



PONTIFICIA  
UNIVERSIDAD  
CATÓLICA  
DE CHILE

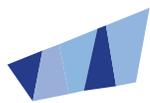
# RESUMEN EJECUTIVO

## DISEÑO DEL MODELO DE EVALUACIÓN DE LA NUEVA EDUCACIÓN PÚBLICA Y LEVANTAMIENTO DE LÍNEA BASE

CENTRO DE ESTUDIOS DE POLÍTICAS Y  
PRÁCTICAS EN EDUCACIÓN- CEPPE UC

---

DIRECCIÓN DE ESTUDIOS SOCIALES  
DESUC



**Centro UC**  
Estudios de Políticas y Prácticas  
en Educación - CEPPE

**DESUC**  
Dirección de Estudios Sociales

Este estudio fue realizado por CEPPE UC y DESUC para el Ministerio de Educación vía licitación **N° 034/2017** del Programa de las Naciones Unidas para el Desarrollo - PNUD. La contraparte técnica del estudio correspondió al Centro de Estudios MINEDUC.

**CEPPE UC** tiene como misión prioritaria realizar investigación sobre políticas y prácticas en educación, para contribuir al desarrollo del sistema educativo chileno. En particular, CEPPE UC impulsa una amplia agenda de proyectos de investigación, tanto avanzada como aplicada, que abordan problemas estratégicos de la educación chilena desde una perspectiva multidisciplinaria. Para más información: [www.ceppeuc.cl](http://www.ceppeuc.cl)

**DESUC** es una unidad de estudios y ejecución de proyectos y servicios profesionales del Instituto de Sociología de la Pontificia Universidad Católica de Chile (ISUC), integrado por un equipo interdisciplinario de profesionales, investigadores y consultores especialistas en diversas áreas y materias de opinión pública, educación, comunicaciones y estudios sociales, en aquellos proyectos que requieren un acercamiento metodológico de alto estándar y generación de modelos de investigación ad hoc. Para más información: [www.sociologia.uc.cl/desuc/](http://www.sociologia.uc.cl/desuc/)

#### EQUIPO DE INVESTIGACIÓN

##### **Jefe de Proyecto**

Alejandro Carrasco, Director CEPPE UC

##### **Investigadores Senior**

Ernesto Treviño, CEPPE UC: Evaluación de Impacto

Magdalena Browne, DESUC: Levantamiento de Datos

##### **Investigadores**

Consuelo Béjares, CEPPE UC

Angélica Bonilla, CEPPE UC

Vicky Rojas, DESUC

Cristóbal Villalobos, CEPPE UC

Ignacio Wyman, CEPPE UC

**Como citar:** CEPPE UC & DESUC (2018). Diseño del Modelo de Evaluación de la Nueva Educación Pública y Levantamiento de Línea Base. Informe Final, CEPPE UC. Santiago, Chile.

## METODOLOGÍA Y LÍNEA BASE PARA EVALUAR LA NUEVA EDUCACIÓN PÚBLICA

Los cambios que involucra la Ley 21.040 que crea el Nuevo Sistema de Educación Pública (NEP) en Chile pueden considerarse una de las reformas educacionales más importantes implementadas desde la primera mitad de los años ochenta. En términos resumidos, la NEP constituye una profunda transformación institucional, al traspasar la administración de los establecimientos educacionales desde las municipalidades al Estado, a través de la creación de 70 Servicios Locales de Educación (SLE), que en adelante serán las nuevas instituciones sostenedoras de la educación pública. Estos Servicios Locales de Educación serán coordinados por una Dirección de Educación Pública, y tendrán, además, un Consejo Directivo Local y un Consejo Local de Educación, en los que participarán distintos actores involucrados en el proceso educativo. El total presupuestado para el año 2018 de esta iniciativa es de \$427.709 millones (Dirección de Educación Pública, 2018). La implementación de esta política contempla un proceso gradual de puesta en marcha a partir del año 2018 hasta el 2025, y en casos muy excepcionales hasta 2030, que incluirá a las 346 comunas del país y a los establecimientos educacionales públicos (que en el 2017 fueron 6.807), lo que significa un importante desafío tanto para ejecutores como para todas las comunidades escolares del país. Considerando la magnitud y gradualidad del cambio, la legislación (Ley 21.040, artículo séptimo) establece la creación de un Consejo de Evaluación del Sistema de Educación Pública, nombrado por el Presidente de la República, con la finalidad de hacer seguimiento a la implementación de los SLE y ofrecer una evaluación intermedia de su desarrollo, generando el marco jurídico para la evaluación de esta política.

