

Evaluación de Principio a Fin

Curso Evaluación de Impacto - Fundación ARU / ASIES

Profesor: Ryan Cooper

CIENTIFIKA

rcooper@cientifika.com

Twitter: @cientifika1

www.cientifika.com

Evaluación de Principio a Fin

Pasos:

1. Entender el **contexto**
2. Tener una **pregunta**
 1. ¿Qué pregunta busca contestar la Evaluación?
 2. Identificar problema y proponer solución
 3. Identificar a la población objetivo
3. **Diseñar** la Evaluación:
 1. Marco Muestral y cálculos de poder estadístico
 2. Instrumentos de Medición

Evaluación de Principio a Fin

Pasos:

4. Implementar y Evaluar **Procesos:**

1. Monitoreo de Implementación
2. Documentación de Implementación

5. **Levantar Datos:**

1. Línea base
2. Línea de salida

6. **Medir Resultados:**

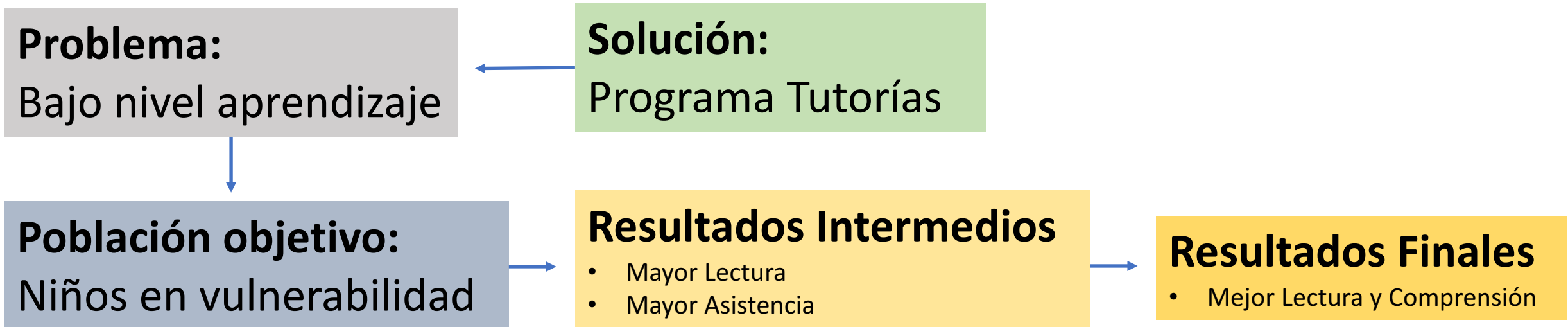
1. Impacto muestra completa
2. Impacto por subgrupos

Evaluación de Principio a Fin

Pasos:

Tener una **pregunta**

1. ¿Qué pregunta busca contestar la Evaluación y resultados esperados?
2. Identificar problema y proponer solución
3. Identificar a la población objetivo



Evaluación de Principio a Fin

Pasos:

1. Entender el **contexto (Servicio País en Educación)**

1. Población Vulnerable
2. Educación Pública
3. Alianza Gobierno y ONG



1. Entender el Contexto

- ¿Cómo mejorar la educación en contextos vulnerables?
 - Es una de las preguntas de política pública más relevantes.
 - Grandes esfuerzos en gasto público en Latinoamérica, con escasas mejoras en resultados educacionales.
- Dos puntos de vista:
 1. Sólo las intervenciones tempranas sirven.
 2. Intervenciones tardías pueden funcionar.
 - Algo de evidencia hay a favor de ambas perspectivas.
 - ¿Puede ser que ambos enfoques sean correctos?



