

Innovaciones en la Ciencia de la Evaluación para Informar la Política Pública: La Evaluación Experimental de Un Buen Comienzo

Hirokazu Yoshikawa, M. Clara Barata, Marycatherine Arbour,
Harvard Graduate School of Education and Harvard Medical School

María José Ramirez, Ernesto Treviño & Soledad Concha,
Universidad Diego Portales

Andrea Rolla, Catherine E. Snow
Harvard Graduate School of Education

Lorenzo Moreno & Julieta Lugo-Gil, Mathematica Policy Research, Inc.

María Graciela Garrido, Fundación Educación Oportunidad
Traducción: Maria Ramos Olazagasti, Diana Leyva

Universidad Diego Portales
Santiago de Chile
October 24, 2008
E-mail: hiro_yoshikawa@harvard.edu

Otros Colaboradores

- Steve Reifenberg, Director of Harvard's David Rockefeller Center for Latin American Studies (DRCLAS) regional office, Santiago, Chile
- Fundación Educacional Oportunidad: Elizabeth Farrelly
- Harvard Graduate School of Education and Medical School: Catherine Ayoub, Judith Palfrey, Barbara Pan, Alice Lorch, Carol Da Silva, Jennifer DiBara, Diana Leyva, Carolina Buitrago, Richard Murnane, John Willett.
- Implementation and evaluation teams in Chile: Marcela Marzolo, Bernardo Martorell, Bernardita Mayo, Carolina Melo, Mercedes Rivadeneira, Carolina Corthorn, Gabriela Barra, María Helena Salas, Paulina Fernández, María José Prat-Corona, Natalia Velásquez, Oriana Berrios, María Soledad Pinochet
- The World Bank: Emiliana Vegas
- Mathematica Policy Research: Kimberly Boller, Margaret Caspe
- DRCLAS: Christopher Barron and Maximiliano Mauriz.

Financiadores

- Evaluación experimental del impacto (2008-2011): Fundación Educacional Oportunidad, Universidad Diego Portales, Harvard Graduate School of Education, Howard Hiatt Global Health Equity Residency in Internal Medicine, Brigham and Women's Hospital, Harvard University
- Evaluación piloto y de proceso (2007-2008): Fundación Educacional Oportunidad, UNICEF Chile office, the World Bank, ACEV Foundation, Center on the Developing Child en Harvard University, Harvard Graduate School of Education, Howard Hiatt Global Health Equity Residency in Internal Medicine, Brigham and Women's Hospital, Harvard University

Visión general

- Evaluación de impactos experimentales para orientar las políticas públicas: Una historia muy breve
- Preguntas guías para la evaluación del impacto de UBC
- Teoría de cambio – por qué es importante una evaluación del impacto?
- Teoría de cambio de UBC
- Innovaciones en la evaluación del impacto de Un Buen Comienzo:
 - Asignación aleatoria de grupos
 - Ausencia de condición “sin tratamiento”
 - Innovaciones en instrumentos de medición
 - Mezcla de métodos cuantitativos y cualitativos
- Inferencias causales para guiar la política pública en la educación infantil temprana en Latinoamérica y alrededor del mundo

Inferencia Causal y Política Pública

- ¿Por qué es importante para la política pública la causalidad?
 - Inversiones en salud y bienestar público pueden ser erróneas si se basan en programas que en realidad no causan resultados positivos