La envergadura de esta iniciativa pública, su importancia para el desarrollo nacional y forma de implementación en el país, requiere un estudio sistemático sobre el alcance e impacto del Nuevo Sistema de Educación Pública. La evaluación de este tipo de iniciativas constituye un elemento central en el ciclo de las políticas públicas. Evaluar las políticas o programas permite identificar si los cambios observados en el tiempo en una población determinada son o no efecto de una intervención en particular (King y Behrman, 2009; Borman et al., 2003), de modo que la toma de decisiones se base en evidencia sobre la eficacia o la capacidad que tienen determinadas políticas de generar cambios en la calidad de vida de la población (Weiss, 1998) y de la eficiencia que tiene, por tanto, en el diseño de la intervención y uso de recursos públicos.

El presente estudio puede considerarse un primer paso en esta dirección, en tanto supuso el diseño de una metodología para la evaluación de la política de Nueva Educación Pública y el levantamiento de la línea base asociada tanto a la evaluación de procesos como de impacto. Para dar cuenta de los resultados de esta investigación, el presente Resumen Ejecutivo sintetiza los principales aspectos estudiados, dando cuenta de los elementos conceptuales, metodológicos y empíricos más relevantes del estudio.

Para ello, el Resumen Ejecutivo se estructura en seis apartados, adicionales a esta introducción. El primer apartado resume la evidencia y discusión académica relevante para contextualizar el estudio en términos del diagnóstico que justifica la desmunicipalización y los principales nudos críticos sobre la futura implementación del Nuevo Sistema de Educación Pública. El segundo apartado ofrece una discusión sobre enfoques y métodos de evaluación de impacto y de procesos que enmarcan

las decisiones metodológicas del estudio. El tercer apartado se focaliza en describir la teoría del cambio construida –es decir, el encadenamiento lógico entre necesidades, acciones y resultados- para evaluar el impacto y los procesos de la Nueva Educación Pública, dando cuenta de los supuestos e hipótesis de la evaluación. El cuarto apartado sintetiza la propuesta metodológica del estudio para realizar tanto la evaluación de impacto como de procesos de la NEP, describiendo las metodologías de selección utilizadas, técnicas de recolección de información y análisis desarrollados. El quinto apartado detalla los principales resultados del estudio, describiendo aquellos relacionados a la evaluación cuantitativa y cualitativa realizada. Finalmente, el último apartado describe las principales recomendaciones y conclusiones generadas del estudio con la aspiración de contribuir a la toma de decisiones que deben efectuar diferentes actores responsables de la implementación de la NEP (Dirección Nacional NEP, MINEDUC, Consejo de Evaluación, Directores de SLE, entre otros).

## 1. CONTEXTO DEL ESTUDIO. LA EDUCACIÓN PÚBLICA Y EL NUEVO SISTEMA DE EDUCACIÓN PÚBLICA (NEP).

Los crecientes procesos de cooperación pública-privada en la provisión educativa han impulsado la discusión sobre el significado de la educación pública en un contexto global en donde su completo predominio ha tendido, aunque lentamente, a debilitarse dando lugar a esquemas híbridos de provisión educacional. En términos conceptuales, existe relativo acuerdo sobre los elementos que definen una educación pública y que la diferencian de esquemas mixtos o híbridos: i) la propiedad; ii) su control democrático; iii) la gestión y financiamiento; iv) los propósitos o intereses últimos que busca satisfacer un proyecto educativo; v) y su carácter sistémico en un territorio nacional (e.g. Lubienski y Lubienski, 2014). Según esa definición, se pueden considerar como establecimientos públicos aquellos donde estos cinco componentes son controlados y/o asegurados por el Estado o por organismos públicos, en grados diversos de autonomía respecto del nivel central.

En Chile, la educación pública ha experimentado importantes transformaciones en las últimas décadas. Desde los años ochenta, los establecimientos son administrados por los municipios, conocido como el “proceso de municipalización” (e.g. Marcel y Raczynski, 2009). Desde esa fecha, la tendencia general de la tasa de matrícula del subsistema de educación pública ha ido a la baja y ha mostrado deficiencias estructurales. Así, por ejemplo, mientras en 1981, cerca del 15% de la matrícula era privada, en 2010 este número se incrementaba por sobre el 40%. En 2015, el 38,5% de los estudiantes chilenos asistían a establecimientos públicos municipales, el 51% a particulares subvencionados y el 10,5% a colegios particulares pagados (OECD, 2016).