1. Entender el Contexto

- Esta evaluación se inscribe en la segunda perspectiva.
- De hecho, el Servicio País en Educación (SPE):
 - Programa enfocado en 4° grado (~10 años).
 - Duración de 3 meses /15 sesiones.
 - Enfoque en mejorar niveles de lectura.
 - Durante el horario de clases (sustituye horas en el colegio)

2. La Pregunta

- Es clave:

1. Pensar qué se quiere contestar con la evaluación.
2. Preguntarse sobre la relevancia de la pregunta.
3. Asegurar que la pregunta no haya sido contestada.

- Esta evaluación se plantea 3 preguntas:

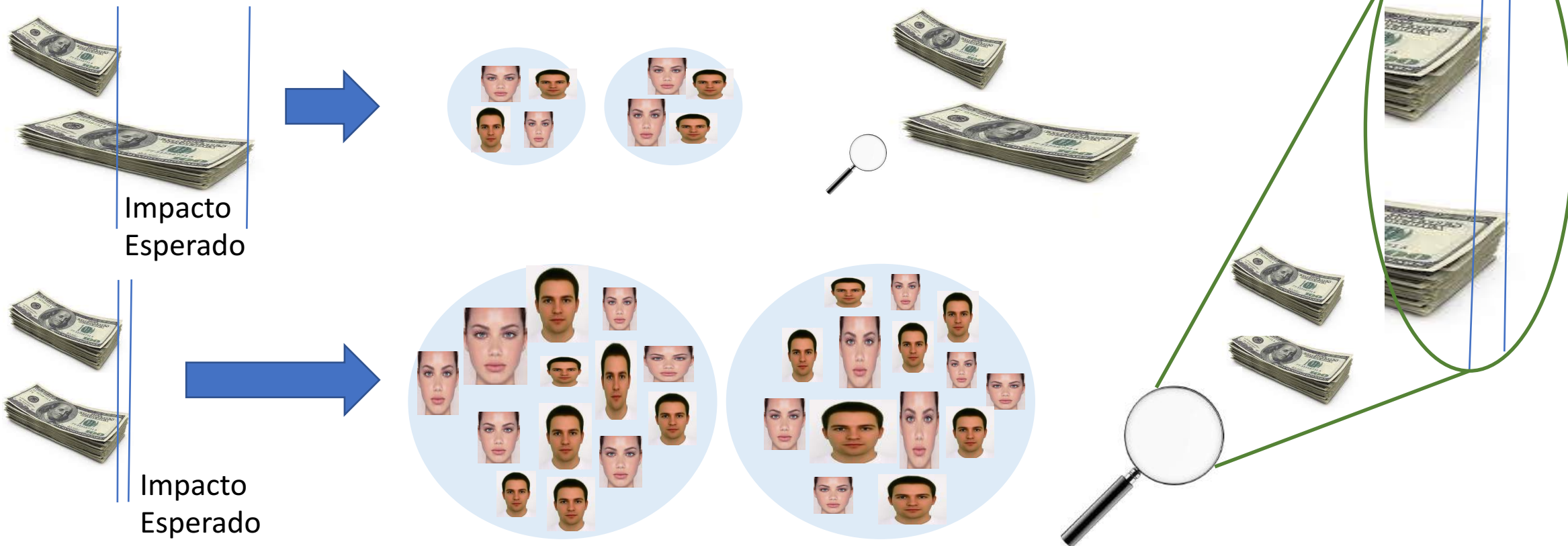
1. ¿Puede el SPE mejorar las habilidades cognitivas y no cognitivas relacionadas con la lectura?
2. ¿Existe alguna relación entre estas habilidades?
3. ¿Qué tan costo-efectivo es el programa?

