrales, permitiendo la reflexión libre de los participantes, así como para obtener propuestas de mejora y opiniones más informadas. Se seguirá un guión que se puede consultar en el anexo 2.

Se realizarán reuniones de padres, madres y profesores/as de distintos colegios de los pueblos del programa, pero con la peculiaridad de que ellos entre sí no se conocerán y no tendrán ningún tipo de relación. Se busca así que tengan total libertad para expresar sus opiniones.

También en este caso se pedirá que firmen el consentimiento informado para usar sus opiniones de forma anónima y se grabarán las entrevistas en su totalidad.

### **Encuestas colectivas al alumnado**

Éstas encuestas están enfocadas a la recolección masiva de datos de los dos grupos identificados para posteriormente realizar el análisis. Serán encuestas de línea de base (inicio y final) y de monitoreo para conocer cómo avanza el programa.

El equipo de campo será el encargado de organizar esta recogida de información, en coordinación con los representantes de los colegios, la asociación Reflexe y los/as profesores/as/as. Ellos serán los encargados de explicar el motivo de las encuestas y su importancia, cómo se contestan de manera adecuada y el sistema de identificación.

Aunque las encuestas estarán principalmente enfocadas en los resultados, también se obtendrán datos sobre la implementación del programa.

Los cuestionarios estarán formulados de forma que no generen problemas a los padres, madres, profesores/as, implementadores/as y alumnos. Que no haya barreras culturales que impidan el buen desarrollo de la prueba. Las preguntas se consensuarán con los organizadores/as e implementadores/as de SocialClown, Reflexe y profesores/as.

Estos cuestionarios se deberán probar antes para asegurarse de que funcionan bien, no generan dudas y recogen la información necesaria de manera adecuada.

La realización de una prueba piloto es necesaria para asegurar que la herramienta no da problemas y cumple. Todos los encuestadores tomarán parte de la prueba. Los resultados de la prueba piloto pueden servir para validar las preguntas, proponer cambios y saber cuánto tiempo aproximadamente lleva responder a toda la encuesta.

El instrumento para la recolección de datos se puede consultar en el anexo 3. Es importante que sea fácil de manejar y que la información pueda ser posteriormente recopilada electrónicamente para su tratamiento. Se plantea de partida la utilización de cuestionarios en hojas preparadas para lectores ópticos ya que este formato tiene un coste más bajo, aunque se prefiere las encuestas realizadas en soporte electrónico pues aportan facilidad de manejo de datos, reducción de tiempos de gestión, menos errores y buena custodia de la información.

## Encuestas a madres o padres

En las entrevistas que se realizarán a algunos padres y madres, se realizará un pequeño cuestionario sobre asuntos relacionados con el programa para obtener su punto de vista y contrastarlo con los datos aportados en las entrevistas individuales.

El número de ellas que se realizará está por determinar, pero no será grande pues no se busca que tenga significación estadística.

## Observación directa

Para obtener más información proveniente de otra fuente y reforzar los datos se realizarán jornadas de observación directa en algunos colegios del grupo de tratamiento con la intención de poder reforzar los datos obtenidos de las encuestas. Estas sesiones abordan temas delicados como aspectos de higiene, lavarse las manos después de ir al baño o antes de comer y aspectos similares, así como ver qué tipo de comidas toman los alumnos/as. Es por esto por lo que se ha de prestar especial cuidado en este asunto para no molestar a nadie, es por esto por lo que se deberá extremar la prudencia en la observación.

Se elegirán colegios en donde tengan todas las instalaciones sanitarias disponibles y tiempo para la comida.

## 6. Estrategia de muestreo

El contexto, a tenor de las declaraciones de los implementadores y los organizadores, es el mismo para todos los pueblos de la provincia, entorno rural, producción eminentemente agrícola y ausencia de buenas condiciones de vida como son luz eléctrica, agua corriente y saneamiento mejorado. También es sabido que los entornos rurales son más homogéneos que las urbanas (INE.Osuna, 1991, pág. 28)

Por motivos de presupuesto se han elegido por parte de los organizadores, SocialClown y Reflexe, tres pueblos separados entre sí y a una distancia similar de la capital de provincia Bobo Dioulasso. El mejor escenario hubiese sido un proceso de asignación aleatoria pero no ha podido ser así. Los colegios que recibirán el Programa de sensibilización están concentrados en tres pueblos que se han elegido por criterios de presupuesto, proximidad a la capital de provincia y producción agrícola. Este tipo de selección condiciona cómo se conformará la muestra para la evaluación.

Los pueblos elegidos están separados geográficamente y ni ellos ni los colegios sabrán de su elección ni si son grupo de control o de tratamiento. El resto de los pueblos y colegios de la provincia podrían haber sido elegidos perfectamente si el presupuesto del programa hubiese sido otro.

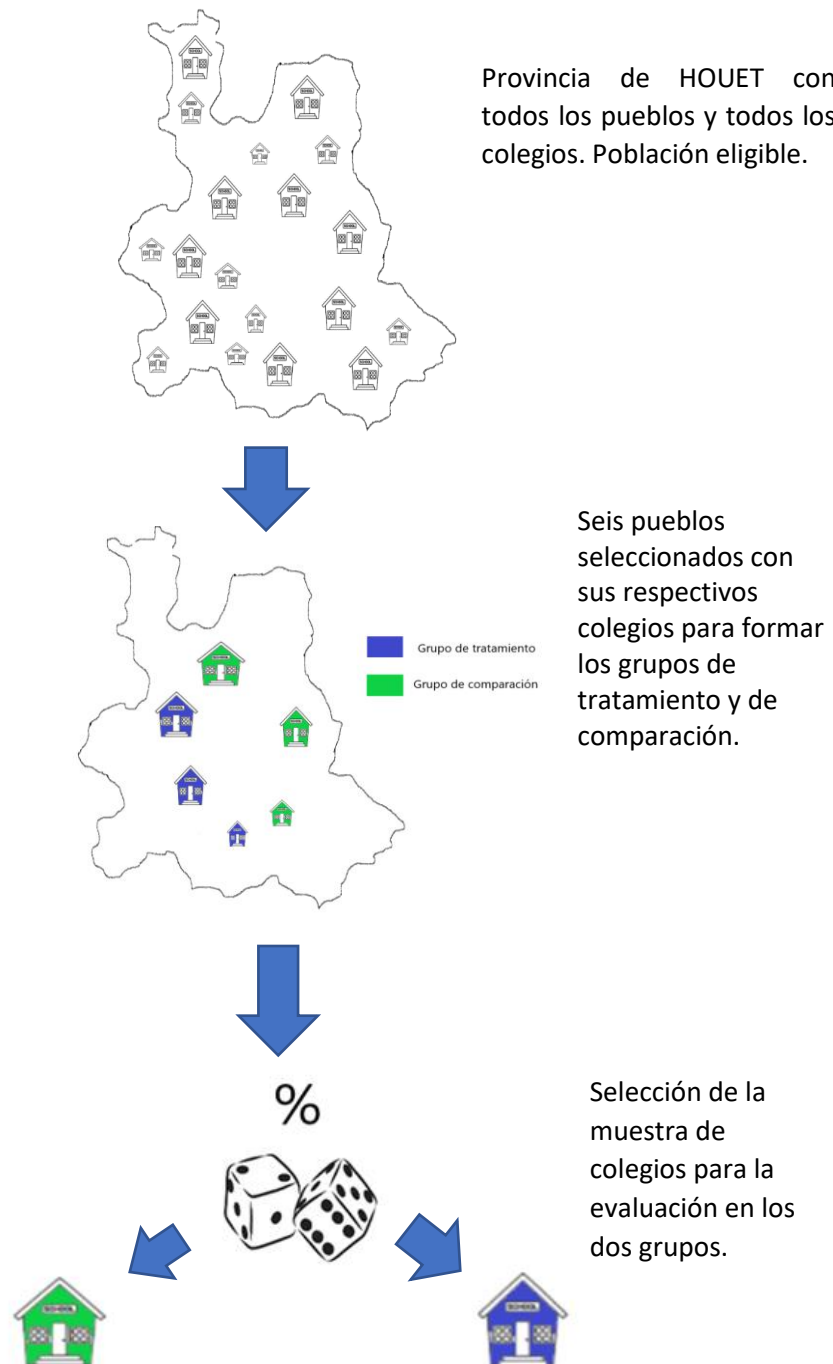
Los pueblos que recibirán el programa de sensibilización son Bama con 40 colegios, Peni con 33 y Toussina con 22. Entre los tres pueblos hay 95 colegios. La población que recibirá el programa son 10500 alumnos/as, esto significa que en los colegios hay de media unos 110 alumnos de ciclo medio 1 que recibirán el programa. Se sabe que el tamaño medio de las clases de CM1 es de 50 alumnos/as, por lo que habrá de media dos clases por colegio.

Para la selección de la muestra se propone hacer un **muestreo por conglomerados monoetápico proporcional aleatorio** en el total de los 95 colegios que recibirán el programa. Cada pueblo conformará un estrato geográfico teniendo así tres estratos, Bama, Peni y Toussiana. Cada colegio se constituirá como un conglomerado. La muestra se distribuirá de manera proporcional al peso de los estratos. A partir del listado de los colegios de cada estrato se extraerán de manera

aleatoria simple los conglomerados que tomarán parte del proceso evaluativo. Los cuestionarios se aplicarán a todos/as los/as alumnos/as de las clases de ciclo medio 1 de los colegios seleccionados. La distribución quedaría como sigue:

- el 42% de los conglomerados serían de Bama
- el 35% de los conglomerados serían de Peni
- y el 23% de los conglomerados serían de Toussiana.

La selección de la muestra para la realización de las encuestas y muestreos es como sigue:



*Figura -2 Distribución de la muestra.*

## 7. Cálculo del tamaño de la muestra. Efecto mínimo detectable

Se debe llegar a un acuerdo con los representantes regionales de Salud, Educación, con los médicos locales, los organizadores y los/as profesores/as para determinar cuáles son los indicadores más adecuados para medir el impacto. Además, hay que concretar con todos/as ellos/as cuales serían los cambios mínimos necesarios en los indicadores para poder considerar que el impacto del programa ha sido positivo en el caso de la malnutrición, de la adquisición de buenos hábitos de higiene y alimentación y que por tanto justificaría la inversión realizada. Este efecto mínimo detectable es el que la evaluación de impacto tiene que ser capaz de identificar (Paul J. Gertler, 2017, págs. 305-306). Estos cambios mínimos son necesarios para el cálculo del tamaño de la muestra. En este caso, por falta de información y comunicación, no ha podido ser posible conocer esos datos mínimos por lo que el cálculo del tamaño de la muestra se ha realizado desde una estimación.

Se han realizado cálculos de potencia estadística y tamaño de muestra para la definición del tamaño del efecto y del efecto mínimo detectable. Para ello se han usado el programa G\*Power de la universidad de Dusseldorf y una herramienta del Banco Interamericano de Desarrollo.

La intención de estos cálculos es saber si la magnitud del impacto causado, si el tamaño del efecto ha sido significativo, más que si en los análisis el p-valor tiene significación estadística. El p-valor indica si ha habido un efecto significativo, pero no dice cómo de grande o importante ha sido.

