

EVIDENCIA SOBRE USOS Y CONSECUENCIAS DE UN PROGRAMA DE EVALUACIÓN DE APRENDIZAJES

CONTRERAS, J.,
TORRES, Á.
MACHTIG, P.,
MUÑOZ, M.,
BARRIENTOS, P.,
CONTARDO, V.,
ACUÑA, V.,

CONTENIDOS

I. Introducción: ¿Por qué estudiar los usos y consecuencias de las evaluaciones estandarizadas?

II: Antecedentes teóricos y empíricos

- Concepto de validez
- Evidencia de validez de usos y consecuencias: teoría de acción del programa y estudios de usos.
- Problematizando la noción de “consecuencias”

III. Preguntas y objetivos

IV. Metodología

V. Resultados

VI. Discusión



I. INTRODUCCIÓN

1. ¿Por qué estudiar los usos y consecuencias de las evaluaciones estandarizadas?

- **Contexto del sistema educacional chileno**

Instalación de un modelo de mercado durante los años 1980 (Corvalán & García-Huidobro, 2016; Villalobos & Quaresma, 2015).

Sistema de Aseguramiento de la Calidad: Políticas de rendición de cuentas (Falabella, 2016; Parcerisa & Falabella, 2017).

I. INTRODUCCIÓN

- Evaluaciones de altas y bajas consecuencias en Chile

| Nivel Analítico | Naturaleza de las Consecuencias | |
|------------------|---------------------------------------|---|
| | Altas | Bajas |
| PROPÓSITO | Certificación y accountability | Toma de decisiones y mejora del aprendizaje |
| Individuo | PSU/ calificaciones escolares | Evaluaciones formativas internas y externas |
| Escuela | SIMCE/clasificación de escuelas | Información de monitoreo de gestión y logros de aprendizaje |

¿Por qué reunir evidencia de validez consecucional en evaluaciones de bajas consecuencias?

II. ANTECEDENTES TEÓRICOS Y EMPÍRICOS

1. Concepto de Validez

Estándares (AERA, APA y NCME, 2014)

“El grado en el cual la evidencia y la teoría apoyan las interpretaciones de los resultados del test para sus usos propuestos. Validez es, entonces, el aspecto más importante en el desarrollo y evaluación de tests”.

Por tanto :

Si desconocemos el uso, desconocemos el real valor del programa (Hein y Taut, 2010).

Evidencia para las consecuencias de la medición (AERA, APA y NCME, 2014)

- Las declaraciones del programa respecto de los usos e interpretaciones propuestos
- Los usos y consecuencias no esperados.

II. ANTECEDENTES TEÓRICOS Y EMPÍRICOS

2. Evidencia de validez de usos y consecuencias: teoría de acción del programa y estudios de usos.

Teoría de la acción: formulación que da cuenta del encadenamiento de procesos a través de los cuales el programa espera lograr los resultados propuestos (Patton, 2008; 2012), desde las declaraciones de los diseñadores y desde la literatura (Wylie, 2017). Suele estar implícita.

La validez de las interpretaciones de puntajes de los tests no solo depende de los usos sino que especialmente de las declaraciones que subyacen a la teoría de acción de esos usos (AERA, APA y NCME, 2014)

II. ANTECEDENTES TEÓRICOS Y EMPÍRICOS

3. Problematizando la noción de “consecuencias”

¿Qué debe considerarse una consecuencia? ¿Para quién?

Evaluación como

- **Producción de diferencias interindividuales** de rendimiento escolar; de éxito y fracaso; de clasificación, de mérito (Duru-Bellat, 2009; Bernstein, 1998; Perrenoud, 1996).
- Tecnología de poder en la producción de verdad científica y control poblacional: Gubernamentalidad educacional (Carrasco, 2013; Ramos, 2018).


No solo las evaluaciones de “altas consecuencias”

- contribuyen a la función de selección y distribución social (Dubet, 2011).
- Participan en el entramado que produce verdad científica sobre el aprendizaje, su mejora y la calidad de la educación (Ramos, 2018).