Evaluación de Principio a Fin

Pasos:

3. Diseñar la Evaluación:

1. Marco Muestral y cálculos de poder estadístico



Evaluación de Principio a Fin

Pasos:

3. Diseñar la Evaluación:

1. Marco Muestral y cálculos de poder estadístico
2. Instrumentos de Medición

Resultados Producción

- Tiempo tutorías
- Cumplimiento con protocolos programa
- Asistencia niños a tutorías



Resultados Intermedios

- Lectura
- Asistencia
- Satisfacción Educación
- Satisfacción Lectura



Resultados Finales

- Velocidad Lectura
- Comprensión Lectura
- Rendimientos Matemática
- Rendimiento Ciencias

3. El Programa

- Identificar el problema y proponer una solución.
 - Visitas en terreno.
 - Focus groups.
 - Discusiones con profesores, estudiantes, investigadores.
- Identificar a la población relevante.
 - Trabajo de Campo
- Hacer pilotos.

- Este trabajo ya había sido hecho al momento de decidir evaluar el programa...

3. El Programa

- Programa implementado por la Fundación para la Superación de la Pobreza (FSP).
- 15 sesiones de tutorías en tiempo de clase.
- Los tutores son voluntarios universitarios.
- Grupos de ~6 estudiantes por tutor.
- Enfocado en escuelas vulnerables.
- Objetivos:
 1. Mejorar actitudes frente a la lectura.
 2. Mejorar niveles de lectura.
 3. Impactar en los voluntarios.
 - Esta evaluación sólo contesta las dos primeras...

Evaluación de Principio a Fin

Pasos:

4. Implementar y Evaluar **Procesos:**

1. Monitoreo de Implementación
2. Documentación de Implementación



Resultados Producción

- Tiempo tutorías
- Cumplimiento con protocolos programa
- Asistencia niños a tutorías

= Diseño – Implementación

Implementar y Evaluar Procesos

- Es muy importante:
 - Armar un buen equipo.
 - Documentar.
 - Obtener aprobación del Comité de Ética.
 - Manuales de Campo exhaustivos.
 - Entrenamientos exhaustivos.
 - Monitorear la Intervención.

Implementar y Evaluar Procesos

- Armar un buen equipo...
- ¿Qué se necesita?
 - Investigadores Principales:
 - Es clave que se involucren en todo el proceso.
 - Asistente de Investigación:
 - Una mezcla de conocimiento técnico y de terreno.
 - Coordinador de Terreno:
 - Encargado de organizar levantamiento de datos y monitoreo del programa.
 - Administra equipos que trabajan en terreno.

Implementar y Evaluar Procesos

- Consideraciones Éticas:
- ¿Por qué es importante?
 - Ha habido casos en que otros estudios han causado graves perjuicios a las personas involucradas.
 - Un Comité de Ética experto puede detectar potenciales daños.
 - Se debe pedir autorización a las personas para acceder y utilizar su información personal.

Implementar y Evaluar Procesos

- El monitoreo del SPE se basó en visitas aleatorias a las escuelas.
 - Observación de Tutorías.
 - Cuestionario a Tutores.
 - Cuestionario a Profesores.
- Además, se recopiló información administrativa de la FSP, de asistencia de alumnos y tutores.

Implementar y Evaluar Procesos

- Resultados del Monitoreo:

Panel A: Some Statistics				
Statistic	Students' Attendance	Sessions per Student	Sessions per School	Tutors per Student
Mean	0.76	9.13	12.04	2.62
Median	0.85	10.00	12.00	2.00
Standard Deviation	0.30	3.73	1.35	1.77
Minimum	0	0	9	0
Maximum	1	15	15	9

- Hubo mucha heterogeneidad en la implementación del programa!
 - Queda claro que fue importante revisar esta información.

Implementar y Evaluar Procesos

- ¿Hay diferencias regionales?

Panel B: Breakdown by Region				
Region	Students' Attendance	Sessions per Student	Sessions per School	Tutors per Student
RM	0.78	9.41	12.13	2.96
VIII	0.73	8.63	11.84	1.94
Diff	0.05	0.78	0.29	1.02**

- Sí, y estadísticamente relevantes.
- En la VIII región se implementó “mejor” que en la RM.

Implementar y Evaluar Procesos

- ¿Comparten los alumnos este enfoque?