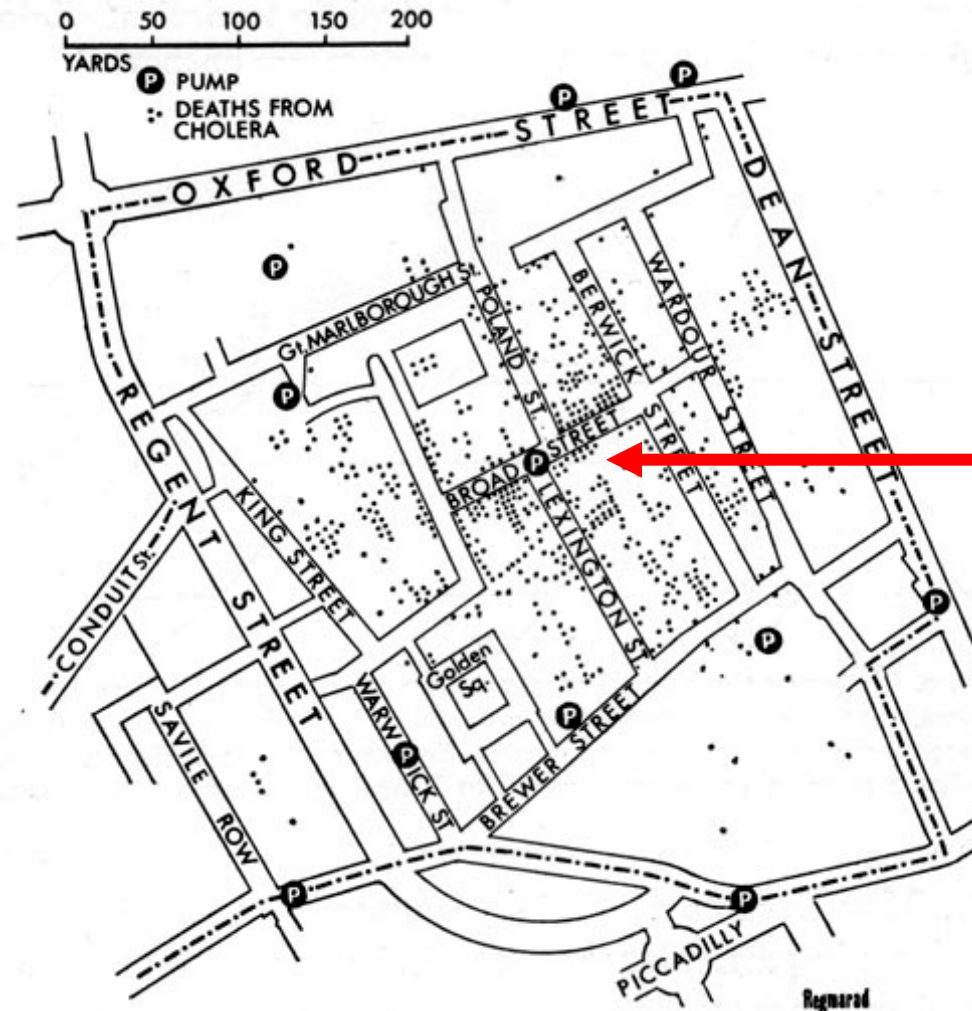
Causalidad, Ciencia y Política Pública: Una Historia Muy Breve

(Michalopoulos, 2005; Pearl, 2000)

- Antes de 1920: Correlación = Causalidad
 - Avance en los años 1890: Correlación parcial (OLS) para descartar hipótesis alternas particulares
 - Pero sólo se pueden descartar hipótesis alternas que se han medido. Qué pasa con las hipótesis que no se han medido?
 - Excepción : Cuando ocurren experimentos naturales (por ejemplo, rol causal de la bomba de agua en Broad Street en la epidemia de cólera en Londres, 1854)

1854: Dr. John Snow y un “Experimento Natural” - revelado a través de cartografía de casos de cólera

Datos espaciales (los puntos representan muertes por cólera, las “Ps” son las bombas de agua) demostraron que la epidemia de cólera en Londres fue causada por agua contaminada de una bomba de agua en la calle Broad, y no por aire contaminado (teoría miasma). Las personas que vivían cerca de otra bomba tenían una probabilidad mucho más alta de sobrevivir.



Dr. Snow le quitó la manija de la bomba de agua



Este fue
el fin de la
epidemia

Pero es difícil encontrar experimentos naturales; no se puede contestar muchas preguntas de política pública

Ronald Fisher (1925) y la primera evaluación experimental

- El estadístico Ronald Fisher fue el primero en utilizar asignación aleatoria para los efectos de evaluación.
- La pregunta era: ¿Cómo separar el efecto de los fertilizantes de los otros miles de factores que pueden influenciar el rendimiento de los cultivos?
- Fisher asignó aleatoriamente los campos: unos recibieron fertilizantes y otros no. Después midió el rendimiento de los cultivos
- El diseño descartó tanto las hipótesis alternas observables como las inobservables. Su hipótesis: que su intervención causa cambios en los resultados de interés