III. OBJETIVO Y PREGUNTAS GUÍA


Este estudio aborda el problema de los efectos intencionados y no intencionados de las mediciones educacionales, de los usos e interpretaciones de los resultados y de la utilidad de sus reportes, tanto desde el punto de vista de los diseñadores como desde la percepción de sus beneficiarios.

Obj: Determinar el grado de correspondencia entre la teoría de acción del programa SEPA y sus interpretaciones, usos y efectos en los destinatarios.

1. ¿Cuáles son las principales declaraciones de SEPA respecto de sus resultados esperados?
 2. ¿Cuáles son las declaraciones de los usuarios respecto de los usos y consecuencias de SEPA?
 3. ¿Cuál es la relación entre la teoría de acción del programa y la evidencia empírica sobre sus usos y consecuencias?
- 

IV. METODOLOGÍA: DISEÑO GENERAL

- **Caso de evaluación**
 - Prueba estandarizada de bajas consecuencias (por diseño).
 - Principalmente formativa
 - Entrega resultado a nivel de alumno, curso (aula) y administrador.
 - Estado y progreso, en 2 materias en 11 niveles.

 - **Diseño general**
 - Alcance descriptivo – relacional, mediante métodos cualitativos.
 - 2 componentes articulados mediante el análisis:
 1. Reconstrucción de la teoría de acción del programa
 - **Evidencia documental (38)** del discurso: búsqueda de documentos públicos y dirigidos a usuarios en web, plataforma de resultados y comunicaciones directas orales y escritas.
 - **Perspectiva de actores clave en la creación e implementación (16)**: Entrevistas individuales y grupales (a equipos)
 2. Recolección de información sobre los usos y consecuencias
 - **Perspectiva de los usuarios (33)**: entrevistas individuales a equipos de directivos y a administradores (tipo de contratante / dependencia / ubicación / rol)
- 

IV. METODOLOGÍA: PARTICIPANTES Y FUENTES DE INFORMACIÓN

1a. Teoría de la acción: Documental (revisión sistemática)

| Fuente/equipo | Total general (n=80) | Incluido en análisis (N=38) | Excluido (N=42) |
|--------------------------------|----------------------|-----------------------------|-----------------|
| Artículos científicos | 10 | 0 | 10 |
| Equipo de usos y resultados | 21 | 15 | 6 |
| Equipo operaciones | 6 | 4 | 2 |
| Transversal (plataforma y web) | 17 | 10 | 7 |
| Equipo Ventas | 10 | 8 | 2 |
| Búsqueda general WWW | 16 | 1 | 15 |

1b. Teoría de la acción: actores clave (colección completa)

| Tipo de actor Clave | Participantes (N=16) | Técnica de producción de datos |
|---------------------------|----------------------|--------------------------------|
| Creación y gestión | 7 | 7 Entrevistas individuales |
| Equipos de implementación | 9 | 3 entrevistas grupales |