Hay que recordar que estas eran las hipótesis de partida:

- **Hipótesis nula,  $H_0$ :** *el programa de sensibilización no reduce la malnutrición de los niños/as de la zona ni provee de hábitos de higiene y alimentación buenos para la salud.*
- **Hipótesis alternativa,  $H_1$ :** *el programa mejora la situación de malnutrición y provee de buenos hábitos de higiene y alimentación para la salud.*

La evaluación quiere aceptar la hipótesis alternativa  $H_1$  y decir que el programa consigue cambios en la población y rechazar la  $H_0$ , que significa que no se producen los cambios. Esto supone conocer el valor del error tipo 2 o  $\beta$  (falsos negativos) y la potencia de la prueba  $1-\beta$ . Se estudia pues la probabilidad de que suceda  $H_1$ .

Si  $\alpha$  expresa la probabilidad de equivocarse al rechazar la  $H_0$ , entonces  $1-\alpha$ , que se denomina nivel de confianza, indica la probabilidad de tomar la decisión correcta al aceptar la  $H_0$ . Entonces,  $1-\beta$  (potencia de la prueba) sería la probabilidad de que  $H_1$  sea verdadera (Camacho, s.f.).

Por tanto, se quiere tener un riesgo bajo al rechazar  $H_0$ , eso significa que  $\alpha$  es pequeña y una probabilidad alta de acertar con  $H_1$  y esto significa que  $1-\beta$  es grande. Gráficamente para dos muestras independientes sería como sigue:

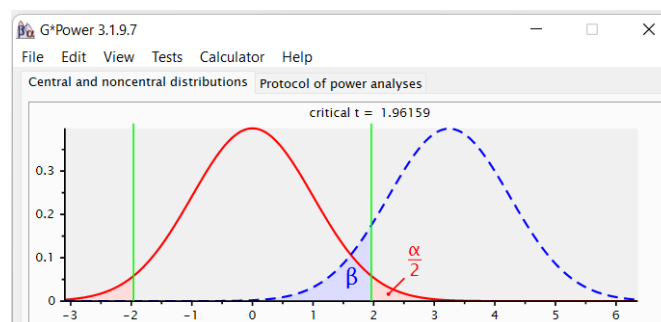


Figura -3 Diferencia de medias. Tamaño del efecto y Potencia (fuente: G\*Power)

Para conocer el efecto mínimo detectable (en adelante, EMD) en los indicadores de resultados se ha utilizado el programa G\*Power y la calculadora del BID (Desarrollo, 2022). Los estudios existentes, por ejemplo en educación, afirman que un aumento de 0,1 desviaciones típicas en pruebas estandarizadas es un impacto relativamente pequeño, mientras que un aumento de 0,5 desviaciones es relativamente grande. (Paul J. Gertler, 2017, págs. 305-306).

Se propone el uso de la calculadora del BID para muestreos por conglomerados y compararla con los resultados del programa G\*Power y tomar la opción más segura para la evaluación.

En el caso del programa G\*Power, se pretende calcular el tamaño de la muestra en función del tamaño del efecto para dos grupos independientes de una población finita. Se ha conversado con los organizadores y piensan que el efecto del programa es pequeño, por lo cual, para hacer los cálculos se ha estimado un efecto leve del 0,2 según la escala de Cohen (Cohen, 1977). Para ello se ha tomado un nivel de confianza del 95% y una potencia del 90% (valores habituales en este tipo de estudios). Con estos datos, el programa G\*Power arroja un resultado de 527 individuos para el grupo de tratamiento y 527 para el de control.

| Test family                                                                      |                                   | Statistical test                                             |           |
|----------------------------------------------------------------------------------|-----------------------------------|--------------------------------------------------------------|-----------|
| t tests                                                                          |                                   | Means: Difference between two independent means (two groups) |           |
| Type of power analysis                                                           |                                   |                                                              |           |
| A priori: Compute required sample size – given $\alpha$ , power, and effect size |                                   |                                                              |           |
| Input Parameters                                                                 |                                   | Output Parameters                                            |           |
| Determine =>                                                                     | Tail(s) Two                       | Noncentrality parameter $\delta$                             | 3.2465366 |
|                                                                                  | Effect size d 0.2                 | Critical t                                                   | 1.9622215 |
|                                                                                  | $\alpha$ err prob 0.05            | Df                                                           | 1052      |
|                                                                                  | Power ( $1 - \beta$ err prob) 0.9 | Sample size group 1                                          | 527       |
|                                                                                  | Allocation ratio N2/N1 1          | Sample size group 2                                          | 527       |
|                                                                                  |                                   | Total sample size                                            | 1054      |
|                                                                                  |                                   | Actual power                                                 | 0.9003604 |

Figura -4 EMD, Potencia y tamaño de la muestra

Con estos datos, 0.2 y 527 unidades, se han llevado a la calculadora del BID. Esta calculadora ha sido desarrollada por el BID para conocer cuál es el efecto mínimo detectable (EMD) para muestreos por conglomerados tanto para prevalencias como para variables continuas en función de los datos de los conglomerados, de la potencia estadística y el nivel de confianza.

Para los cálculos del EMD se asumen los mismos valores de parámetros estándar recomendados por el BID, como son una potencia estadística de 0,90, un nivel de significancia del 0,05, un 95% de nivel de confianza, una correlación intra-conglomerados del 0,05 (algo normal para estos casos) y una desviación estándar de 1 (recomendada). Con estos datos, el tamaño del Efecto Mínimo Detectable de la evaluación sería una variación de 0,2 en las mismas unidades que se haya expresado la desviación típica.

Es decir, el proceso evaluativo si se realiza correctamente, tendría una sensibilidad de 0,2 variaciones en las variables consideradas. Es posible que se den cambios en la población que al ser más pequeños al 0,2 pasarían desapercibidos. Es por ello que se hace crucial y necesario hablar con las autoridades en salud y los organizadores para definir este criterio como se comenta al inicio de este punto.

A continuación, se presentan los cálculos en forma de tabla:

#### Para un indicador cuantitativo (una variable continua)

|                                 |                       |                                    |
|---------------------------------|-----------------------|------------------------------------|
| Idioma                          | 2                     | ( Español )                        |
| Parámetros de diseño            |                       |                                    |
| Potencia                        | $1 - \beta = 90,0\%$  | $(t_{1-\beta} = 1,2816 )$          |
| Nivel de confianza              | $1 - \alpha = 95,0\%$ | $(t_{1-\alpha/2} = 1,9600 )$       |
| Características de la muestra   |                       |                                    |
| Tamaño de los conglomerados     | $m = 50$              | Para datos no agrupados, use $m=1$ |
| Conglomerados de tratamiento    | $k_T = 35$            | $(n_T = 1.750 )$                   |
| Conglomerados de control        | $k_C = 35$            | $(n_C = 1.750 )$                   |
| Correlación Intra-Conglomerados | $ICC = 0,05$          | $(Deff = 3,45 )$                   |
| Características del indicador   |                       |                                    |
| Desviación estándar             | $\sigma = 1,00$       |                                    |
| Mínimo Efecto Detectable        | $\Delta = 0,20$       |                                    |

#### Para una prevalencia

|                                 |                       |                                    |
|---------------------------------|-----------------------|------------------------------------|
| Idioma                          | 2                     | ( Español )                        |
| Parámetros de diseño            |                       |                                    |
| Potencia                        | $1 - \beta = 90,0\%$  | $(t_{1-\beta} = 1,2816 )$          |
| Nivel de confianza              | $1 - \alpha = 95,0\%$ | $(t_{1-\alpha/2} = 1,9600 )$       |
| Características de la muestra   |                       |                                    |
| Tamaño de los conglomerados     | $m = 50$              | Para datos no agrupados, use $m=1$ |
| Conglomerados de tratamiento    | $k_T = 35$            | $(n_T = 1.750 )$                   |
| Conglomerados de control        | $k_C = 35$            | $(n_C = 1.750 )$                   |
| Correlación Intra-Conglomerados | $ICC = 0,05$          | $(Deff = 3,45 )$                   |
| Características del indicador   |                       |                                    |
| Prevalencia                     | $P = 1,0\%$           | $(\sigma = 9,9\% )$                |
| Mínimo Efecto Detectable        | $\Delta = 2,0\%$      |                                    |

*Tabla 1 Cálculo efecto mínimo detectable. Fuente BID*

Después de un proceso iterativo de búsqueda se determina que número adecuado de conglomerados para ésta evaluación sería de 35 con un tamaño de 50 alumnos/as, es decir, 35 colegios de un tamaño de 50 encuestas por colegio, lo que hace un total de 1750 encuestas.

Estos datos se pueden consultar en la siguiente URL:

<https://www.iadb.org/es/topics-effectiveness-improving-lives/evaluacion-de-impacto>

La literatura consultada, BID (Paul J. Gertler, 2017, págs. 312-313), dice que “se requieren al menos entre 40 y 50 conglomerados en cada grupo de tratamiento y comparación para obtener potencia suficiente y garantizar la similitud de las características de línea de base al usar métodos de asignación aleatoria”.

El resultado final sería de 1750 encuestas para que la muestra pueda detectar un efecto leve de 0,2 y que sea representativa en cada uno de los grupos, pero para estar en concordancia con los cálculos anteriores del tamaño de la muestra y asegurar el cumplimiento de los requerimientos para la medición del impacto, se ha decidido tomar una opción conservadora y proponer hacer 2300 encuestas. En esta cantidad se tiene en cuenta un 5% de encuestas que podrían estar mal rellenas y también las necesidades para medir el impacto mediante el método de pareamiento del cual se hablará más adelante y que se estima en un 80%. Por lo tanto, y de manera proporcional, se propone finalmente la realización de las siguientes encuestas en los pueblos como sigue:

- Bama con 950 encuestas en 19 clases/colegios.
- Peni con 800 encuestas en 16 clases/colegios.
- Toussiana con 550 encuestas en 11 clases/colegios.

El grupo de control lo conformarán otras 46 clases de colegios de los pueblos de la zona que tienen similares características y que serán elegidos de la misma manera; proporcional y aleatoria. Se realizarán el mismo número de encuestas que al grupo de tratamiento y en las mismas fechas. Para ello, serán los/as implementadores/as y los/as representantes regionales de Salud y de Educación quienes determinarán estos colegios y pueblos del grupo de control.

Se entiende que este incremento de encuestas supone un coste importante en el presupuesto y también se asegura que el proceso evaluativo es de calidad y los resultados podrán usarse para hacer inferencias al resto de la población de la provincia.

Para reducir costes se propone que en aquellos colegios en donde haya dos o tres clases de ciclo medio 1 de 50 alumnos/as se aproveche el transporte y la logística para abaratar y reducir los tiempos y se realicen allí las encuestas que se puedan. Todos los cálculos anteriores están hechos en base a clases de 50 alumnos/as o colegios de 50 alumnos/as de CM1.

Es sabido que hay indicadores que precisan de muestras más grandes que otros para advertir las diferencias. Es una práctica habitual realizar cálculos de potencia para varios indicadores de resultados para estimar el tamaño de la muestra y del efecto. En este caso y a falta de datos no se realizarán más cálculos de potencia.