El uso de la evaluación experimental para informar a la política pública

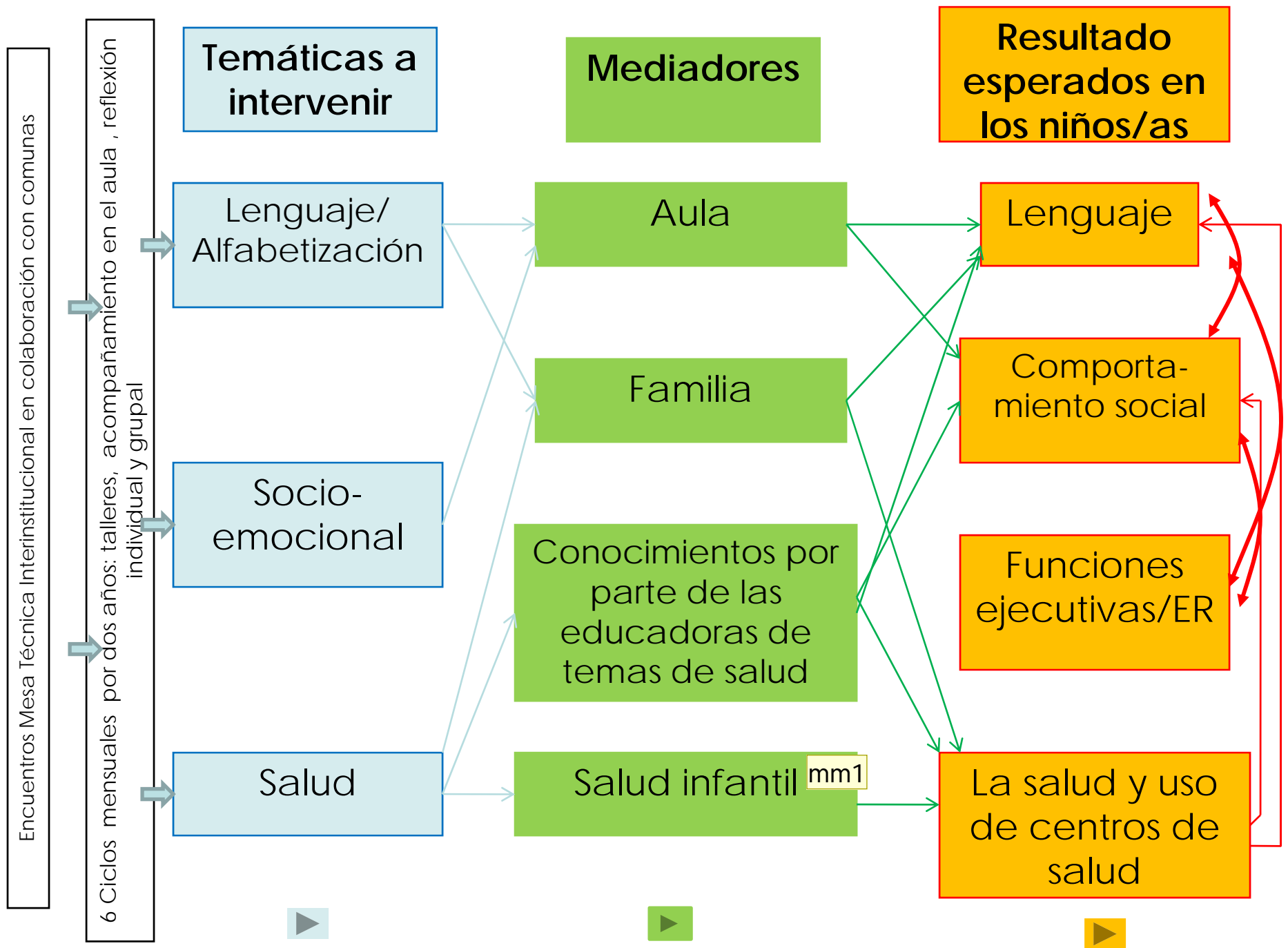
- Años 1940 – 2000: Se expandieron las aplicaciones a programas sociales y a la política pública
 - Primero en la medicina; luego en la política de impuestos, empleo, programas de asistencia pública, educación infantil temprana; más recientemente, educación primaria y secundaria
- Experimento social famoso en Latinoamérica: El programa Progreso / Oportunidades para mejorar la asistencia escolar, salud y nutrición (Levy, 2006)
- Consenso reciente: Diseño experimental sigue siendo el diseño de investigación más poderoso en la ciencia de la evaluación (Bloom, 2005; Cook, 2005)