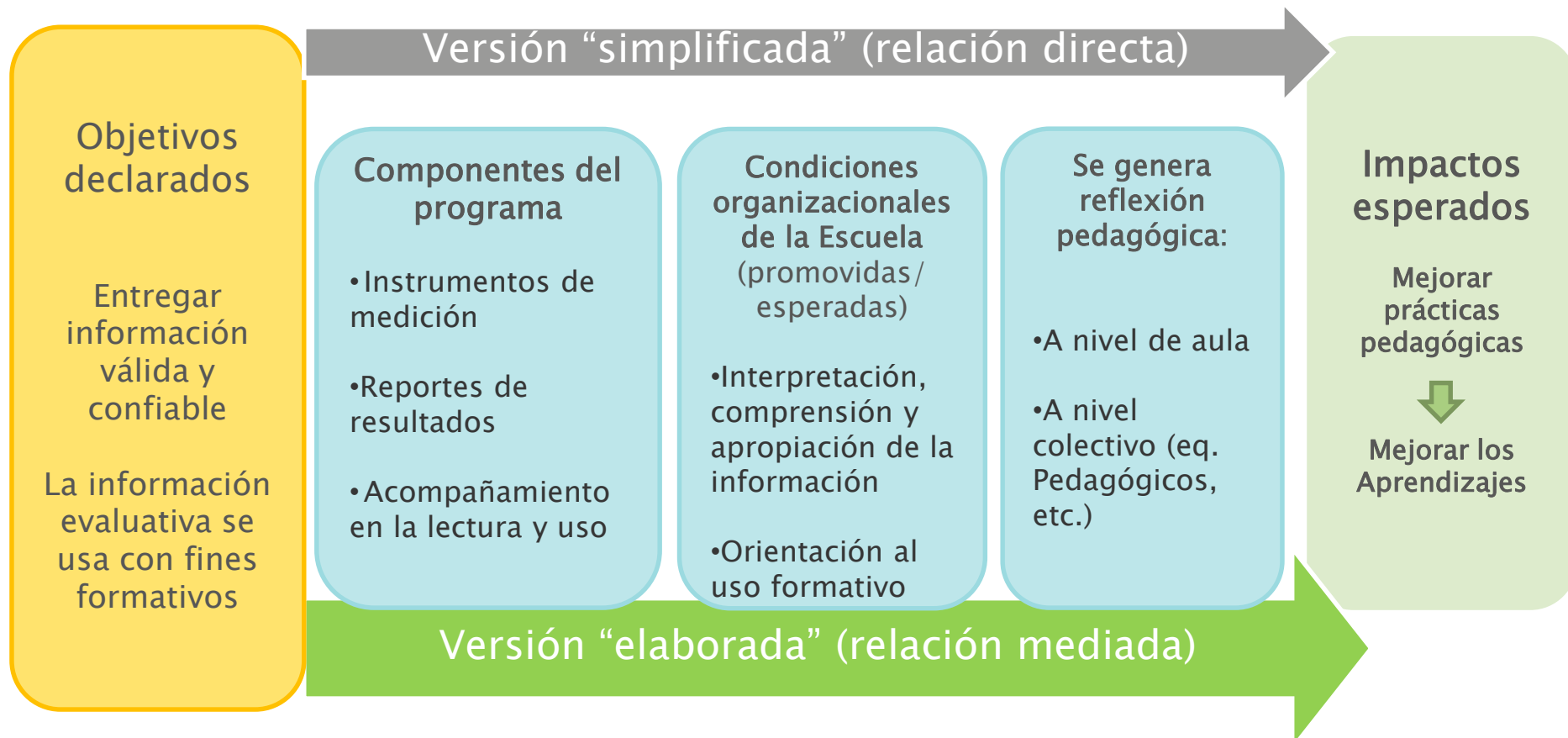
IV. METODOLOGÍA: PARTICIPANTES Y FUENTES DE INFORMACIÓN

2. Usos y consecuencias (Muestreo determinado a priori: Tipo de contratante/dependencia/ubicación/rol)

| Tipo contratante | Dependencia | Ubicación | Total (N=33) |
|---------------------------------|---------------|---------------|--------------|
| Administrador grupo de colegios | Pública | Metropolitana | 5 |
| | | Región | 5 |
| | Privada | Metropolitana | 1 |
| | | Región | 4 |
| Establecimiento individual | Pública | Metropolitana | 2 |
| | | Región | 4 |
| | Privada | Metropolitana | 2 |
| | | Región | 2 |
| | Subvencionada | Metropolitana | 4 |
| | | Región | 2 |
| Investigación | Investigación | Metropolitana | 2 |