## **8. Metodología para la medición del impacto. Operacionalización**

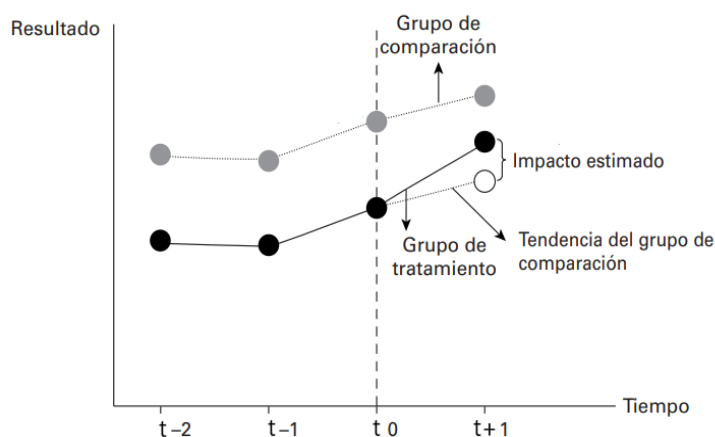
La operacionalización es el proceso que explica cómo se van a contestar las preguntas de evaluación o cómo se van a medir los resultados. A continuación, se explica cómo será este proceso:

Debido al tipo de selección efectuada por parte de los organizadores de los pueblos y colegios y puesto que no ha habido una regla explícita de cómo realizar la asignación del programa, el método que se propone, y que los manuales como el del BID (Paul J. Gertler, 2017, pág. 143) recomiendan es el método de **Diferencias en diferencias**, combinado con el **Pareamiento** estadístico. Se busca así reducir el riesgo de sesgo de selección y confirmar los hallazgos.

En el caso del método de Diferencias en diferencias, se comparan las tendencias de los dos grupos tratamiento y control antes y después del programa. La diferencia entre esas tendencias es el impacto. Es decir, en ausencia del programa, los resultados mostrarían tendencias iguales en los grupos tratamiento y comparación.

Antes de que dé comienzo la implementación del programa se llevarán a cabo encuestas en los dos grupos, tratamiento y control para conocer la situación de partida. Es lo que se llama una **encuesta de línea de base**. Con ella se obtiene la información necesaria con la que se irán haciendo las comparaciones entre los dos grupos. Por lo que su buena calidad es la clave de la evaluación.

Para este método de Diferencias en diferencias, hay que asegurarse de que las tendencias de los dos grupos son más o menos paralelas a lo largo de la línea temporal. Según el manual del BID (Paul J. Gertler, 2017, págs. 305-306), sería necesario realizar dos observaciones de línea de base antes de la intervención (t-2 y t-1) otra antes del inicio del programa (t-inicio o t 0) y otra posterior (t-final o t+1) para los dos grupos, tratamiento y control. Así, el supuesto de igualdad de tendencias para ambos grupos estaría confirmado y se podría decir que los dos grupos estarían en los mismos puntos en ausencia del programa. De esta manera se están controlando las características observables y no observables de cada alumno/a al quedar todo resumido en una tendencia.



*Figura -5 Método de diferencia en diferencias. Fuente BID*

El método por **Pareamiento** busca reforzar la veracidad de los resultados encontrados en el método de Diferencias en diferencias y reducir el problema del sesgo por selección. Para ello, trata de encontrar la mejor pareja, mejor contrafactual, para cada individuo que ha recibido el programa dentro del grupo de control.

El impacto en este caso sería la diferencia promedio en los resultados de los dos grupos, pero sólo de los que han encontrado pareja en el proceso.

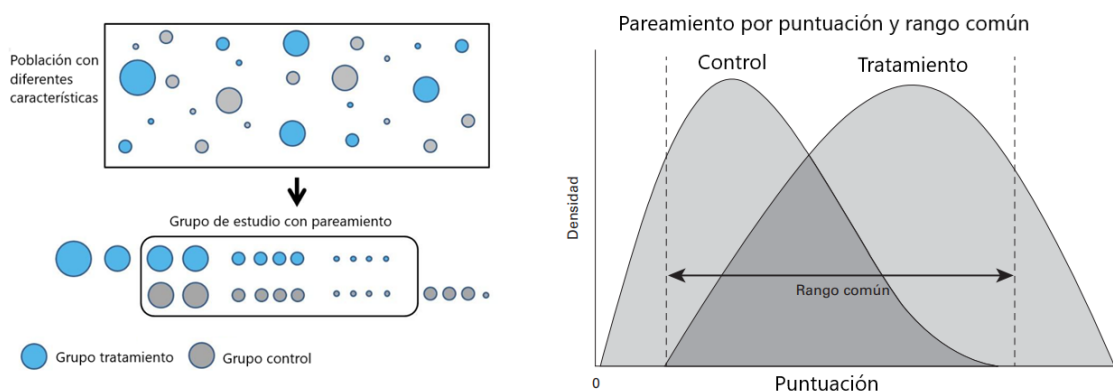
Este método trata de parecerse lo máximo posible a un experimento aleatorio entre individuos que tienen las mismas características. Para llevarlo a cabo se usará el programa SPSS que realiza

el **match** entre las parejas de manera automática una vez la información está cargada en la base de datos. El proceso sigue los siguientes pasos:

- Disponer de encuestas muy similares en los dos grupos.
- Estimar la probabilidad o puntuación de cada individuo que participa, tanto en el grupo de tratamiento como en el de control.
- La muestra resultante estará compuesta de aquellos individuos que se encuentran en un rango común de puntuación entre los dos grupos.
- Identificar las parejas con puntuaciones similares.
- Comparar los resultados entre los grupos de tratamiento y control.
- La media de los impactos individuales da una estimación del impacto promedio del programa.

Todas las variables que se utilizan para realizar el pareamiento son características observables de los individuos y que evolucionan igual en el tiempo que dura el programa. La identificación de estas variables es fundamental para el buen resultado del proceso y están todas recogidas en el anexo 4 llamado encuesta de pareamiento.

A continuación, aparece representado de forma gráfica el pareamiento de individuos de los dos grupos con características similares y el rango común que forman las parejas formadas entre los dos grupos.



*Figura -6 Pareamiento por puntajes de propensión y rango común. Fuente BID*

Es muy probable que el número de parejas sea bastante menor a las 2300, pues en muchos casos no se encontrará una pareja en el grupo de control para cada participante del grupo de tratamiento. Además, habrá un porcentaje de encuestas que estén mal completadas.

El método del Pareamiento se basa en el **Propensity score-matching o puntaje** que se reproduce con SPSS. Éste genera un puntaje o promedio ponderado entre 0 y 1 en función de todas estas variables observables en los dos grupos y genera un nuevo grupo de control (Mackee, 2015), que es con el cual se hará la comparación. Será en este nuevo grupo donde estará la pareja más cercana a cada elemento del grupo de tratamiento.

La puntuación o puntaje de propensión (propensity score) es la probabilidad de que a un individuo con unas características específicas se le asigne el grupo de tratamiento y no el de control. Por eso va de 0 a 1. El programa SPSS calcula esa probabilidad para cada individuo contemplando esas variables.

En la tabla siguiente se muestra un ejemplo realizado sobre tres variables observables donde se aprecia como de un grupo de control de 5 elementos y otro de tratamiento de 4, al final del

proceso el programa compara los 4 elementos del grupo de tratamiento con dos elementos del grupo de control que tienen puntuaciones o puntajes similares.

| Grupos      |   | Variable 1 | Variable 2 | Variable 3 | Puntaje  |
|-------------|---|------------|------------|------------|----------|
| Tratamiento | 1 | 10         | 0,5        | 0,01       | 0,460948 |
|             | 1 | 15         | 0,6        | 0,02       | 0,708763 |
|             | 1 | 22         | 0,7        | 0,01       | 0,876257 |
|             | 1 | 19         | 0,6        | 0,02       | 0,733563 |
| Control     | 0 | 25         | 0,6        | 0,01       | 0,711045 |
|             | 0 | 19         | 0,5        | 0,02       | 0,458197 |
|             | 0 | 4          | 0,1        | 0,04       | 0,011918 |
|             | 0 | 8          | 0,3        | 0,05       | 0,089903 |
|             | 0 | 6          | 0,2        | 0,04       | 0,033546 |

Tabla 2 Ejemplo de puntaje de propensión.

Después se ha de verificar la calidad del emparejamiento, se llaman pruebas de equilibrio del emparejamiento. Para ello se puede hacer comparando las medias de las variables mediante pruebas t y también viendo las diferencias estandarizadas de las variables antes y después del emparejamiento. Se mide mediante el PSTEST según (Ferrer, 2011) y (Leuve y Sianesi, 2003). Se calcula para cada variable la reducción del sesgo. Las diferencias estandarizadas máximas propuestas oscilan entre 10-25%. El cálculo se realiza de la siguiente manera:

$$dif = \frac{100 (\hat{X}_{Trat} - \hat{X}_{Control})}{\sqrt{\frac{(S_{Trat}^2 + S_{Control}^2)}{2}}}$$

Donde *dif* es la diferencia,  $\hat{X}_{Trat}$ , es la media de grupo de tratamiento,  $\hat{X}_{Control}$  es la media del grupo de control,  $S_{Trat}^2$  y  $S_{Control}^2$  es la desviación estándar del grupo de tratamiento y del grupo de control. En la siguiente tabla se muestra un ejemplo:

| Variable |                                            | Medias      |         | % sesgo | % Reducción sesgo |
|----------|--------------------------------------------|-------------|---------|---------|-------------------|
|          |                                            | Tratamiento | Control |         |                   |
| Manos    | No emparejada<br>(toda la variable)        | 0,47932     | 0,29856 | 38,5    | 97                |
|          | Emparejada<br>(sólo individuos con pareja) | 0,47932     | 0,48234 | -0,9    |                   |

Tabla 3 Ejemplo de prueba de equilibrio del emparejamiento de la variable Manos

Lo ideal es que todos los participantes del grupo de tratamiento encuentren pareja, se estima que un 75% de rango común se considera un buen emparejamiento (raymarshallcenter.org, s.f.). En este caso, debido a lo homogénea que es la población se ha estimado en un 80% para los cálculos.

Si por cualquier motivo esto no fuera posible, habría que consensuar con los organizadores y responsables en salud cuáles son las variables observables más importantes para tener en cuenta en la conformación de las parejas para que el número de estas sea suficientemente representativo pues quizás se hayan incorporado muchas variables para realizar las parejas. Se analizarán los datos de la zona común solamente. Por este motivo, se propone tomar una muestra mayor a la calculada en la potencia estadística y en el efecto mínimo detectable.

## Bloque III. Ejecución

### 1. Trabajo de Campo

El trabajo de campo será eminentemente realizado por un equipo contratado en el país. El equipo de Burkina realizará todas esas tareas en colaboración con el equipo de España. Se establecerán los mecanismos oportunos mediante teleconferencias o similar para tener las reuniones de coordinación que sean necesarias. El equipo será externo al programa y a los organizadores.

Para la realización del trabajo de campo debe haber un compromiso alto por parte de quien lo realice pues el programa es largo y el calendario de actividades también por lo que una alta rotación en el personal podría ocasionar pérdida de información y aprendizajes muy valiosos.

Será el equipo de Burkina el que concierte todas las citas con los actores claves en los tiempos previstos y tratará de llegar a acuerdos de consenso sin problemas siempre en comunicación con la parte española.

#### **Equipo de campo**

El equipo de campo en Burkina estará formado por al menos un evaluador senior con experiencia demostrable en evaluaciones de impacto, un equipo de evaluadores junior y un coordinador.