Preguntas guías para la evaluación del impacto de *Un Buen Comienzo*

- ¿Cómo podemos mejorar la educación parvularia de tal modo que maximicemos sus beneficios para la próxima generación (especialmente generaciones con bajos recursos) de ciudadanos chileños/as?
- ¿Cuáles son los efectos causales de UBC en:
 - ¿Los niños(as)?
 - ¿Las maestras y las interacciones en el salón de clase?
- ¿Cómo se producen estos efectos? ¿Cuáles son los mecanismos a través de los cuales UBC ejerce su influencia?
- ¿Varían los efectos de UBC en los subgrupos particulares? Grupos definidos por:
 - Características del niño(a) y de la familia (por ejemplo, género; los niveles previos de riesgo de familia o niño/a)
- ¿Cuáles son las vivencias de los niños, los padres y los maestros que participaron en UBC?

La Teoría de Cambio de *Un Buen Comienzo*

- Teoría de cambio: Un modelo conceptual acerca de cómo funciona un programa (Weiss, 1995)
- No es sólo una lista de resultados de interés
- Indica los *mecanismos* mediante los cuales esos resultados van a ser afectados
- Puede ayudar a hacer explícitas las suposiciones y la cadena de efectos causales en las que se basa un programa
- Es una guía para determinar qué medir y cómo y cuándo medirlo



Diapositiva 13

mm1

Esto más que un nivel de intervención, me parece que corresponde a una temática: no creo que debiera ir aquí

mmarzolo, 17-10-2008

Temáticas a intervenir

```
graph TD; A[Temáticas a intervenir] --> B[Lenguaje / Alfabetización]; A --> C[Desarrollo Socioemocional]; A --> D[Salud];
```

Lenguaje / Alfabetización:-

- Evaluación formativa
- Lenguaje en la organización del tiempo diario
- Lectura y escritura emergente
- Extensión del lenguaje oral
- Involucramiento de la familia
- Integración curricular