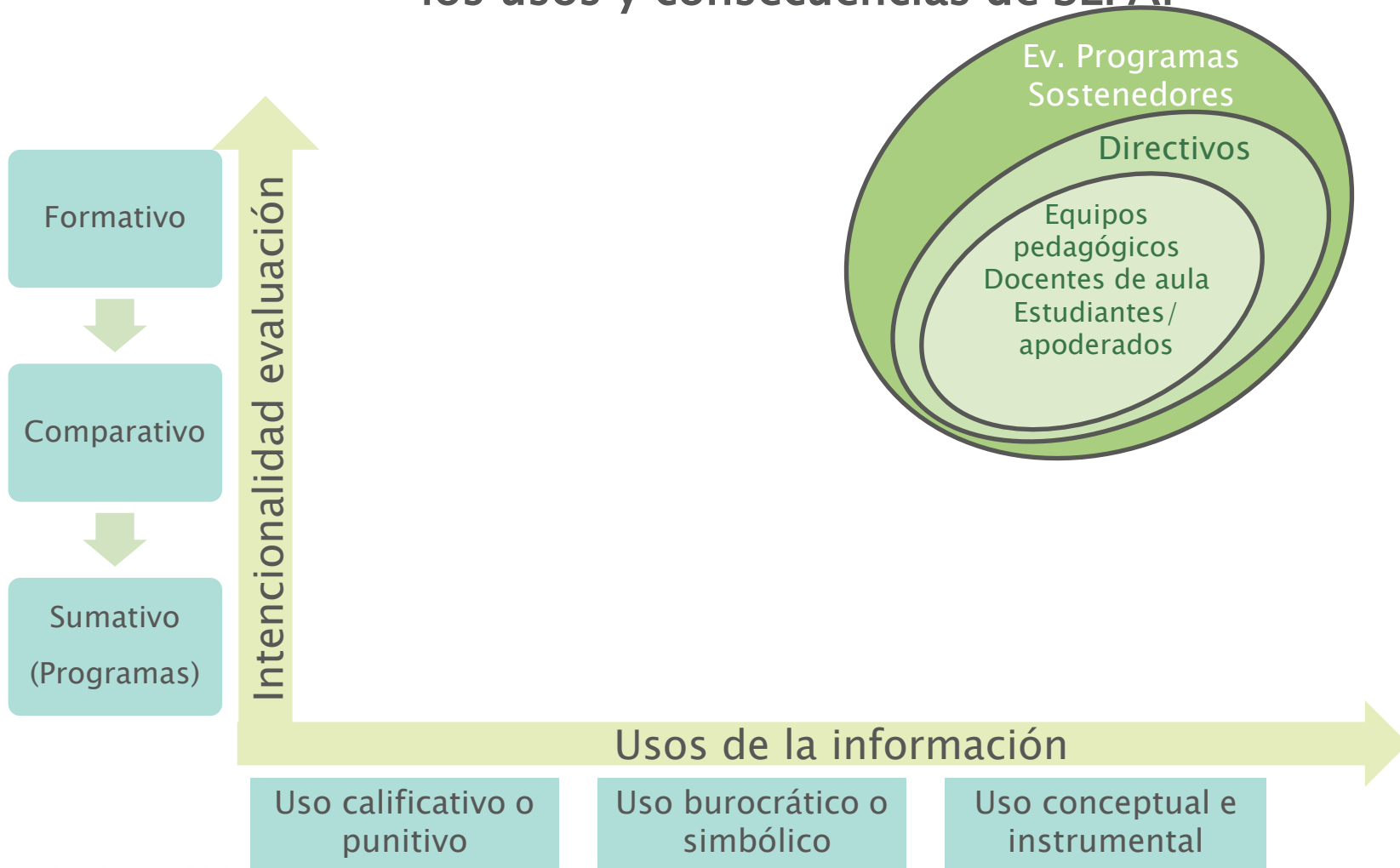
V. RESULTADOS

1. ¿Cuáles son las principales declaraciones de SEPA respecto de sus resultados esperados?