Para la conformación del equipo de campo se llevará a cabo una selección cuidadosa del personal contratado pues serán los responsables de la buena calidad de los datos y resolverán todas las dudas a los/as profesores/as, a las personas en los hogares y al alumnado todas las dudas con las encuestas y entrevistas. Junto con los organizadores se establecerán unos criterios tales como:

- Experiencia demostrable en realización de entrevistas personalizadas cara a cara, por ordenador y/o por teléfono de al menos 2 años.
- Experiencia demostrable en realización de estudios similares relacionados con la salud.
- Conocimientos en gestión de datos y encuestas.
- Conocimientos en alimentación y promoción de la higiene.
- Conocimientos de los hábitos de alimentación e higiene locales.
- Conocimiento de la lengua local y las costumbres
- A ser posible deberían ser locales, de la provincia de Bobo.
- En el caso del evaluador senior los criterios serán distintos, más exigentes y todos demostrables.

El equipo de campo recibirá entrenamiento específico para el trabajo a realizar en las encuestas y en las entrevistas. Se estudiarán las encuestas y se simularán entrevistas, conocerán las técnicas y las actividades realizadas en el programa. Se realizarán pruebas antes del trabajo real para probar la calidad de estas.

Esta formación será llevada a cabo por el coordinador del equipo de campo y el evaluador senior contratado. El horario necesario para completar esta parte no está incluido en el presupuesto pues no se sabe la formación de partida de los evaluadores junior y el tiempo necesario para completarla.

Además, los encuestadores/as, serán los encargados de realizar las valoraciones antropométricas de los niños y niñas por lo que deberán ser adiestrados por un médico local para aprender a realizar las mediciones de manera correcta y registrarla en los cuestionarios.

Si la contratación de este personal no fuera posible se encargará a una empresa especializada la recogida de datos de las encuestas de los alumnos/as y el resto del trabajo lo haría un pequeño equipo de evaluación de dos personas con experiencia en evaluaciones de impacto, el evaluador senior local y un ayudante.

La recogida de datos de calidad ha de ser la preocupación principal del equipo de campo pues de ellos/as depende mucha parte del proceso. Una mala recogida echaría por tierra todo el trabajo realizado, el dinero invertido y los resultados serían malos, dejaría sin efecto la evaluación.

### **Observación directa**

Se realizarán visitas a los colegios que han recibido el programa para comprobar si determinados alumnos tienen conductas distintas a otros en asuntos de higiene o alimentación.

Por supuesto, deberán estar autorizadas por la dirección del colegio y por la autoridad educativa regional. Y se llevarán a cabo con la mayor discreción posible, pues se trata de asuntos de carácter sensible en algunos casos.

No ocuparán mucho tiempo y sólo las pueden realizar personas experimentadas y entrenadas en esta labor, ya que son aspectos sutiles del comportamiento que son difíciles de detectar. Será el evaluador/ra senior local quien realice esta labor.

### **Encuesta línea de base alumnos/as**

El personal contratado será el encargado de explicar las encuestas a los/as alumnos/as, así como de la recogida y custodia de la información. Se realizarán muchas encuestas en diferentes lugares lo que requerirá mucha coordinación con los colegios y los/as profesores/as para entorpecer lo menos posible el horario escolar y ocupar el tiempo sólo lo necesario.

La logística de los colegios y los transportes tienen que estar resueltas de antemano pues los colegios están bastante dispersos por el territorio. Además, quién realice la recogida de los datos deberá coordinarse con la asociación Reflexe, que es quien implementa el programa, así como con las autoridades educativas.

Se establecerá un calendario adecuado a las distintas actividades de acuerdo con la autoridad educativa y los organizadores. Hay cuatro momentos clave en los dos grupos para realizar las encuestas que son t-2, t-1, t-inicio y t-final. Están claramente definidos en el calendario y están explicados anteriormente en el apartado de análisis cuantitativo, diferencias en diferencias.

Es muy importante disponer de datos de línea de base para el proceso. La intención es alcanzar a todo el alumnado de los dos grupos en la misma época del año para evitar posibles sesgos debidos por ejemplo a la estacionalidad, las lluvias, la recogida de cosecha, etc.

Según la distribución realizada, completar una ronda de encuestas en cada uno de los grupos llevará cerca de una semana a tres encuestadores/as, por lo que se necesitan dos semanas máximo en cada momento clave.

La recogida de datos ha de ser de la mayor calidad posible, por lo que es imprescindible que quien la haga tenga experiencia suficiente. El procedimiento de recogida de datos deberá ser el mismo para ambos grupos

Los datos de línea de base son parte esencial de esta evaluación para probar que su diseño es bueno y que el programa obtiene los resultados esperados.

### **Entrevistas semiestructuradas**

Estas entrevistas las realizará el equipo de campo y se espera que sean completadas en tres o cuatro días. Se harán entre 10 y 20 entrevistas a madres o padres de niños o niñas que han participado en el programa y a 10 profesores/as que hayan estado en contacto directo con él. Dependiendo de las encuestas de los niños/as, se tomarán padres y madres que hablen francés si es lo habitual, sino se harían en lengua local.

También se realizarán con el personal de SocialClown y Reflexe.

En el caso de los/las docentes se necesitará permiso de la autoridad educativa regional, así como del director/a del centro escolar y se coordinarán con las agendas de los/as profesores/as/as para no interferir en los horarios escolares.

Todas serán grabadas y transcritas para su posterior análisis. El guión se puede consultar en el anexo 3. Aunque se permite cierta flexibilidad, que el entrevistador usará para conseguir la mejor información posible dependiendo del entrevistado, se intentará contestar a las preguntas del guión. Las entrevistas serán conducidas por el evaluador/ra senior local con experiencia suficiente en estas labores.

### **Entrevistas de grupo**

Estarán conformados por unas seis personas y un moderador. Éste será una persona especializada en estas labores que planteará las cuestiones importantes sobre las que se quiere debatir y recabar información. En este caso, este papel lo desempeñará el/la evaluador/ra senior local. En el anexo 2 se puede consultar el guión con los temas sobre los que se quiere obtener información.

Las reuniones serán grabadas para poder transcribirlas y realizar luego el análisis del discurso mediante el programa Atlas-Ti, así como para poder triangular con otros datos. No serán muy largas, del orden de una y media a dos horas. El equipo de terreno tendrá que coordinarse con los padres y madres de los diferentes pueblos y con los/as profesores/as/as y localizar un lugar adecuado para todos. También se encargará de la logística del lugar, de tener preparado todo lo necesario para la grabación, así como un aperitivo y unas bebidas para todos los participantes.

Les comunicarán por adelantado la cita y la hora donde se realizarán y en qué consiste la actividad, así como cuánto durará y cuáles son los objetivos.

### **Monitoreo**

Esta fase es tan importante como el resto, pues asegura que los datos del alumnado que forman parte de los grupos no varían y, en caso de hacerlo se detecta, y se tiene en cuenta en el análisis posterior de los datos. Además, sirve para saber cómo se está implementado el programa y cómo se usan los recursos.

La encuesta de seguimiento se asegura que los factores externos no han variado de forma distinta para un grupo y otro. Puede que no muestre el impacto, pero podría dar cierta información del impacto a largo plazo y sobre todo informará del programa.

Puesto que el programa lo conforman cuatro actividades a lo largo del año escolar, separadas temporalmente de forma equitativa, se propone la realización de tres tomas de datos en el monitoreo, intercalados entre estas actividades.

## 2. Análisis e interpretación de datos

A continuación, se puede ver un esquema del proceso de análisis de la información con los diferentes caminos que se siguen en función del método empleado, las herramientas y las fuentes:

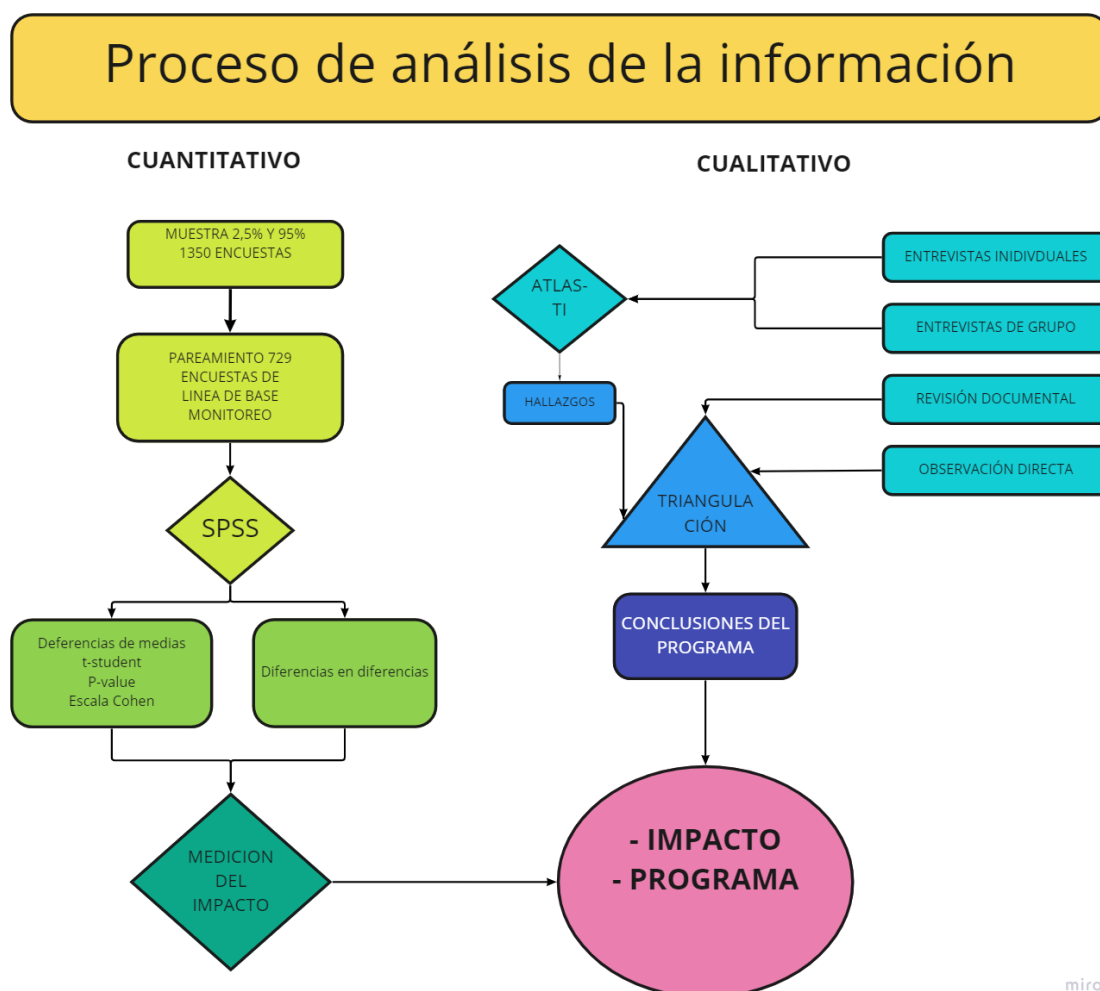


Figura 7 Proceso de análisis de la información obtenida

### Triangulación

La triangulación de los datos cualitativos y cuantitativos se aplicará para profundizar en el conocimiento del programa. Este método (Denzin, 1978; Fassio, Pascual y Suárez, 2004) es la combinación de distintas técnicas de recolección de información, de fuentes de datos distintas y de métodos de investigación diferentes en una misma evaluación. El análisis en una matriz de doble o triple entrada es lo más habitual. El éxito de esta parte del trabajo depende de la buena organización del material informativo.