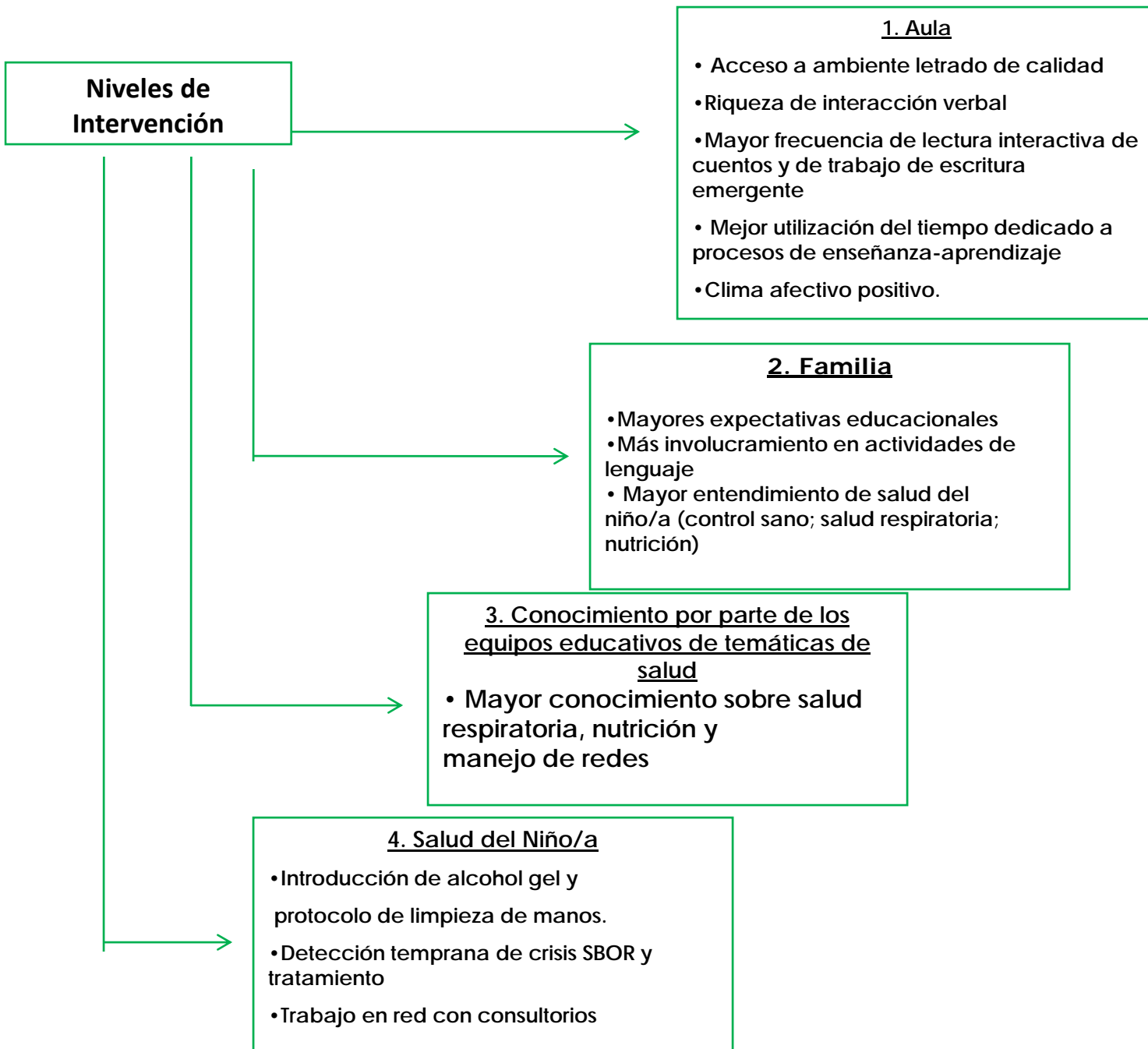
Desarrollo Socioemocional

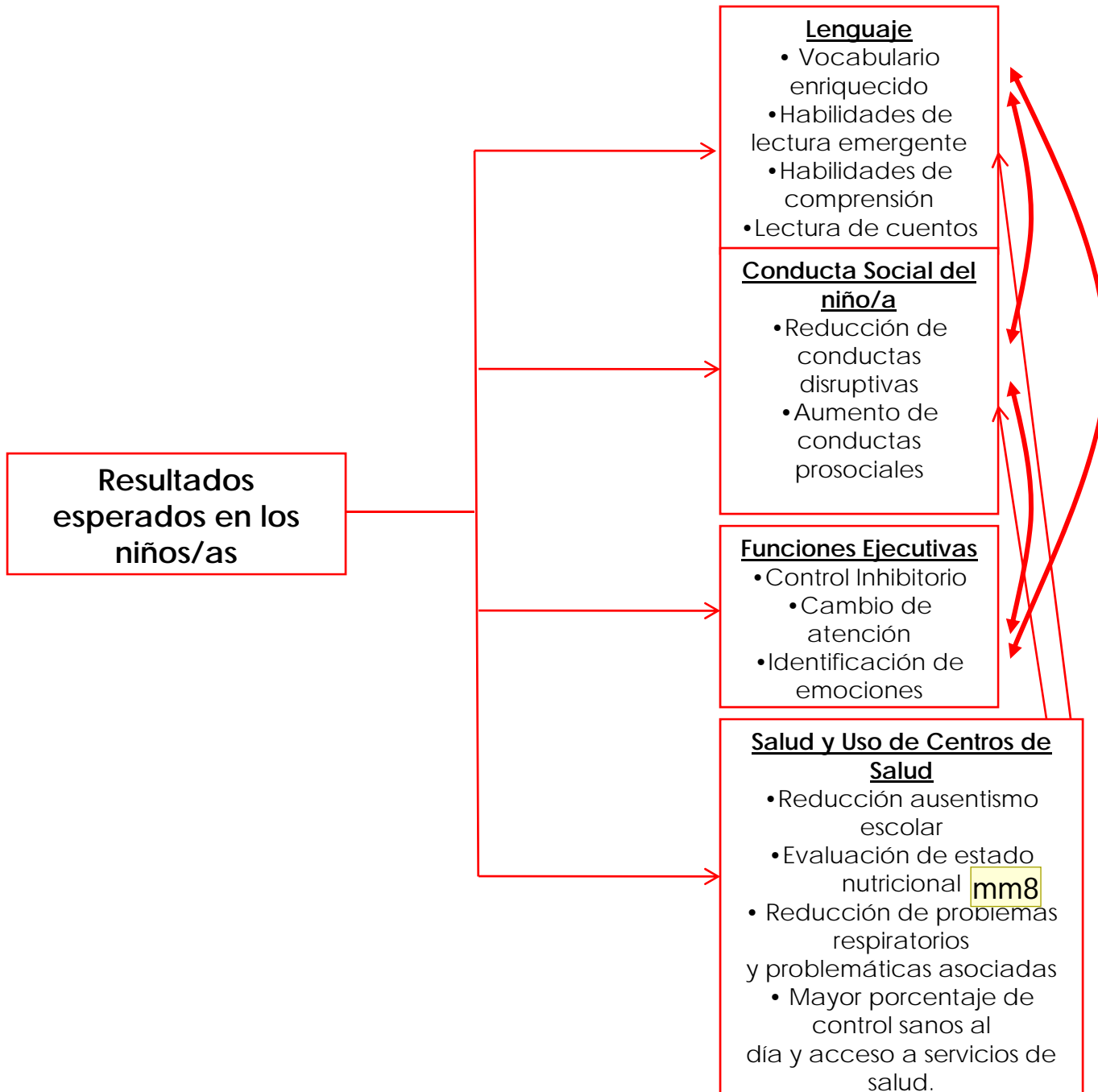
- Clima afectivo positivo
- Manejo de casos individuales en conjunto con la familia