	Like the Program?	Improved Reading?	Improved Writing?	Like Tutors?	Tutors' Skills?
Sessions per Student	0.086*** (0.026)	0.054*** (0.018)	0.021* (0.012)	0.079*** (0.019)	0.048** (0.021)
Tutors per Student	-0.053 (0.041)	-0.068** (0.033)	-0.047 (0.033)	-0.066 (0.056)	-0.047 (0.043)

- Sí, evalúan mejor el programa cuando:
 - Reciben más sesiones de tutorías.
 - Reciben menos rotación de tutores.

4. Diseño de la Evaluación

- Lo central, ¿qué metodología de evaluación utilizar?
 - Evaluar aleatoriamente es lo más confiable, pero..
 - No siempre se puede aleatorizar.
 - Es difícil evaluar políticas universales de esta forma.
 - Si el programa es muy pequeño en escala, puede ser mejor esperar a que crezca.
 - Si la implementación del programa ya comenzó, es muy tarde para aleatorizar.
- En este caso, se daban las condiciones necesarias, y se llevó a cabo una evaluación aleatoria.

4. Diseño de la Evaluación

- Luego, ¿qué grupos incluir en la evaluación?
 - Depende de la pregunta.
- En este caso, ¿Cuál es el impacto del SPE?
 - Comparamos niños con SPE vs niños sin SPE.



Tratamiento
Con SPE

Control
Sin SPE

4. Diseño de la Intervención

- Luego, hay que terminar de diseñar la evaluación:
 - Marco Muestral.
 - Unidad de Aleatorización.
 - Estratificación.
 - Cálculo de Poder.

4.1. Marco Muestral

- Se construye una lista de escuelas elegibles con base en los criterios de la FSP:
 - Escuelas de bajo desempeño académico.
 - 10 municipios en 2 regiones (RM y VIII).
 - Dependencia de la escuela en algunos casos.
 - Escuelas de 100+ alumnos en 4° grado o parejas de escuelas cercanas que sumen 100+ alumnos.

4.1. Marco Muestral

- Sólo se consideran en la evaluación escuelas que cumplan con todos estos criterios.
- ¿Qué pasa si los colegios rechazan el programa?
 - Plan B: Se genera lista de reemplazo aleatoria, previendo no aceptación del programa/evaluación.
- Se contactan escuelas con el apoyo del Ministerio de Educación.
 - ¿Cómo asegurar el interés del Grupo Control?
 - Se le ofrece un informe de resultados y recomendaciones.

4.2. Unidad de Aleatorización

- ¿Cómo decidir a qué unidad aleatorizar?
 - El programa se implementa en 4° grado.
 - Es difícil en la práctica excluir alumnos del programa en un colegio en que esté se está implementando (Riesgo de Contaminación).
 - Restricciones operativas de la FSP.
- Por ello, aparece como una opción idónea el aleatorizar a nivel de escuela.

4.3. Estratificación

- ¿Mediante qué variables estratificar?
 - Restriciones Operativas de la SPE: Estratificar por comuna en que están las escuelas.
 - Además:
 - Estratificar por Nivel Socioeconómico de las escuelas.
 - Estratificar por Desempeño educativo de las escuelas.

4.4. Aleatorización

- La aleatorización construyó la siguiente muestra:

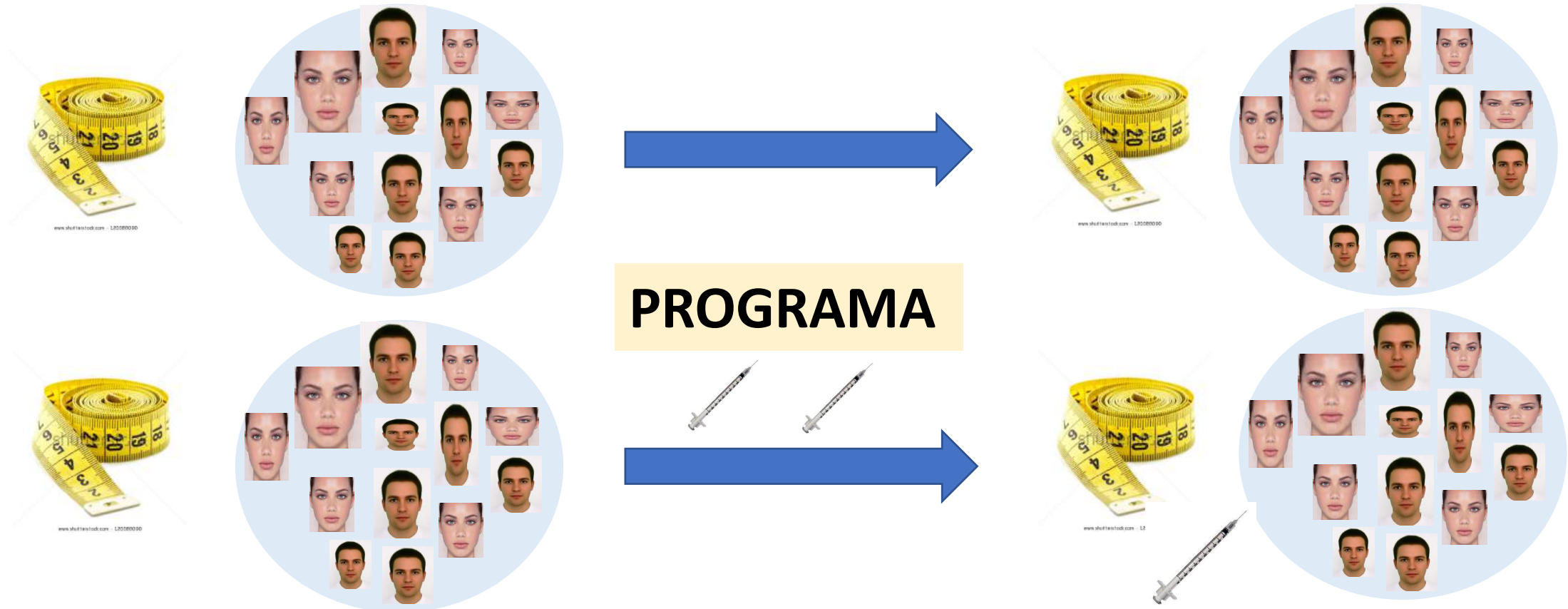
Region	County	Treatment	Control
RM	Santiago	9	5
	Estación Central	7	7
	Lo Espejo	2	1
	Maipú	5	7
	La Florida	4	5
	San Bernardo	2	4
	Total RM		29
VIII	Concepción	5	4
	Coronel	2	1
	Hualpén	4	3
	Talcahuano	5	3
	Total VIII		16

Evaluación de Principio a Fin

Pasos:

5. Levantar Datos:

- Línea base y Línea de salida



Instrumentos de Medición

- ¿Cómo medir las variables de interés?
 - Se debe hacer una revisión de los instrumentos disponibles.
 - Usamos cuatro fuentes.

1. Comprensión de Lectura y Producción de Textos (CL-PT):

- Usamos una prueba estandarizada que mide:
 - Comprensión de Lectura (RC)
 - Uso de la Lengua (UL)
 - Producción de Textos (TP).
- A éstas las llamamos “habilidades cognitivas”.

Instrumentos de Medición

2. Gusto por la Lectura (GPL):

- Construimos un instrumento nuevo con base en otros:
 - Autopercepción como Lector (SPR).
 - Disfrute de la Lectura (ER).
 - Interés por la Lectura (IR).
 - Percepción de la Lectura en la Escuela (PRS).
- A éstas las llamamos “habilidades no cognitivas”.