En este caso, serán datos cualitativos y cuantitativos extraídos de encuestas, entrevistas, observación directa, documentación y recolectados por evaluadores distintos y con métodos diferentes. El objetivo es lograr la complementariedad del enfoque cuantitativo y cualitativo para aportar mayor profundidad en el análisis del programa.

Es un método de convergencia de hallazgos donde unos datos complementan a otros para confirmar resultados. Para realizar esta triangulación, la información ha de estar bien organizada, clasificada por categorías de forma «estandarizada», pues el volumen que se generará es muy grande y será necesario poder localizarla y contrastarla fácilmente.

### Análisis cuantitativo

Para el análisis de los distintos datos de los cuestionarios se utilizará el programa SPSS y Excel. Las gráficas y las tablas de los documentos se realizarán en formato de Power-Bi o Excel según convenga.

Los datos cuantitativos que se derivan de la matriz de evaluación se deben a los indicadores seleccionados y se usarán en la triangulación. Como las preguntas de evaluación de las entrevistas individuales y de las entrevistas de grupo sólo se refieren al programa, sólo se harán al grupo de tratamiento, por lo que no hay comparación posible.

Una vez realizadas todas las encuestas de línea de base y monitoreos a los/as alumnos/as, se tratarán los datos obtenidos con el programa SPSS y mediante las variables que se determine se valorará el impacto de la siguiente manera:

- Se utilizará **las diferencias de medias** y la **t-student** entre los dos grupos para saber si estas diferencias son significativas estadísticamente. Para ello se seleccionarán las variables que se consideren más oportunas y se compararán entre los dos grupos.
- También se podrá conocer la magnitud del cambio con la escala de Cohen (Cohen, 1977) y así poder calcular el efecto del impacto (**EI**) mediante la escala de Cohen en cada una de las variables:

$$d = \frac{\bar{X}_{Tratamiento} - \bar{X}_{Control}}{\sqrt{\frac{S^2_{Tratamiento} + S^2_{Control}}{2}}}$$

Donde  $\bar{X}$  es la media de cada uno de los grupos y  $S^2$  es la desviación estándar de cada grupo. Dicho autor establece los siguientes valores como referencia en términos de desviación estándar:

- d entre 0,2 y 0,5 es un efecto leve.
- d entre 0,5 y 0,8 es un efecto medio.
- d mayor a 0,8 es un efecto grande.
- Se utilizará la **Diferencia en diferencias** también para saber cuál ha sido el impacto del programa en cada una de las variables. Estas tendrán lecturas del impacto como sigue por ejemplo para la variable Manos:

$$Impacto_{Manos} = [\bar{Y}_{ManosTrataDespues} - \bar{Y}_{ManosControlDespues}] - [\bar{Y}_{ManosTrataAntes} - \bar{Y}_{ManosControlAntes}]$$

| Variable Manos ( $\bar{Y}$ ) | Control (A) | Tratamiento (B) |
|------------------------------|-------------|-----------------|
| Antes del programa (1)       | $\bar{Y}A1$ | $\bar{Y}B1$     |
| Después del programa (2)     | $\bar{Y}A2$ | $\bar{Y}B2$     |

$$Impacto_{Manos} = (\bar{Y}B2 - \bar{Y}A2) - (\bar{Y}B1 - \bar{Y}A1)$$

$$Impacto_{Manos} = (\bar{Y}B2 - \bar{Y}B1) - (\bar{Y}A2 - \bar{Y}A1)$$

Las diferencias se pueden hacer de dos maneras, en un caso entre grupos y en otro entre periodos, aunque el resultado del impacto es el mismo.

Se considerarán que los valores de **p-valor** inferiores a 0,05 tienen un resultado significativo estadísticamente. Es decir, ha habido un efecto significativo del programa con una un nivel de confianza del 95%, sin embargo, eso no dice cómo ha sido de grande el efecto. Para saberlo, se usará **el efecto del impacto (EI)** mediante la escala de Cohen.

Hay que recordar aquí que el efecto mínimo detectable según los cálculos explicados antes es de 0,17 uds. Podría pasar que hubiese cambios por valor de 0,16 uds, lo que sería un efecto muy leve en términos de la escala de Cohen pero que no se podrían apreciar con los medios y datos que se tienen.

Todas las variables serán dicotómicas. Las que tienen más de dos opciones de respuesta en la encuesta se codificarán de forma que sean 1 ó 0. Sucede lo mismo tanto en el apartado de higiene como en el de alimentación.

Las variables siguientes son un ejemplo sobre la encuesta de línea de base y serán las comparadas en ambos grupos:

| Higiene         |               |                  | Alimentación |                |                  |
|-----------------|---------------|------------------|--------------|----------------|------------------|
| Manos           | HigieneCuerpo | OtraHigiene      | CuandoComo   | CuantoComo     | MasAlimentacion  |
| EnseñadoManos   | Dientes       | NecesitaConocer1 | Horario      | AguaSuficiente | ImporComerSano   |
| ImportanteManos | Ducha         | Bienvisto        | VecesComer   | AguaBuena      | Preocupa         |
| Transmitenferme | Pelo          | Porque           | Hambre       | Leche          | TodaComida       |
| ContarManos     | Uñas          | Jabón            | TodoslosDias | Fruta          | NecesitaConocer2 |
| CuandoManos     | Intimo        |                  |              | Carne          |                  |
| QuienManos      | Ropa          |                  |              | Verdura        |                  |
|                 | Zapatos       |                  |              |                |                  |

*Tabla 1 Variables posibles y asociaciones para análisis*

Mediante SPSS se podrán realizar distintos agrupamientos con estas variables para crear nuevas variables como en este caso (**Manos, HigieneCuerpo, OtraHigiene, CuandoComo, CuantoComo, MasAlimentacion**) que reúnan información que permitan una visión más general del conocimiento y determinar así otros resultados. Se podrá realizar la desagregación por sexos, estudios de los padres, de las madres, de tipos de casas (con baño, sin él, con agua, sin ella, etc). También se podrán hacer simulaciones para controlar variables que puedan estar sesgando el estudio, etc. Con todos estos datos se determinará el impacto en la población que ha recibido el programa en contraste con la que no lo ha recibido.

Además de estas variables se estudiarán las valoraciones antropométricas que se han registrado y que son indicadores directos de malnutrición, aunque se sabe que los resultados a largo plazo son los más interesantes y esta evaluación no los recoge.

Los resultados habrá que interpretarlos como una probabilidad mayor o menor de reducir la malnutrición con la adquisición de los nuevos hábitos de higiene y alimentación debidos al programa.

### Análisis cualitativo

Se analizará el discurso o relato de las entrevistas individuales y de grupo. El objetivo de esta técnica cualitativa es encontrar la relación que hay entre lo dicho y la realidad de la cual se habla,

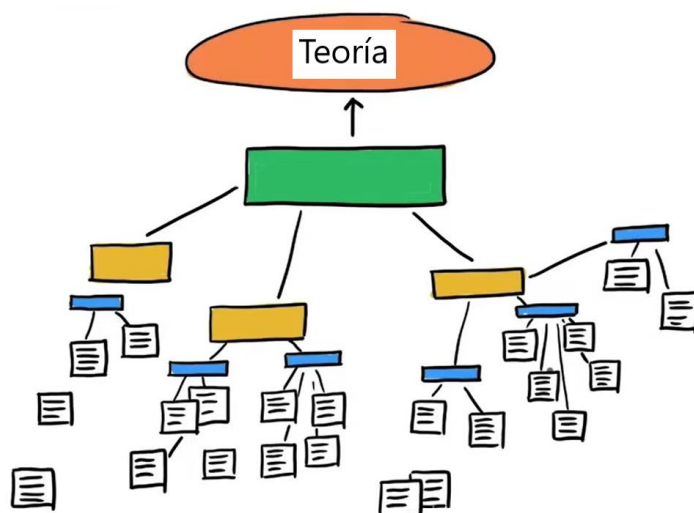
para encontrar de dónde y por dónde fluyen las ideas. El método empleado será la saturación de ideas, las recurrencias en los relatos, la repetición y agrupación. Es decir, la suma de ideas de similares características en torno a un concepto que los diferentes hablantes poco a poco van conformando en grupos homogéneos de ideas, denominados campos semánticos. Estas ideas comunes pueden ser intercambiables en los discursos de las personas entrevistadas. Por eso tienen tanto valor.

Es habitual que se produzcan contradicciones y opiniones muy diversas que enriquecen el proceso dando más puntos de vista. La intención es encontrar los puntos comunes a todos para conformar un relato del programa que converja con los datos e ideas obtenidos por otras fuentes.

Para este método se utilizarán las entrevistas en grupo y las individuales, todas ellas grabadas y transcritas. Se usarán junto con el programa **Atlas-ti**, que realiza parte de este trabajo de formar los campos semánticos. El proceso para realizar este análisis sigue las siguientes fases (Ligero, 2021):

- Todas las entrevistas han de estar transcritas y anonimizadas para descontextualizar el discurso del momento en que se produce.
- Realizar una lectura minuciosa para identificar las ideas que emergen e identificar los símbolos o palabras clave, llamados incidentes (Trinidad y otras, 25-26).
- Disponer los grupos según las relaciones percibidas en el discurso.
- Nombrar los campos semánticos.
- Escribir una síntesis narrativa que dé sentido a los diferentes campos de ideas. Es decir, construir una nueva sintaxis.
- Elaborar una narración más amplia que explique la síntesis, pudiendo incorporar textos literales de los/as informantes (verbatim).
- Elaboración de una teoría fundamentada en los datos recolectados y analizados anteriormente.

En el siguiente dibujo se intenta resumir toda esta idea denominada Grounded theory, donde aparecen los trozos del discurso asociados a las ideas comunes que conforman el campo semántico y que al final dan lugar a la teoría. Se parte desde los datos hacia la teoría.



*Figura 8 Esquema del análisis del discurso por campos semánticos*

En la evaluación orientada por la teoría del cambio, tras el análisis, se debería estar en condiciones de contestar a varias preguntas sobre el programa como son (Ligero, 2017):

- ¿Se han logrado los efectos finales deseados?
- ¿Se han logrado los productos previstos?
- ¿Los procesos se han ejecutado bien? En caso contrario, ¿Cuáles han sido los motivos de una mala ejecución?
- ¿Los procesos son capaces de explicar los resultados provocados por el programa?
- ¿Existen relaciones casuales entre las dimensiones del programa, estructura, procesos y resultados que expliquen las claves del éxito?

Una vez terminado el análisis cuantitativo y cualitativo es el momento de integrar toda esta información en un nuevo esquema interpretativo (Serrano y otros, 2009) con la intención dar un sentido a todas las informaciones obtenidas plasmándolas en un texto, relato o discurso resumido que refleje todo lo recogido de una forma clara y concreta. Esta interpretación deberá responder a los propósitos principales de la evaluación. Si no es posible responder a alguna de las preguntas de evaluación es necesario explicar los motivos.

Todas las interpretaciones deberán estar expresadas con lógica en base al análisis de los datos. Cualquier suposición implícita en el análisis de los datos debe explicarse claramente para que no genere confusión.