Salud

- Prevención de Síndrome Bronquial Obstructivo Recurrente (SBOR) y Coordinación entre la familia y la escuela
- Conexión con redes de salud
- Nutrición sana: prevención de sobrepeso y obesidad







Diapositiva 16

mm8

Esto tampoco hasta donde yo sé...

mmarzolo, 17-10-2008

Innovaciones:

1) Asignación Aleatoria de Grupos

- Intervención diseñada para implementarse a nivel de la escuela preescolar para poder contestar la pregunta de cómo mejorar la calidad de la educación parvularia
- No se puede asignar a los niños a pre-escolar versus no pre-escolar
- No se puede asignar aulas dentro de la escuela a UBC vs. no UBC (las maestras se comunican entre ellas; los efectos pueden extenderse de una condición a la otra)
- Diseño de asignación aleatoria de grupos (Bloom, 2005; Raudenbush, Martínez, & Spybrook, 2007; Schochet, 2005) en el que se asignan pre-escolares enteros a las distintas condiciones y se hace seguimiento en el tiempo
- Nuestra evaluación es, hasta donde tengo entendido, la primera asignación aleatoria de grupos fuera de los EEUU en educación temprana.

Innovaciones: 2) No hay una Condición sin Tratamiento

- Condición Intervención 1: Modelo UBC completo
- Condición Intervención 2: Talleres y provisión de materiales literarios
- Ayuda a abordar las preocupaciones éticas

La Comunicación del Diseño Experimental a los Grupos de Interés

- Nos involucramos en un proceso de colaboración con comunas con el fin de comunicar los objetivos y el contenido del programa
- Comunicamos la imparcialidad de utilizar un sorteo para asignar recursos limitados y determinar cuáles pre-escolares estarían asignados a las 2 condiciones:
Intervención 1 e Intervención 2

- ¿Pero cómo funciona la asignación aleatoria?
- ¿Cómo se hace?

El Sorteo: 2008 Marzo



Un esfuerzo colaborativo entre los líderes de la comuna y las escuelas



Innovaciones: 3) Medición

- Estudio inicial en 2007 demostró que las medidas de los niños(as) eran confiables y válidas, las cuales incluyen medidas chilenas e internacionales (Yoshikawa et al., 2008 informe para UNICEF Chile)
- Medición del proceso y recursos educacional en aula
 - Proceso de interacción de maestro y estudiante
 - Recursos en el salón de clase y acceso a materiales
- Medidas de una serie de resultados de salud y utilización de servicios de salud
- Medidas precias de asistencia (más difícil de lo que parece!)
- Medidas de calidad de implementación de UBC