Asimismo, se ha dado cuenta de distintos problemas en la gestión y desarrollo de la educación municipal, especialmente en aspectos como: i) bajas capacidades profesionales y disparidad técnica de los equipos municipales en términos financieros, administrativos y de apoyo técnico-pedagógico; ii) alto nivel de atomización de los establecimientos públicos y escala inadecuada de implementación de la educación pública; iii) alta heterogeneidad de la oferta pública en términos de su calidad y complejidad; iv) débil gobernanza ante esquema de autoridad diluida entre MINEDUC y municipios, así como alta dependencia del ciclo político de alcaldes; v) mayor concentración de estudiantes en desventaja socioeconómica, generando ingresos inestables y alto costo de provisión (Gallego y Seebach, 2007; Román, 2008; Marcel y Raczynski, 2009; Bellei, 2015; Villalobos y Quaresma, 2015; Carrasco et al. 2015; CEPPE UC, 2017).

## 1.1. EL SISTEMA DE NUEVA EDUCACIÓN PÚBLICA. PRINCIPIOS Y CARACTERÍSTICAS CENTRALES

La situación resumida anteriormente ha motivado una serie de reformas, siendo el envío del Proyecto de Ley de Nueva Educación Pública (NEP) durante el año 2016 (y aprobado el 2017 en la Ley 21.040), el más relevante para este propósito (Gobierno de Chile, 2015). En este se entiende que la educación pública deberá estar organizada bajo nueve principios articuladores: i) calidad integral; ii) mejora continua de la calidad; iii) cobertura nacional y garantía de acceso; iv) desarrollo equitativo e igualdad de oportunidades; v) colaboración y trabajo en red; vi) desarrollo de proyectos educativos inclusivos, laicos y de formación ciudadana; vii) pertinencia local, diversidad de los proyectos educativos y participación de la comunidad; viii) formación ciudadana y valores republicanos y; ix) integración con el entorno y la comunidad.

Para el desarrollo de estos principios, la Ley de Nueva Educación Pública propone un rediseño organizacional e institucional del sub-sistema de educación pública. Así, en el nuevo Sistema de Educación Pública se establece la instalación de un esquema organizacional integrado por los establecimientos educacionales que actualmente dependen de las municipalidades (a través de DAEM/DEM y corporaciones municipales), los cuales serán traspasados a los Servicios Locales de Educación Pública (SLE), servicios públicos descentralizados cuyo objeto único será la provisión del servicio educacional en sus respectivos territorios de competencia. Adicionalmente, se establece la creación de una Dirección de Educación Pública que tendrá por objeto coordinar a los Servicios Locales y asegurarse que provean una educación pública de calidad a lo largo de todo el país (Gobierno de Chile, 2017).

Los Servicios Locales de Educación estarán distribuidos a lo largo de todo el territorio nacional y agruparán, en promedio, cuatro comunas colindantes, teniendo dentro de sus principales funciones: i) proveer, a través de los establecimientos educacionales de su dependencia, el servicio educacional en los niveles y modalidades que corresponda de conformidad a la ley; ii) administrar los recursos humanos, financieros y materiales del servicio y los establecimientos educacionales de su dependencia; iii) desarrollar la oferta de educación pública en el territorio que le corresponda; iv) diseñar y prestar apoyo técnico-pedagógico y a la gestión de los establecimientos educacionales de su dependencia; v) fomentar el trabajo colaborativo y en red de sus establecimientos educacionales y; vi) ejecutar acciones orientadas a fomentar la participación de los miembros de la comunidad educativa y de las comunidades locales; entre otras.

Adicionalmente, la ley establece la creación de los Consejos Locales de Educación Pública. Estos Consejos son “una instancia que tiene como objetivo estimular y canalizar la participación de la comunidad educativa en el proyecto educativo y en las demás áreas que estén dentro de las esferas de sus competencias” (Ley 21.040, art. 15). Los Consejos estarán compuestos por representantes de los estudiantes, centro de padres, madres y apoderados, asistentes de la educación, docentes, directivos y sostenedores, así como también otros organismos, instituciones y actores relevantes de la comunidad y territorio donde se encuentre el establecimiento educacional. De esta manera, el nuevo Sistema de Educación Pública fomenta la participación de la comunidad local en la estructura local, lo que constituye un aspecto crítico de los sistemas de educación pública (Cuppen, Broekhans y Enserink, 2012).