### **3. Enjuiciamiento**

Las valoraciones que se hagan deben estar apoyadas por los análisis anteriores, mostrando una línea clara de razonamiento que sustenten los hallazgos y conclusiones. Se utilizará un lenguaje fácil, claro y que exprese los juicios con asertividad. Sobre todo, estas valoraciones tienen que ir encaminadas a proporcionar más información sobre los propósitos de la evaluación que en este caso es su impacto y saber cómo funciona el programa.

Se establecerá un sistema de valoración (positivo, negativo y neutro) para explicar en cada uno de los hallazgos como estos contribuyen o no al impacto del programa, al logro de los resultados,

a la implementación del programa y a los procesos, si estos están bien ejecutados o no y si los diferentes elementos que conforman la estructura del programa son adecuados o no. El objetivo es mostrar los elementos positivos del programa y las metas alcanzadas, los elementos que no aportan y aquellos que lo hacen de forma negativa.

Esta valoración se basará en los criterios que deben ser acordados con los organizadores y con las autoridades y con los resultados obtenidos.

Con toda la información analizada, se reconstruirá o se confirmará en su caso el modelo lógico del programa, esta vez, basado en la evaluación realizada y se contrastará con el propuesto en la teoría del programa.

Este enjuiciamiento estará realizado por el evaluador senior del equipo de Burkina Faso y el equipo de España.

#### 4. Comunicación de resultados

Tal y como propone la OCDE, se comunicarán los principales hallazgos a los que se ha llegado tras todo el proceso, las lecciones aprendidas, las buenas prácticas realizadas por el programa y sus implementadores, así como las recomendaciones. (DAC,2010:13)

Se sugiere la presentar un primer informe preliminar a los organizadores, implementadores y profesores/as con los principales hallazgos e interpretación de los datos para no malinterpretar la situación. Se dará la oportunidad de que lo puedan consultar el informe y hacer aportaciones y comentarios que podrán ser incluidos posteriormente en el informe final si se considera pertinente.

La intención es que la evaluación cumpla con sus propósitos iniciales y que además influya en los actores y se apropien algo de ella.

Los informes que se hagan sobre la evaluación han de ir encaminados a incrementar el conocimiento sobre el programa y sus efectos. Es más probable que se produzcan cambios en el caso de que sean necesarios si se entiende claramente.

#### 5. Recomendaciones

Las recomendaciones se harán en base a los hallazgos y conclusiones de manera que sea clara la intención de lo que se quiere cambiar o mejorar.

Estarán dirigidas a los responsables de salud, educación, a SocialClown, a Reflexe y los/as profesores/as.

Serán claras, realistas, alcanzables y enfocadas al stakeholder que compete. Siempre que se pueda se darán herramientas para poder conseguirlas.

Se indicarán aquellas partes del programa que funciona y necesita mejorarse y aquello que no funciona. Además, es interesante que tengan un seguimiento para ver que se consiguen los resultados que se esperaban.

Sería preferible tener pocas recomendaciones y si no es así se harán por categorías o prioridades para el programa.

## Bloque IV.

### 1. Calendario

Para llevar a cabo la evaluación en tiempo y forma se proponen las siguientes fechas para realizar cada una de las recogidas de información necesarias acopladas al calendario escolar. Es necesario tener el visto bueno de la autoridad educativa, de la dirección del colegio y la participación generosa de los profesores y alumnos. En la tabla siguiente aparece resumida toda la información.

| Fecha aprox.                     | Técnica                               | Descripción                                                               | Muestra                                                    |
|----------------------------------|---------------------------------------|---------------------------------------------------------------------------|------------------------------------------------------------|
| Noviembre-22                     | Encuesta, cuestionario en el colegio  | Encuesta línea de base t-2                                                | Grupo de tratamiento 2300 y 2300 Grupo de control          |
| Mayo-23                          | Encuesta, cuestionario en el colegio  | Encuesta línea de base t-1                                                | Grupo de tratamiento 2300 y 2300 Grupo de control          |
| Septiembre-23                    | Encuesta, cuestionario en el colegio  | Encuesta línea de base INICIO Pareamiento                                 | Grupo de tratamiento 2300 y 2300 Grupo de control          |
| SEPT-23-ACTIVIDAD TEATRO FORUM-1 |                                       |                                                                           |                                                            |
| Octubre-23                       | Encuesta, cuestionario en el colegio  | Encuesta monitoreo                                                        | Grupo de tratamiento 2300                                  |
| NOV-23-ACTIVIDAD TEATRO FORUM-2  |                                       |                                                                           |                                                            |
| Enero-24                         | Encuesta, cuestionario en el colegio  | Encuesta monitoreo                                                        | Grupo de tratamiento 2300                                  |
| FEB-24-ACTIVIDAD TEATRO FORUM-3  |                                       |                                                                           |                                                            |
| Marzo-24                         | Encuesta cuestionario en el colegio   | Encuesta monitoreo                                                        | Grupo de tratamiento 2300                                  |
| ABR-24-ACTIVIDAD TEATRO FORUM-4  |                                       |                                                                           |                                                            |
| Mayo-24                          | Entrevistas individuales cuestionario | Entrevistas a profesores y padres que han tenido contacto con el programa | 20 profesores y 20 madres o padres                         |
| Mayo-24                          | Observación directa                   | Visita a varios colegios para observación directa de hábitos              | Diferentes colegios que han recibido el programa           |
| Junio-24                         | Encuesta, cuestionario en el colegio  | Encuesta línea de base FINAL                                              | Grupo de tratamiento 2300 y 2300 Grupo de control          |
| Junio-24                         | Entrevistas grupales                  | Entrevistas a profesores y padres que han tenido contacto con el programa | 2 grupos. Uno de 6 padres/madres y otro de 6 profesores/as |

*Tabla 2 Calendario aproximado del trabajo de campo*

En el peor de los casos, suponiendo que las encuestas del alumnado se realicen entre tres personas, se necesitaría del orden de una semana para poder recorrer todos los colegios de un grupo y otra semana para los del otro grupo para cada una de las recogidas de información.

Un/a encuestador/a podría realizar cuatro visitas al día a diferentes colegios y podría completar 120 encuestas por lo que al cabo de cinco días haría las encuestas en Bama y en otros cuatro haría Peni y en tres días podría hacer Toussiana. Para el grupo de control sucedería más o menos lo mismo.

Para las entrevistas individuales hay más flexibilidad pues son 20 las que hay que realizar y se realizarían en la parte final del programa, cerca de junio-2024.

## 2. Presupuesto

Para el cálculo del presupuesto se ha utilizado la herramienta de [www.evaluationcost.org](http://www.evaluationcost.org). El precio de hora de evaluador/a senior en Burkina se ha estipulado en 15€, para el/la junior, 10€, sabiendo que según [www.fdex-es.com](http://www.fdex-es.com) el salario medio de un abogado está en torno a los 200€ mensuales y el de un médico, en los 220€ mensuales. En la parte de la evaluación que se realiza desde España el precio es de 30€/h para el evaluador/a senior y de 20€/h para el/la junior.

En los costes extras están incluidos costes para las entrevistas individuales y las de grupo, ya que según SocialClown los padres y madres no las hacen gratis, sino que cobran algo por ello. Además, en este dinero se ha incluido el transporte de los evaluadores a los pueblos para la realización de las encuestas. En las entrevistas de grupo quizás haya que alquilar algún local y preparar alguna merienda o similar para hacer atractivo el evento a los/las asistentes.

Hay un coste que no se ha podido estimar y que es muy importante, se trata la traducción de las entrevistas individuales y las de grupo para poder hacer el análisis del discurso de la mejor manera posible. Una traducción mediocre puede echar al traste todo el esfuerzo anterior. Las encuestas de los alumnos/as pueden ir solo numeradas y en el idioma local pues es un lector óptico quien toma los datos de SÍ o NO. En España, el precio por hora de trabajo oscila desde los 12€ a los 15€ la página de unos 2100 caracteres.

Tampoco se ha podido estimar el coste que tendría si hay que traducir todo a la lengua local.

El precio de las transcripciones se puede mejorar mucho con los programas informáticos como los que hay en el paquete de Outlook que permiten hacerlo de manera rápida y fácil por muy poco dinero. Pero para ello tienen que estar en castellano o en francés que es el idioma que habla en Burkina.

El presupuesto calculado aquí es para un equipo de evaluación confeccionado entre España y Burkina. Si se ha de contratar a una empresa especializada en la recogida de datos para las encuestas el presupuesto sería otro distinto.

El presupuesto pormenorizado se puede consultar en el anexo 7. La versión resumida es la que aparece a continuación, con estos precios/hora como referencia.

|                     | Eva-Senior-España | Eva-Junior-España | Eva-Senior-Burkina | Eva-Junior-Burkina | Coste       |
|---------------------|-------------------|-------------------|--------------------|--------------------|-------------|
| <b>Total</b>        | 304,43 h          | 88,23 h           | 167,73 h           | 836,41 h           | 16.907,02 € |
| - Encargo           | 37,50 h           | 20,00 h           | 18,00 h            | 18,00 h            | 1.795,05 €  |
| - Diseño            | 34,57 h           | 28,23 h           | 38,73 h            | 46,07 h            | 2.219,50 €  |
| - Trabajo de campo  | 0,00 h            | 1,00 h            | 41,00 h            | 667,34 h           | 3.916,67 €  |
| - Análisis de datos | 83,36 h           | 7,00 h            | 26,00 h            | 25,00 h            | 3.025,80 €  |
| - Interpretar       | 13,00 h           | 4,00 h            | 0,00 h             | 0,00 h             | 510,00      |
| - Enjuiciar         | 11,00 h           | 0,00 h            | 0,00 h             | 0,00 h             | 330,00 €    |
| - Recomendaciones   | 4,00 h            | 0,00 h            | 0,00 h             | 0,00 h             | 120,0 €     |
| - Informes          | 111,00 h          | 8,00 h            | 0,00 h             | 0,00 h             | 3.490,00 €  |
| - Otros costes      | 10,00 h           | 20,00 h           | 40,00 h            | 80,00 h            | 1.500,0 €   |

*Tabla 3 Presupuesto estimado de la evaluación*

### 3. Protocolos éticos

Toda la información se tratará con sumo cuidado y confidencialidad. Los datos proporcionados se conservarán durante los años necesarios para cumplir con las obligaciones legales. Los datos no se cederán a terceros salvo en los casos en que exista una obligación legal.

Como se trata de datos personales, las personas involucradas en el proceso tienen derecho a acceder a sus datos personales, rectificar los datos inexactos o solicitar su supresión cuando los datos ya no sean necesarios.

Todo el personal encargado del tratamiento de los datos se encuentra obligado a guardar secreto profesional.

Los evaluadores respetan las principales normas y códigos de conducta profesionales y éticos que les atañen. La evaluación se lleva a cabo con integridad y honradez. Los directores y gestores de la evaluación y los evaluadores deben respetar los derechos humanos y las diferencias culturales y en las costumbres, creencias y prácticas religiosas de todos los actores implicados. (OCDE, 2010).

### 4.Utilización de los Resultados

Según David Bonbright (Bonbright, 2012) “Es más probable que se usen las evaluaciones de impacto cuando se han anticipado y planificado sus usos durante las primeras etapas de la

evaluación y, aún mejor, durante las etapas de planificación del trabajo que se está evaluando” a posteriori no funcionan tan bien.