Innovaciones: 4) Mezcla de métodos cualitativos y cuantitativos

- Tendencia reciente en las evaluaciones de impactos experimentales: Incluir estudios cualitativos para examinar cómo se vivencia el programa:
 - Los que lo implementan
 - Los que participan de él
- ¿Por qué?
 - Ayuda a explicar efectos inesperados en la evaluación cuantitativa
 - Ayuda a explicar diferencias en impactos para diferentes subgrupos
 - Ayuda a explicar la variación en implementación del programa
- Es particularmente poderoso si la investigación cualitativa es longitudinal y puede seguir el cambio en niños(as), familias y educadoras (Duncan, Huston, & Weisner, 2007; Yoshikawa, Weisner, & Lowe, 2006)
- De esta manera, datos cualitativos pueden contribuir a nuestro entendimiento de procesos causales de un experimento (Gibson-Davis & Duncan, 2005; Yoshikawa, Weisner, Kalil, & Way, 2008)

El Poder de los Datos Cualitativos en un Diseño Experimental

- **New Hope:** Un programa de lucha contra la pobreza mediante el suministro de suplementos de ingresos, el cuidado de los niños(as) y subvenciones para la atención de la salud (Duncan, Huston, & Weisner, 2007)
- Los efectos sobre los niños(as) se evaluó en un experimento aleatorio; un estudio cualitativo de 44 familias fue incluido en la muestra mayor de 750 familias
- El análisis cuantitativo mostró que los efectos positivos en el rendimiento escolar, la reducción de la conducta agresiva es más marcada para los niños que las niñas (Bos et al., 1999)
- **¿Por qué?** Análisis cuantitativos no pudieron detectar una respuesta
- Datos etnográficos mostraron: los padres estaban más preocupados por el tiempo después de la escuela de los niños (por ejemplo, la participación en pandillas) - tenían más probabilidades de poner los niños en actividades extra-curriculares después de la escuela
- Esto puede haber jugado un rol en los efectos positivos en los niños (Duncan, Huston, & Weisner, 2007)
- *Las preocupaciones de los padres no se midieron en la encuesta, pero surgió en la entrevista más abierta de los datos cualitativos*

El Rol de los Datos Cualitativos en la Evaluación de UBC

- En 2007, proporcionó información esencial sobre cómo los maestros y directores vivenciaron UBC
- Nuestro plan es continuar con este tipo de recopilación de datos a través de todo el estudio con el fin de complementar los datos de la encuesta:
- ¿Cuál es la calidad de la vivencia de UBC? ¿Positiva / o no tan positiva la vivencia? Objetivo: mejorar continuamente el programa
- ¿Cómo se producen los efectos? ¿Existen mecanismos que no hemos medido con otros métodos que se reflejan en las descripciones de los educadoras, técnicos y las familias?
- ¿Por qué podría ser UBC una vivencia diferente para los distintos subgrupos de profesores o familias?

Conclusión: Una Evaluación Experimental Aleatoria para Influir las Políticas Públicas

- Simplemente aumentar el acceso a la educación preescolar no garantiza resultados positivos para los niños(as) (Yoshikawa, McCartney et al., 2007, informe sobre México para UNICEF).
- Mejorar la calidad de la Educación Inicial es el objetivo internacional fundamental para los millones de niños(as) pequeños que no alcanzan su potencial de desarrollo debido a la pobreza, la salud o los riesgos ambientales (Engle et al., 2007; Grantham-McGregor et al., 2007).
- Desarrollar y evaluar enfoques para la mejora de la calidad representa el siguiente reto en la atención de la primera infancia y la educación a nivel internacional.
- La evaluación de Un Buen Comienzo utiliza múltiples métodos dentro de un diseño experimental para responder a la pregunta de cómo mejorar la calidad de la educación temprana y el desarrollo de los niños(as), no sólo en Chile, pero en el resto de América Latina y América del Sur y el mundo.