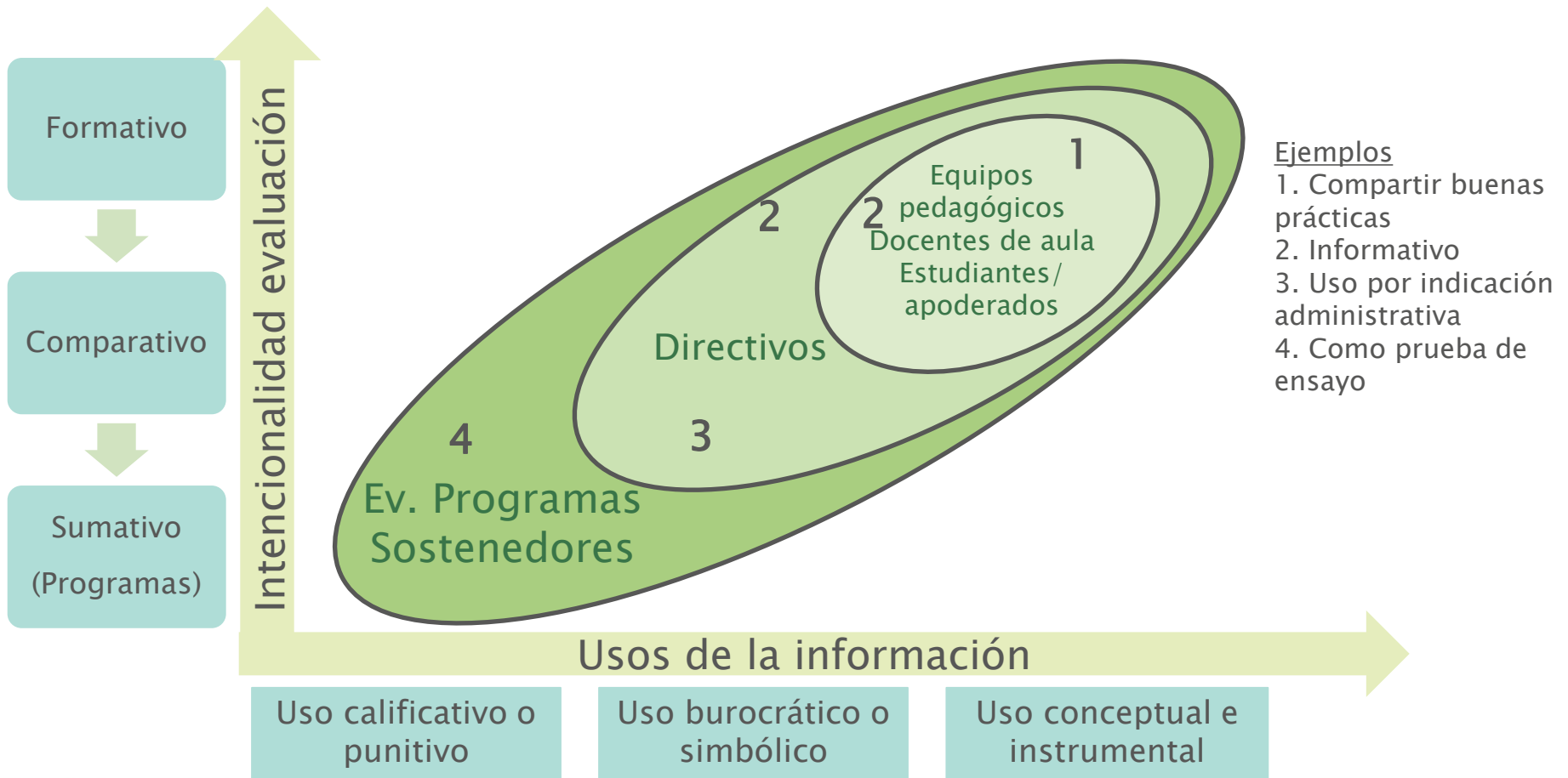
V. RESULTADOS

2. ¿Cuáles son las declaraciones de los usuarios respecto de los usos y consecuencias de SEPA?