El hecho de incorporar a los usuarios de la evaluación en los diferentes procesos hace que se produzca cierto compromiso con el uso de esta y se incorporan mejor los hallazgos, conclusiones y recomendaciones. Parte de la intención es que produzca una cierta apropiación de la evaluación y sea tanto de los usuarios como del equipo evaluador. Cuanto más se consiga esto, mayor uso y aceptación tendrá.

## 5.Limitaciones de la evaluación

### **Idioma**

La lengua local en los colegios de la zona es la misma y no se sabe si todos los padres y todos los niños manejan el francés a la perfección, de no ser así, habría que traducir muchas encuestas y entrevistas del francés al idioma local para luego después hacer el camino inverso y más costoso para llegar a tenerlo en castellano con la considerable pérdida de información por el camino, en la expresión y en las palabras. La traducción al francés supondrá un reto importante que a todos los niveles afectaría de manera importante, incluido al presupuesto.

El idioma para los asuntos de coordinación entre los dos equipos, el de Burkina y el de España, puede que sea otra barrera para considerar.

De no solventarse con holgura esta barrera idiomática desde el principio dará muchos problemas a la larga y complicará mucho la evaluación haciéndola inservible.

### **Distancia, coordinación**

La distancia entre Burkina Faso es grande y el costo de los billetes es alto acompañado del visado. Esto hace que la coordinación virtual y telefónica tenga que ser excelente entre los dos equipos, de lo contrario será un fracaso.

### **Disponibilidad de especialistas**

Si no se encuentra especialistas para conformar el equipo de Burkina para realizar el trabajo de campo sería necesario hacerlo con el equipo de España y contratar a una empresa que haga todo el trabajo de campo. Eso sería muy costoso y podría suceder que no diese los mejores resultados al externalizar parte del proceso y tener alta rotación de personas.

Es importante la fidelización al proceso para percibir los cambios. La alta rotación de personal haría que se perdiese la experiencia y parte de información.

### **Presupuesto**

El hecho de que el proceso evaluativo sea tan largo hace que sea complejo de llevar a cabo y costoso sin tener en cuenta las limitaciones anteriores mencionadas.

La recolección y tratamiento de todas las encuestas previstas suponen un coste importe y probablemente sea mayor al estimado aquí. Dependiendo de las herramientas disponibles en el lugar para la gestión de las encuestas y las entrevistas los costes serán mayores o no.

Este punto y el del idioma son cruciales en esta evaluación.

### **Falta de información, bases de datos**

La malnutrición es un asunto muy complejo con muchas variables a tener en cuenta de las cuales las autoridades de la zona no tienen información a tener de las declaraciones de la organización SocialClown. Hay otras organizaciones trabajando en la zona en asuntos de higiene y alimentación de las cuales no se sabe si tienen información precisa y reciente. Sería un asunto para consultar.

### **Duración en el tiempo**

El hecho de que dure más de un año y medio y que la recogida de datos esté espaciada en ese tiempo hace que sea complejo. Esto hace difícil que los equipos sean los mismos con lo que las lecciones aprendidas y experiencias no se acumulen y se pierdan. Eso podría afectar a la calidad de la recogida de datos y al resultado de la evaluación.

La falta de un horizonte temporal más amplio para hacer mediciones evita que se pueda determinar con claridad el impacto del programa sobre todo en los indicadores de malnutrición, pues los registros de los centros de salud tardarán en repercutir en sus registros cualquier cambio positivo que se produzca como consecuencia de nuevos comportamientos en la práctica de hábitos de higiene y de alimentación y el efecto que se pretende de diáspora de conocimientos con los niños/as que pasan por el programa. Este punto es muy importante.

### **Reflexiones personales**

Recientemente, un grupo de alumnos/as del máster y yo hemos terminado una evaluación para la AECID sobre el MAP 2018-2021 de Bolivia-España. Quería dejar aquí algunos de los aprendizajes que creo pueden ser importantes.

1. Cuando un/a evaluador/a solicita información a la persona que como en este caso está en Burkina implementando estos programas, la sensación es como si fuese una auditoría y hay cierta reticencia a proporcionar todo lo que se reclama. La única intención del evaluador/a es hacer su trabajo lo mejor posible para mejorar el programa, para rendir cuentas más apropiadamente o para enfocar mejor los procesos y conseguir más impacto. No se evalúa el trabajo de nadie. La evaluación es una oportunidad para aumentar el impacto del programa y para mejorar el conocimiento que se tiene de él, de quienes están en contacto con él. No es una amenaza. Este mensaje debería quedar cristalino en las reuniones previas para que las personas en contacto con la evaluación colaboren tranquilas, si es posible y se apropien de ella.

El evaluador/a no debe hacer una labor de investigación interna en la sombra si no le ponen las cosas fáciles, no encontrará eso que le ocultan, la evaluación será un fiasco y nadie saldrá contento. No se mejorará, que es el objetivo.

2.El equipo evaluador mantuvo muchas reuniones con la AECID, pero nunca se estableció un sistema de actas o similar para consultarlas en caso de duda. Se alcanzaron muchos acuerdos de manera rápida, pero no fueron muy buenos para el equipo y la independencia del equipo evaluador quedó comprometida. No se supo dejar claro cuáles eran las líneas que el equipo no se estaba dispuesto a cruzar.

Se aconseja hacer un breve resumen de las reuniones con los acuerdos alcanzados, fechas, etc. y que se envíe a todos los participantes, así como tener mucho cuidado en las decisiones que se

toman en las reuniones. Sería mucho mejor darles una pensada aparte y contestar rápido para evitar problemas y confusiones.

3. Del trabajo de campo he obtenido diferentes observaciones que creo conviene tener en cuenta cuando se conduce una entrevista individual semiestructurada aparte de otras muchas que son conocidas por todos/as:

- Las preguntas tienen que referir claramente a la matriz de evaluación, si no es así, las respuestas solo servirán para redactar recomendaciones, por ejemplo.
- Las preguntas de la matriz deben ser fáciles de comprender para que el/la entrevistado/a responda a lo que se pregunta y no divague porque no la he entendido. Cuanto más concretas, mejor.
- Las entrevistas largas dejan preguntas bien contestadas al principio y no tan bien al final. Es importante elegir bien el orden de las preguntas. El tiempo de duración de la entrevista es una variable importante a tener en cuenta. El número máximo de preguntas estaría en torno a 20 para que no se alargue y la persona entrevistada no se canse y pierda interés. Es importante que la persona sepa que en una hora más o menos se acaba. Una hora es una entrevista larga.
- Cuando en los TdR la parte contratante propone una matriz con 20 preguntas y dobles preguntas en muchas de ellas, esto es un problema pues acaba convirtiéndose en muchas más y eso hace que las entrevistas tengan que ser muchas o muy largas para recoger información.  
En ocasiones, hay que realizar muchas entrevistas a realizar y una matriz grande dificulta tener respuestas de muchos actores de cada pregunta para el cruzado de información posterior. Se dispersan las respuestas y la precisión se reduce. Es preferible menos preguntas y contestadas por más gente. Un elevado número de respuestas a una pregunta asegura tener buena información. Dos o tres respuestas a una pregunta puede tener sesgo, la saturación mejora la calidad o certeza de la respuesta.
- El volcado de información de las respuestas debe ser eficaz para que no se alargue en el tiempo y que la custodia de información sea lo más escrupulosa posible. Las fichas de sistematización han de ser conocidas y seguidas por todos los involucrados en esta tarea y deben ser utilizadas de la misma manera. Cuanta menos gente participe de estas tareas, más precisas serán.
- La transcripción de las respuestas se hace fundamental para no perder detalles de las entrevistas y así poder hacer revisiones textuales cuando sea necesario. Es un dinero que merece la pena invertir, al menos en las entrevistas clave.
- El/la entrevistador/a nunca debería interpretar las preguntas y formularlas de otra manera al entrevistado/a. Interesa la respuesta a la pregunta de evaluación y que el entrevistado/a responda lo que crea. Las preguntas que mejor se responden son las cortas y claras.
- Si, por cualquier motivo, alguien nos pidiera la explicación de la pregunta, esta ha de hacerse sin emitir juicios, ideas o ejemplos. Por ejemplo, si se pregunta por coordinación cada uno entiende una cosa y eso no es explicable. Es más, interesa y mucho saber lo que el/la entrevistado/a entiende por coordinación, no lo que entiende el entrevistador/a. Esto explicará asuntos que a veces se dan por hecho y en realidad necesitan de un trabajo más profundo o un trato especial y que afectan a otros ámbitos del programa.
- Es aconsejable identificar de antemano los actores principales y que ellos sean los primeros en ser entrevistados, pues darán mucha información valiosa para tener las “coordenadas” precisas del programa.
- Si hay muchas personas a las que entrevistar porque así lo requiere la situación, sería conveniente elegir unas pocas entrevistas clave que sean las que más peso tengan en el

proceso. No todas son igual de importantes.

- La agenda de entrevistas ha de estar bien ajustada a todas las personas involucradas y, además, deben contar con los consentimientos informados, los cuestionarios, los tiempos, el enlace de conexión, la confirmación de presencia, etc. Todo tiene que estar probado de antemano.
- En caso de tener que reconducir una entrevista debido a que el entrevistado/a divaga en exceso, hay que ser educado y elegante para volver a la senda de preguntas sin molestar. Es importante que los/as entrevistados/as no consuman tiempo en contar asuntos que no son de la evaluación a no ser que claramente se vea que ahí hay información interesante.

## Referencias bibliográficas

- Tres métodos de Evaluación de programas y servicios. Juicios finales sumativos, Teoría del cambio y Evaluación orientada a los actores implicados. Juan Andrés Ligeró Lasa. 2017
- Rayuela. Juan Andrés Ligeró Lasa, Alma Porta, Nuria Muñoz Muñoz, María Bustelo Ruesta. Ed. Máster de Evaluación de Programas y Políticas Públicas, Universidad Complutense de Madrid. Madrid 2019.
- Evaluación del impacto de los proyectos de desarrollo en la pobreza (Baker, Judy; Banco Mundial, 2000).
- Simkus, J. (2022, Jan 03). *Cluster Sampling: Definition, Method and Examples*. Simply Psychology. [www.simplypsychology.org/cluster-sampling.html](http://www.simplypsychology.org/cluster-sampling.html)
- [https://www.unsiap.or.jp/elearning/el\\_material/5\\_Agri/1511\\_Sampling2\\_IDN/Materials/M2\\_S2\\_4\\_%20Cluster%20Sampling.pdf](https://www.unsiap.or.jp/elearning/el_material/5_Agri/1511_Sampling2_IDN/Materials/M2_S2_4_%20Cluster%20Sampling.pdf)
- <https://www.formpl.us/blog/cluster-sampling>
- Diseño de evaluaciones de impacto: Perspectivas diversas Robert Chambers, Dean Karlan, Martin Ravallion and Patricia Rogers Julio 2009
- Introducción a los métodos mixtos de la evaluación de impacto. Michael Bamberger. 2012
- Potencia estadística y cálculo del tamaño del efecto en G\*Power: complementos a las pruebas de significación estadística y su aplicación en psicología. Manuel Cárdenas Castro<sup>1</sup>; Héctor Arancibia Martini<sup>2</sup> (Universidad De Valparaíso, Valparaíso, Chile) 2014.
- Manual de G\*Power: [http://www.psych.uni-duesseldorf.de/aap/projects/gpower/how\\_to\\_use\\_gpower.html](http://www.psych.uni-duesseldorf.de/aap/projects/gpower/how_to_use_gpower.html)
- Tamaño De Efecto Y Potencia Estadística. Carlos Camacho. Universidad de Sevilla.
- Métodos de muestreo. Jacinto Rodríguez Osuna. Centro de Investigaciones Sociológicas.
- Evaluando Impactos: Midiendo el Impacto. Taller: Evaluación de Impacto en Salud Pública Managua, Nicaragua 12-14 febrero 2009. Grupo Banco Mundial.
- Evaluación de impacto del Programa de Mejoramiento de Caminos Vecinales en Paraguay. Informe de línea de referencia Ana Cuesta Alejandra Caldo Martín Sosa Daniel Martínez Ruth Aquino. Noviembre de 2019. Banco Interamericano de Desarrollo.
- Efectos del Programa de Agua Potable y Saneamiento para Pequeñas Localidades y Comunidades Rurales en Bolivia Resultados de la evaluación de impacto intermedia. Autores: Pablo Celhay, José Luis Márquez, Sebastián Martínez, Cristina Mecerreyes Matías Muñoz, German Sturzenegger, Cecilia Vidal. Banco Interamericano de Desarrollo.
- Evaluación de impacto del programa “Línea directa de Innovación CDTI” (LIC). Juan Andrés Ligeró, Óscar Franco y Alma Porta. 2021
- Evaluación de impacto del programa de desarrollo infantil temprano “Crecer Bien para