Referencias

- Barata, M.C., Yoshikawa, H., Rolla, A., Marzolo, M., & Garrido, M.G. (2008). *Resultados de la Evaluación Cualitativa de Un Buen Comienzo*. Report to the UNICEF Chile office.
- Bloom, H. (2005). Randomizing groups to evaluate place-based programs. In H.S. Bloom (Ed.), *Learning more from social experiments: Evolving analytic approaches* (pp. 115-171). New York: Russell Sage Foundation.
- Contreras, D., Herrera, R. and Leyton, G. (2007). Impacto de la educación preescolar sobre el logro educacional: Evidencia para Chile. Chile, Santiago: Universidad de Chile. http://sechi.facea.uchile.cl/pdf/regulares%20-%20sesiones/Herrera_rher.pdf.
- Cook, T.D. (2005). Emergent principles for the design, implementation, and analysis of cluster-randomized experiments in social science. *Annals of the American Academy of Political and Social Sciences*, 599, 176-198.
- Cunha, F., & Heckman, J. (2007). *The technology of skill formation*. NBER Working Papers 12840. Cambridge, MA: National Bureau of Economic Research.
- Duncan, G.J., Huston, A.C., & Weisner, T.S. (2007). *New Hope for low-income working families and their children*. New York: Russell Sage Foundation.
- Engle, P.L., Black, M.M., Behrman, J.R., DeMello, M.C., Gertler, P.J., Kapiriri, L., Martorell, R., Young, M.E., & the International Child Development Steering Group (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet*, 369, 229-242.
- Fisher, R.A. (1925). *Statistical methods for research workers*. Edinburgh: Oliver & Boyd.
- Grantham-McGregor, S., Cheung, Y.B., Cueto, S., Glewwe, P., Richter, L., Strupp, B., & the International Child Development Steering Group. Developmental potential in the first 5 years for children in developing countries. *The Lancet*, 369, 60-70.
- Levy, S. (2006). *Progress against poverty: Sustaining Mexico's Progreso / Oportunidades program*. Washington, DC: Brookings Institution.
- Michalopoulos, C. (2005). Precedents and prospects for randomized experiments. In H.S. Bloom (Ed.), *Learning more from social experiments: Evolving analytic approaches* (pp. 1-36). New York: Russell Sage Foundation.
- Moreno, L., & Lugo-Gil, J. (2008). *Revised power calculations and additional ICC's*. Unpublished memo, Mathematica Policy Research, Inc., January 2008.
- National Forum on Early Childhood Program Evaluation (2007). *A science-based framework for early childhood policy*. Cambridge, MA: Center on the Developing Child, Harvard University. Available at: http://www.developingchild.harvard.edu/content/downloads/Policy_Framework.pdf.
- Raudenbush, S.W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis*, 29, 5-29.
- Rolla, A., Arias, M., Villers, R., & Snow, C. (2006). Evaluating the impact of different early literacy interventions on low-income Costa Rican kindergarteners' academic skills. *International Journal of Educational Research*, 45, 188-201.
- Snow, C.E., Burns, M.S., & Griffin, P. (1998). (Eds.). *Preventing reading difficulties in young children*. Washington, DC: National Academy Press.
- Umayahara, M. (2006). Early childhood education policies in Chile: From pre-Jomtien to post-Dakar. Background paper prepared for the Education for all *Global Monitoring Report 2007: Strong Foundations Early Childhood Care and Education*. <http://unesdoc.unesco.org/images/0014/001474/147494e.pdf>.
- UNESCO. (2006). *Education for all: Strong Foundations Early Childhood Care and Education*. EFA Global Monitoring Report 2007.
- Vegas, E., & Petrow, J. (2007). *Raising student learning in Latin America: The challenge for the 21st century*. Washington, DC: World Bank. Villalón, M., Suzuki, E., Herrera, M.O., & Mathiesen, M.E. (2002). Quality of Chilean early childhood education from an international perspective. *International Journal of Early Years Education*, 10 (1), 49-59.
- Yoshikawa, H., Barata, M.C., Rolla, A., Da Silva, C.D., Ayoub, C., Arbour, M., & Snow, C.E. (2008). *Un Buen Comienzo, una iniciativa para mejorar la educación preescolar en Chile: Resultados del primer año de implementación*. Report to the UNICEF office, Santiago, Chile.
- Yoshikawa, H., McCartney, K, Myers, R., Bub, K., Lugo-Gil, J., Knaul, F., & Ramos, M. (2007). *Preschool education in Mexico: Expansion, quality Improvement, and curricular reform*. Florence, Italy: UNICEF Innocenti Research Centre Working Paper. Available at: http://www.unicef-irc.org/publications/pdf/iwp_2007_03.pdf.
- Yoshikawa, H., Weisner, T.S., Kalil, A., & Way, N. (2008). Mixing qualitative and quantitative research methods in developmental science: Uses and methodological choices. *Developmental Psychology*, 44, 344-354. .
- Yoshikawa, H., Weisner, T.S., & Lowe, E. (2006). (Eds.). *Making it work: Low-wage employment, family life, and child development*. New York: Russell Sage Foundation.