De acuerdo a la ley, todos estos componentes del Nuevo Sistema de Educación Pública se instalarán de manera gradual a partir del año 2018 hasta el 2025, con la posibilidad de retrasar la entrada de algunos municipios por solicitud expresa del Presidente de la República hasta el año 2030, para casos que cumplan requisitos muy específicos de excepción (Gobierno de Chile, 2018). Esta implementación gradual tiene especial relevancia para el diseño de la evaluación presentada enseguida.

## 1.2. EL DISEÑO DE LA NEP Y LOS DESAFÍOS QUE PLANTEA A SU EVALUACIÓN: GOBERNANZA, ESCALA Y ESCALABILIDAD

Evidentemente, la puesta en marcha de la Ley 21.040 implica una transformación global del subsistema de educación pública, especialmente en términos organizacionales y estructurales. Variada literatura internacional ha mostrado que este tipo de transformaciones tienen que hacer frente a lo menos tres nudos críticos, que deben considerarse elementos claves para cualquier evaluación de esta naturaleza.

En primer término, es necesario afrontar los cambios en la *gobernanza* del sistema. La gobernanza es la distribución de atribuciones y responsabilidades que se construye entre los niveles centrales, los niveles intermedios y los niveles locales donde se entrega el servicio educativo. Conceptualmente, es importante que los sistemas educativos puedan balancear dos principios en la constitución de su proceso de gobernanza: el principio de autonomía con el principio de equidad en la entrega del servicio (Hanushek, Link y Woessmann, 2013; McKinsey y Company, 2010). En el caso chileno, la gobernanza ha sido históricamente un tema complejo, ya que la municipalización hacía que el sistema público tuviera una autoridad diluida, con “dos cabezas” (el Ministerio de Educación y los municipios), con niveles difusos de responsabilización (Román, 2007), aspecto que aborda directamente el Nuevo Sistema de Educación Pública.

Un segundo aspecto responde al tema de la *escala*, entendido como el nivel de anidamiento de los establecimientos públicos. Al igual que en el caso de la gobernanza, en este aspecto también hay dos asuntos que deben balancearse. Por una parte, la necesidad de generar servicios intermedios en una escala que sea eficiente en términos económicos; y por otra, que estos servicios puedan entregar eficientemente el servicio educativo, en cuanto al apoyo, asesoría y gestión pedagógica (Andrews, Duncombe y Yinger, 2002; Robertson, 2007). En el caso chileno, hasta antes de la entrada en vigencia de la Ley 21.040 existían 346 servicios intermedios (cada municipio corresponde a un servicio), lo que también ha sido considerado un aspecto crítico del sistema educativo público (Marcel y Raczynski, 2009), y es abordado por el Nuevo Sistema de Educación Pública, al crear 70 Servicios Locales de Educación, que estarán distribuidos a lo largo del territorio de manera similar a la configuración administrativa-territorial del país y que agruparán, en promedio, cuatro comunas colindantes.

Finalmente, existe un tercer nudo crítico, que puede ser definido como el de la *escalabilidad*. La escalabilidad puede definirse como la capacidad que tienen los programas o reformas de ser replicables en distintos contextos, escenarios y ambientes. Como indica Coburn (2003), el proceso de escalabilidad no implica solo un crecimiento en la complejidad y desarrollo de los programas, sino también importantes transformaciones en la profundidad, sustentabilidad y distribución de los recursos implicados en las reformas. Algunos autores han mostrado la importancia de la participación de los actores (Álvarez et al., 2010) y de las investigaciones (Roschelle, Tatar y Shecht-

man, 2008) en el proceso de escalabilidad. La instalación progresiva del Sistema de Nueva Educación Pública (hasta completarse en el año 2025 con la inclusión de la totalidad de las municipalidades) hace que este aspecto también pueda ser abordado por este proyecto.

De esta manera, el desarrollo e implementación del Sistema de Nueva Educación Pública encierra una serie de desafíos estructurales para el subsistema de educación pública, siendo, por lo mismo, crítico realizar una evaluación sobre esta iniciativa.