3. Información de Procesos del programa.

4. Información administrativa del Ministerio de Educación.

Balance entre Grupos

- Es relevante revisar estadísticamente si los grupos de la evaluación están balanceados.
- ¿En qué dimensiones?
 - En todas las variables con las que se cuente en la línea de base.
- ¿Entre qué grupos?
 - Puede ser relevante mostrar que los “desertores” son iguales entre grupos.
 - En este caso la deserción fue considerable.
 - Para otras subpoblaciones relevantes en el estudio.

i. Balance en Características de la Escuela:

Variable	Mean T	Mean C	Difference
Mothers' Years of Schooling	10.65 (1.34)	10.35 (1.48)	0.30 (0.31)
Household Income	261615.80 (80009.94)	248089.60 (70989.89)	13526.20 (16476.93)
Language SIMCE 2009	257.31 (19.24)	251.15 (24.00)	6.16 (4.82)
Math SIMCE 2009	249.98 (19.98)	243.33 (25.68)	6.65 (5.10)
IVE 2010	60.69 (11.24)	63.97 (12.50)	-3.28 (2.60)
School Size	71.70 (35.44)	74.30 (35.05)	-2.60 (7.70)
School Dependence	0.39 (0.49)	0.28 (0.45)	0.11 (0.10)

ii. Balance en Características de los Estudiantes:

Panel A: All students in the Sample				
Variable	N	Mean T	Mean C	Difference
All Students	6,066			
Attendance 2009		91.56 (6.63)	91.38 (7.68)	0.19 (0.66)
Grades GPA 2009		5.79 (0.57)	5.76 (0.59)	0.03 (0.04)
Grades Arts 2009		6.19 (0.6)	6.12 (0.65)	0.07 (0.05)
Grades English 2009		5.62 (0.93)	5.33 (1.28)	0.29 (0.18)
Grades History 2009		5.46 (0.76)	5.41 (0.81)	0.05 (0.05)
Grades Language 2009		5.31 (0.78)	5.31 (0.79)	0.01 (0.05)
Grades Math 2009		5.27 (0.85)	5.26 (0.85)	0.01 (0.05)
Grades PE 2009		6.43 (0.5)	6.48 (0.55)	-0.05 (0.05)
Grade TE 2009		6.14 (0.68)	6.04 (0.73)	0.10 (0.06)