- Vivir Bien” en Bolivia – Modalidad centros infantiles (BO-L1064) Julia Johannsen, Sebastian Martinez, Cecilia Vidal, Anastasiya Yarygina.
- Estudio de evaluación del impacto del régimen de ayudas a proyectos de I+D del CDTI. Novadays 2020.
  - Diferentes aproximaciones para hacer una evaluación sensible al género y al enfoque basado en derechos humanos para el desarrollo. Juan Andrés Ligerio Lasa, Julia Espinosa Fajardo, Carmen Mormeneo Cortés y María Bustelo Ruesta. AECID 2014.
  - Cuadernos Técnicos. junio 2014 Auditores. Nº 21 Conceptos Básicos de Muestreo.
  - Una guía para las evaluaciones de impacto en agua y saneamiento (Poulos et al.; Banco Mundial, 2006) (en inglés).
  - Development research in practice. The dime analytics data handbook. Kristoffer Bjärkefur, Luíza Cardoso de Andrade, Benjamín Daniels, Maria Ruth Jones. Grupo Banco Mundial.
  - Designing and Evaluating Behaviour Change Interventions. The Scottish Government.2015 ISBN: 978-1-78544-166-0 (web only)
  - <https://ers.fpg.unc.edu/environment-rating-scales>
  - La triangulación aplicada a la gestión: diagnóstico, monitoreo y evaluación de políticas públicas. Adriana Fassio, María Gabriela Rutty. Universidad de Buenos Aires (Argentina)
  - <https://thecompassforsbc.org/sbcc-tools/social-and-behavior-change-capacity-assessment-tool-individuals-facilitators-guide-and>
  - <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6062857/>
  - Directrices y Referencias del CAD. Estándares de Calidad para la Evaluación del Desarrollo. OCDE 2010.
  - Uso de los resultados de la evaluación de Impacto. David Bonbright, Keystone Accountability 2012
  - <https://odpc.ucsf.edu/clinical/behavioral-diagnosis-assessment/evaluation-of-a-behavior-change>
  - <https://www.3ieimpact.org>
  - [https://dimewiki.worldbank.org/Minimum\\_Detectable\\_Effect](https://dimewiki.worldbank.org/Minimum_Detectable_Effect)
  - [https://www.slideshare.net/MarcosVera1/sample-size-calculations-for-impact-evaluations?from\\_action=save](https://www.slideshare.net/MarcosVera1/sample-size-calculations-for-impact-evaluations?from_action=save)
  - <http://www.clinks.org/sites/default/files/UsingOffShelfToolstoMeasureChange.pdf>
  - <http://www.evaluationsupportscotland.org.uk/resources/tools/>
  - <http://inspiringimpact.org/resources/>
  - <https://www.thinknpc.org/starting-to-measure-your-impact/>
  - <https://evaluationsupportscotland.org.uk/evaluation/>
  - <https://beesystrategy.com/watch-out-for-these-16-mistakes-with-measuring-behavior-and-how-to-fix-them/>
  - [https://scielo.isciii.es/scielo.php?script=sci\\_arttext&pid=S0212-16112010000900009](https://scielo.isciii.es/scielo.php?script=sci_arttext&pid=S0212-16112010000900009)
  - <https://www.iadb.org/es/topics-effectiveness-improving-lives/evaluacion-de-impacto>

## ANEXOS

### Anexo 1. Matriz de evaluación

Ver archivo Excel adjunto

### Anexo 2. Instrumentos:

## Guía entrevista individual

Nombre del entrevistador: \_\_\_\_\_

### Para el entrevistador

Recuerda que la ficha de la entrevista tiene que ser llenada al iniciar cada una de las entrevistas que realices. Antes de iniciar es importante que leas y pidas que te firmen el consentimiento para la realización de esta. Es importante subrayar que la información se manejará de manera confidencial.

**ANTES DE LLEGAR AL LUGAR DE LA ENTREVISTA LLENA ESTA PÁGINA**

|                                                         |  |       |  |      |  |
|---------------------------------------------------------|--|-------|--|------|--|
| Fecha de la entrevista                                  |  | Lugar |  | Hora |  |
| Personas presentes en el momento de hacer la entrevista |  |       |  |      |  |

### INFORMACIÓN SOBRE LA PERSONA ENTREVISTADA:

|                               |  |             |  |
|-------------------------------|--|-------------|--|
| Nombre                        |  | Edad y Sexo |  |
| Máximo grado de estudios      |  |             |  |
| Cargo o actividad que realiza |  |             |  |
| Observaciones.                |  |             |  |

### Presentación:

1. ¿Nos puede hablar sobre usted y a qué se dedica?
2. ¿Conoce el programa de sensibilización para la mejora de prácticas de higiene y alimentación?

### Estructura:

3. ¿Considera que es importante un programa así? ¿por qué cree que puede serlo?

4. Desde su punto de vista, ¿cree que el programa tiene el apoyo y cuenta con los medios suficientes?
5. ¿Los acuerdos permiten trabajar todos los aspectos necesarios?
6. ¿Cree que existen barreras culturales que se opongan a los objetivos del programa?
7. ¿El lugar donde se realizan las actividades y el material empleado es el adecuado?
8. ¿El personal que implementa las actividades está capacitado para llevarlas a cabo?

#### **Implementación:**

9. ¿Diría usted que las actividades desarrolladas recogen los aspectos más importantes para una higiene y alimentación adecuadas? ¿Considera que falta o sobra alguna? ¿Cómo por ejemplo...?
10. ¿El lenguaje utilizado para realizar las actividades es el adecuado para las actividades?
11. ¿Las actividades creadas se ajustan bien a los mensajes que se quieren mandar al alumnado? ¿Cree que el mensaje llega a los niños/as?
12. ¿El periodo de formación de los implementadores para realizar las actividades es el adecuado?
13. ¿El número de actividades planificadas cubre toda la temática? ¿Son suficientes para mandar el mensaje y calar?
14. ¿Considera que el método del teatro fórum es el adecuado para este tipo de actuaciones o hay alguna mejor de hacerlo? ¿podría darnos un ejemplo?
15. ¿Cree que el alumnado se siente reflejado en la actividad teatralizada y comprenden los mensajes?
16. ¿De qué forma el alumnado aprecia las actividades del teatro fórum?
17. ¿Los implementadores consiguen la participación del alumnado? ¿Participan muchos activamente?
18. ¿Las tareas que los profesores realizan durante el curso refuerzan los conocimientos adquiridos en los teatro-fórum?
19. ¿Los profesores siguen la guía de actividades?
20. ¿Los carteles están presentes en el colegio y recogen la información adecuada? ¿Cree que ayuda esta información? ¿cómo podría mejorarse? ¿podría darnos un ejemplo?

#### **Resultados:**

21. ¿Cree que al final del curso el alumnado está mejor informado sobre los buenos hábitos y prácticas?
22. ¿Piensa que los alumnos han adquirido nuevos y mejores hábitos alimenticios y de higiene? ¿Cree que el alumnado cuenta sus nuevos conocimientos en casa a sus familiares?
23. ¿Se han reducido las enfermedades digestivas en casa? ¿sabe a qué es debido?
24. ¿Los familiares del alumnado del programa han adoptado mejores hábitos de alimentación e higiene en casa, en el colegio?
25. ¿Se han producido algunos efectos adversos no planificados a causa del programa? ¿Podría darnos un ejemplo?
26. ¿El programa ha alcanzado los objetivos que tenía previstos?
27. ¿Qué factores cree usted que explican los resultados?
28. ¿Considera que los efectos del programa se mantendrán en el tiempo?

# Guía entrevista de grupo

Nombre de los entrevistadores: \_\_\_\_\_

## Para el entrevistador

Recuerda que la ficha de la entrevista tiene que ser llenada al iniciar cada una de las entrevistas que realices. Antes de iniciar es importante que leas y pidas que te firmen el consentimiento para la realización de la misma. Es importante subrayar que la información se manejará de manera confidencial.

**ANTES DE LLEGAR AL LUGAR DE LA ENTREVISTA LLENAD ESTA PÁGINA**

|                                                         |  |       |  |      |  |
|---------------------------------------------------------|--|-------|--|------|--|
| Fecha de la entrevista                                  |  | Lugar |  | Hora |  |
| Personas presentes en el momento de hacer la entrevista |  |       |  |      |  |

## INFORMACIÓN SOBRE LA PERSONA ENTREVISTADA:

|                                                        |
|--------------------------------------------------------|
| Integrantes del grupo (edad, sexo, ocupación y origen) |
|                                                        |
|                                                        |
|                                                        |
|                                                        |
|                                                        |
|                                                        |
|                                                        |
|                                                        |
|                                                        |
|                                                        |
|                                                        |

Para comenzar les vamos a pedir que se presenten diciendo quiénes son, a qué se dedican y de dónde son.

1. ¿Creen ustedes que el programa de sensibilización para la mejora de buenas prácticas de higiene y alimentación es importante para combatir la malnutrición?
2. ¿Creen que el programa Buenas Prácticas de Higiene y Alimentación (BPHA) es capaz de mejorar la situación actual? ¿Creen que mejora a nivel de casa?
3. ¿Consideran que le falta algo al programa BPHA?
4. ¿Creen que se puede hacer de otra manera mejor? Pongan un ejemplo
5. ¿Qué efectos positivos puede tener un programa como el BPHA en la población?  
¿Qué puede cambiar?
6. ¿Cuál es la valoración que tienen del programa BPHA? ¿Cuáles son los aspectos más positivos el programa?

### **Anexo 3. Encuesta de línea de base alumnado**

Ver archivo Excel adjunto.

### **Anexo 4. Encuesta de pareamiento**

Ver archivo Excel adjunto.

### **Anexo 5. Encuesta monitoreo**

Ver archivo Excel adjunto.

### **Anexo 6. Encuesta de padres**

Ver archivo Excel adjunto.

### **Anexo 7. Presupuesto**

Ver archivo Pdf adjunto.