V. RESULTADOS

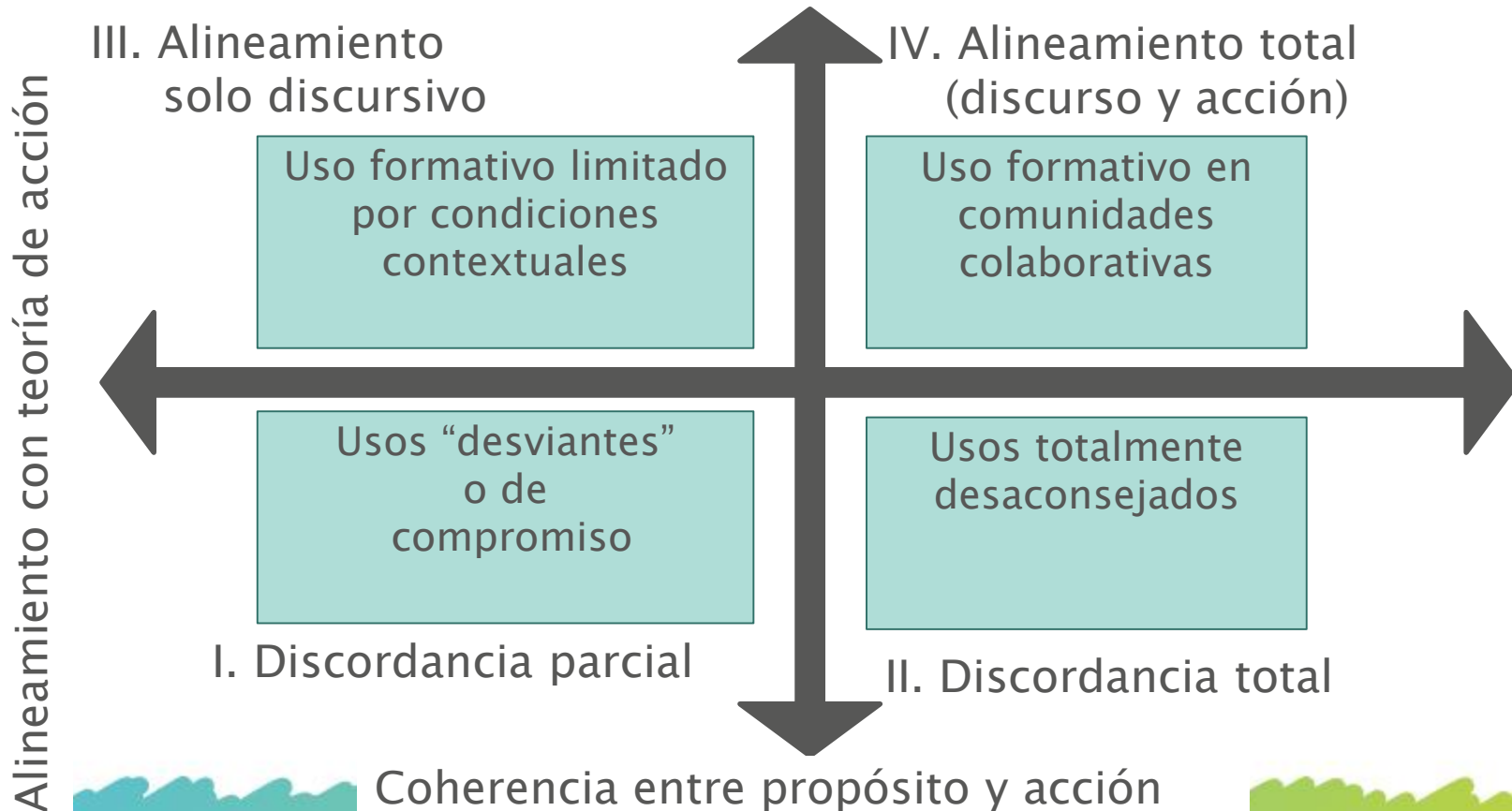
2. ¿Cuáles son las declaraciones de los usuarios respecto de los usos y consecuencias de SEPA?



V. RESULTADOS

3. ¿Cuál es la relación entre la teoría de acción del programa y la evidencia empírica sobre sus usos y consecuencias?

Creencia:
Evaluación mejora el aprendizaje



VI. DISCUSIÓN

- Limitaciones de este estudio
- ¿En qué medida los resultados sobre los usos constituyen evidencia a favor y en contra de la validez de SEPA?
 - Vacíos o ambigüedades en la teoría de acción
 - En este caso, la evidencia de validez de usos no debiese basarse tanto en la concordancia discursiva con el enfoque de la evaluación formativa, pues la adhesión existe (al menos parcialmente), sino en la declaración de acciones.
 - Emerge la pregunta de si es evidencia a favor de validez el hecho de compartir un propósito de la evaluación, a pesar de que no tenga sustento empírico.
- ¿Cuáles son las exigencias/límites de los resguardos de validez de usos de un programa de evaluación educacional de “bajas consecuencias”?