iii. Balance entre Desertores y No Desertores

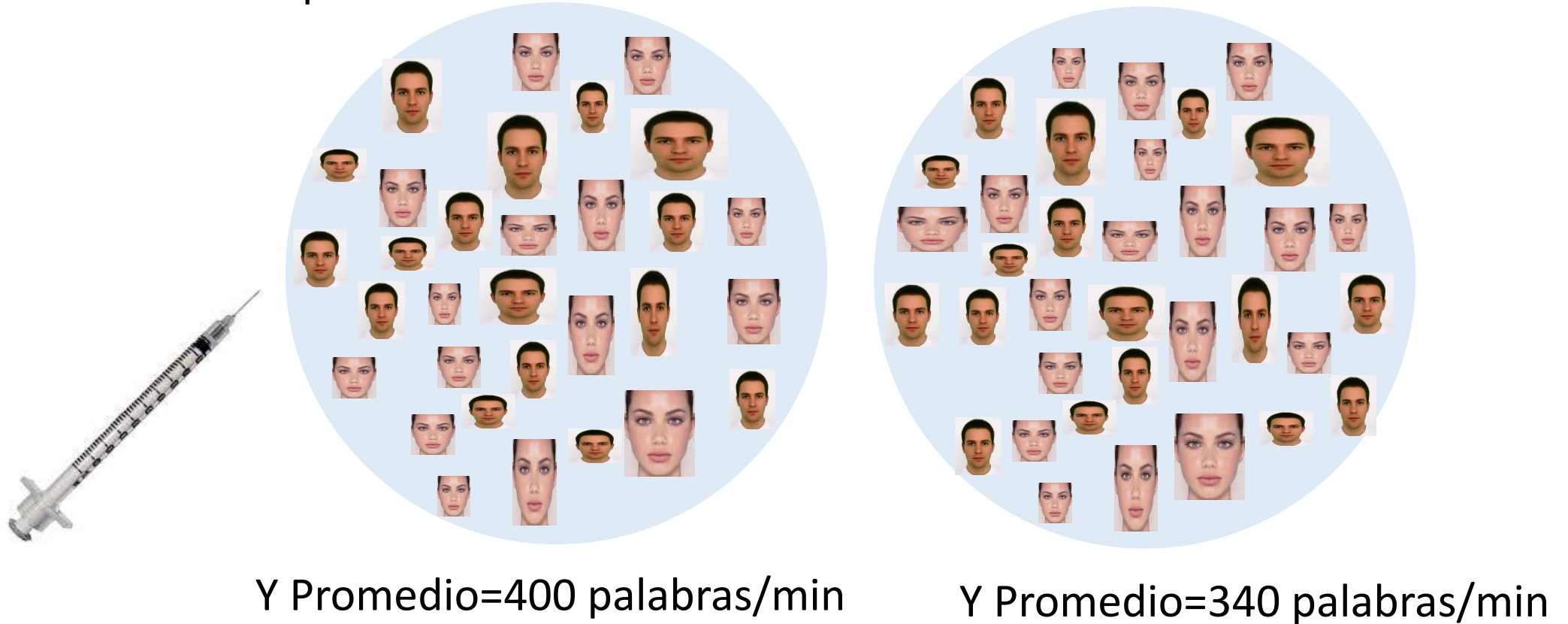
Panel B: Non Attriters and Attriters									
Variable	N	Mean T	Mean C	Difference	Variable	N	Mean T	Mean C	Difference
<u>Non Attriters</u> 4858					<u>Attriters</u> 1208				
Attendance 2009		91.73 (6.49)	91.64 (7.35)	0.08 (0.65)	Attendance 2009		89.77 (7.9)	89.21 (9.61)	0.56 (1.05)
Grades GPA 2009		5.80 (0.57)	5.77 (0.58)	0.02 (0.04)	Grades GPA 2009		5.72 (0.55)	5.67 (0.62)	0.06 (0.06)
Grades Arts 2009		6.20 (0.6)	6.14 (0.64)	0.06 (0.05)	Grades Arts 2009		6.07 (0.66)	5.98 (0.75)	0.09 (0.08)
Grades English 2009		5.63 (0.93)	5.31 (1.29)	0.32 (0.19)	Grades English 2009		5.35 (0.73)	5.31 (0.86)	0.04 (0.08)
Grades History 2009		5.47 (0.77)	5.42 (0.8)	0.05 (0.05)	Grades History 2009		5.55 (0.96)	5.46 (1.17)	0.09 (0.18)
Grades Language 2009		5.32 (0.79)	5.32 (0.79)	0.00 (0.05)	Grades Language 2009		5.25 (0.76)	5.21 (0.83)	0.04 (0.09)
Grades Math 2009		5.27 (0.86)	5.28 (0.85)	-0.01 (0.05)	Grades Math 2009		5.26 (0.83)	5.14 (0.87)	0.11 (0.08)
Grades PE 2009		6.44 (0.51)	6.49 (0.55)	-0.05 (0.05)	Grades PE 2009		6.40 (0.45)	6.42 (0.56)	-0.01 (0.07)
Grade TE 2009		6.15 (0.68)	6.05 (0.73)	0.09 (0.06)	Grade TE 2009		6.09 (0.68)	5.99 (0.75)	0.10 (0.09)

Evaluación de Principio a Fin

Pasos:

6. Medir Resultados:

1. Impacto muestra completa



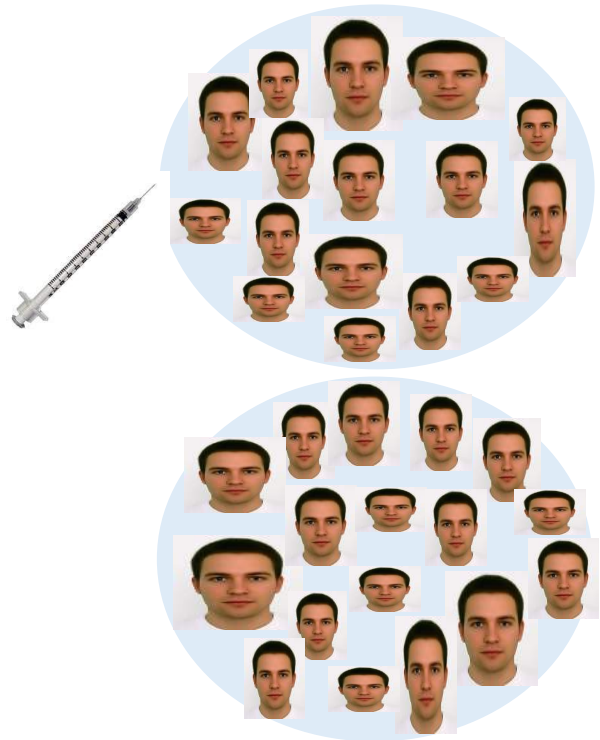
IMPACTO PROMEDIO=400-340=60

Evaluación de Principio a Fin

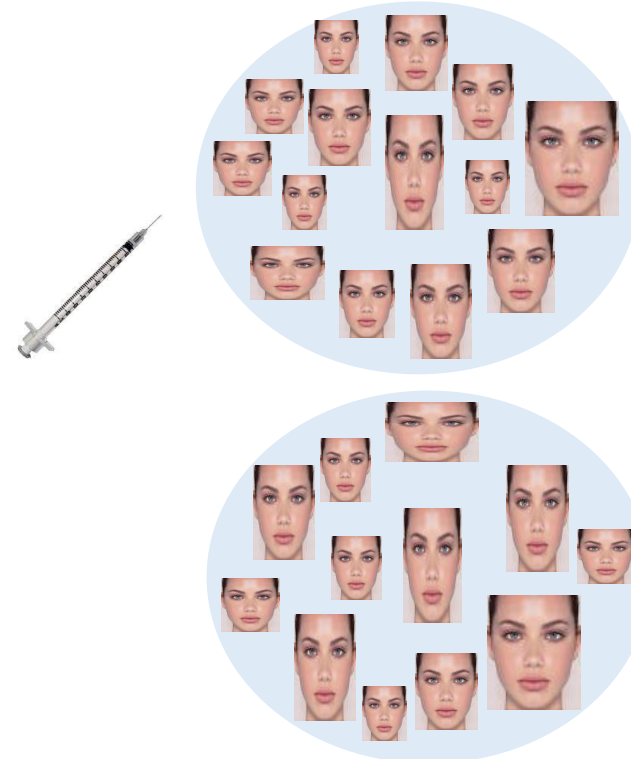
Pasos:

6. Medir Resultados:

1. Impacto muestra completa
2. Impacto x sub muestra



$$\text{IMPACTO} = 400 - 350 = 50$$



$$\text{IMPACTO} = 400 - 330 = 70$$

Evaluación de Principio a Fin

Ejercicio en Grupo: Hipotetizar Evaluación Principio a Fin

1. Entender el **contexto**
2. Tener una **pregunta**
3. **Diseñar** la Evaluación
4. Implementar y Evaluar Procesos
5. Levantar Datos
6. Medir Resultados

Evaluación de Principio a Fin

Curso Evaluación de Impacto - Fundación ARU / ASIES

Profesor: Ryan Cooper

CIENTIFIKA

rcooper@cientifika.com

Twitter: @cientifika1

www.cientifika.com