Producción de “verdad científica”

Participación en entramado de relaciones de poder que definen la calidad de la educación, el aprendizaje, las diferencias individuales, los méritos académicos

SELECCIÓN DE REFERENCIAS

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Araya, C. Taut, S., Santelices, V. & Manzi, J. (2011). Validez consecucional del programa de asignación de excelencia pedagógica en Chile. *Estudios pedagógicos (Valdivia)*, 37(2), 25–42. <https://dx.doi.org/10.4067/S0718-07052011000200001>.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5.
- Brewer, C., Knoepfel, R. C., & Lindle, J. C. (2015). Consequential validity of accountability policy: Public understanding of assessments. *Educational Policy*, 29(5), 711–745.
- Briggs, D. (2017). Learning Theory and Psychometrics: Room for Growth, Assessment in Education: Principles. *Policy & Practice*. 24(3), 351–358. Doi: 10.1080/0969594X.2017.1336987
- Darling–Hammond, L. (2007). Race, inequality and educational accountability: the irony of ‘No Child Left Behind’. *Race Ethnicity and Education*, 10(3), 245–260, DOI: 10.1080/13613320701503207.
- Dubet, F. (2011). *La experiencia sociológica*. Barcelona: Gedisa.
- Duru–Bellat, M. (2009). *Le mérite contre la justice*. París: Presses de Sciences Po.
- Falabella, A. & de la Vega, F. (2016). Políticas de responsabilización por desempeño escolar: Un debate a partir de la literatura internacional y el caso chileno. *Estudios pedagógicos (Valdivia)*, 42(2), 395–413. <https://dx.doi.org/10.4067/S0718-07052016000200023>
- Flick, U. (2004). *Introducción a la Investigación Cualitativa (2ª Edición)*. Madrid: Morata.

SELECCIÓN DE REFERENCIAS

- Hein, A. & Taut, S. (2010). El uso de información evaluativa externa con fines formativos: el caso de establecimientos educacionales chilenos participantes de SEPA. *Revista Iberoamericana de Evaluación Educativa*, 3(2), 160–181.
- Heritage, M. (2010). *Formative Assessment. Making it Happen in the Classroom*. California: Corwin.
- Kane, M. (2008). Terminology, Emphasis, and Utility in Validation. *Educational Researcher*, 37(2), 76–82.
- Kane, M. (2016). Validation Strategies: Delineating and Validating Proposed Interpretations. En S. Lane, M. Raymond & T. Haladyna (Eds.). *Handbook of Test Development*. (2nd ed), pp. 64–80. New York: Routledge.
- Manzi, J., Bogolasky, F., Gutiérrez, G., Grau, V. & Volante, P. (2014). *Análisis sobre valoraciones, comprensión y uso del SIMCE por parte de directores escolares de establecimientos subvencionados*. Santiago: FONIDE.
- Mons, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie*, 169, 99–140. Doi : 10.4000/rfp.1531
- Nkwake, A. (2015). *Credibility, Validity, and Assumptions in Program Evaluation Methodology*. Suiza: Springer.
- Patton, M. (2002). *Qualitative research and Evaluation Methods* (3th ed). California: Sage.
- Perrenoud, P. (1996). *La construcción del éxito y del fracaso escolar*. Madrid: Morata.
- Ramos, C. (2018). Dispositivo de evaluación y gubernamentalidad del sistema educacional: entretejimiento de ciencia social y poder. *Cinta moebio* 61, 41–55. Doi: [10.4067/S0717-554X2018000100041](https://doi.org/10.4067/S0717-554X2018000100041)
- Ravela, P., Picaroni, B. & Loureiro, G. (2017). *¿Cómo mejorar la evaluación en el aula?* Montevideo: Grupo Magro.
- Taut, S., Santelices, V., Araya, C., & Manzi, J. (2010). Theory underlying a national teacher evaluation program. *Evaluation and Program Planning*, 33(4), 477–486.
- Wylie, E. C. (2017). Winsight™ Assessment System: Preliminary Theory of Action. *ETS Research Report Series*.



¡MUCHAS GRACIAS!

Johana Contreras, MIDE UC

jtcontre@uc.cl

Alex Torres, MIDE UC

atorresp@uc.cl