## 2. ENFOQUES PARA EVALUAR LA NEP. EVALUACIONES DE IMPACTO Y DE PROCESO.

En general, las evaluaciones de programas sociales se pueden desarrollar desde tres perspectivas: i) desde la mirada del diseño, donde se busca analizar la coherencia y consistencia de una determinada política; ii) desde la mirada de los procesos, donde lo central es estudiar la forma en cómo una determinada política se está implementando y; iii) desde la mirada de los resultados y del impacto, donde se busca estimar el efecto de una intervención. En nuestro caso, el estudio se centró especialmente en las dos últimas formas de evaluación, cuyos principales elementos se describen a continuación.

### 2.1. DIMENSIÓN EVALUACIÓN DE IMPACTO Y EL SISTEMA DE NUEVA EDUCACIÓN PÚBLICA.

La evaluación de impacto consiste en la estimación del efecto causal de un programa, intervención o tratamiento en una población determinada, lo que permite determinar si este tuvo o no los efectos esperados (Bernal y Peña, 2012; Khandker, Koolwal y Samad, 2010; Rubin, 1974; Murmane y Willet, 2011). En términos genéricos, cualquier evaluación de impacto de un programa trata de responder a la pregunta sobre qué hubiese pasado con la población sin la intervención en cuestión. Para realizar esto, las evaluaciones frecuentemente enfrentan dos problemas esenciales: i) el problema del contrafactual y; ii) el sesgo de selección<sup>1</sup>. La elección de métodos de evaluación de impacto es una decisión que debe basarse en gran medida en las características propias de la intervención y por sobre todo en las características de los grupos de tratamiento y control. En general, la literatura reconoce dos grandes tipos de métodos: i) los métodos experimentales y; ii) los métodos cuasi experimentales (ver por ej. Bernal y Peña, 2012).

Los métodos experimentales -que se caracterizan por la asignación aleatoria a los grupos de tratamiento y control- representan en la actualidad el estándar máximo en cuanto a evaluación de impacto, pues equipara las diferencias de inicio de los grupos de tratamiento y control con una probabilidad conocida y permite llegar a inferencias causales sin la necesidad de controlar por otros factores intervinientes (Bernal y Peña, 2012). En América Latina ya se han realizado este tipo de evaluaciones para distintas políticas públicas en educación -como los casos de *“Una laptop por niño”* en Perú (Cristia, Ibararán, Cueto, Santiago y Severín, 2012) y *“PROGRESA”* en México (Behrman, Sengupta y Todd, 2001). A pesar de estos beneficios, el Sistema de Nueva Educación Pública (NEP) no puede evaluarse según este método, ya que el de gestión financiera de las unidades sostenedoras de los grupos de tratamiento y

<sup>1</sup> El primer problema refiere a la imposibilidad de medir el efecto de una intervención considerando un escenario donde la población participa del programa y otro escenario donde no participa, mientras que el segundo problema dice relación con las características propias de los grupos tratados y de comparación, así como también con las condiciones a las que están expuestos ambos grupos.

control. Los resultados muestran que un porcentaje mayor de establecimientos del grupo de control presenta déficit financiero. Además de la relevancia que por sí mismo tiene este dato, es importante que esta sea una variable a controlar en futuras evaluaciones intermedias y de impacto de la NEP, ya que la capacidad financiera de los establecimientos es una importante variable explicativa de una multiplicidad de procesos y prácticas que ocurren (o no ocurren) al interior de las escuela.

Por este motivo, es necesario explorar el segundo grupo de métodos: los diseños cuasi-experimentales. Estos métodos resultan más apropiados, ya que pueden aplicarse para estimar impacto en grandes poblaciones en contextos donde no es posible aleatorizar la asignación al tratamiento. Los métodos cuasi-experimentales, al tratar con grupos no aleatorizados, y que por tanto pueden no ser idénticas entre sí, incorporan una serie de supuestos y procedimientos que disminuyen la incidencia del sesgo de autoselección en la estimación de impacto. Si bien estos métodos sacrifican los altos niveles de validez interna que tienen los diseños experimentales, permiten obtener resultados altamente útiles para las decisiones de política pública (e.g. Shadish, Cook y Campbell, 2002).

Dentro de los métodos cuasi-experimentales encontramos los experimentos naturales y las técnicas de emparejamiento. Los experimentos naturales son escenarios que producen resultados similares a los experimentos aleatorios, pero donde el método de asignación se asemeja a la aleatorización porque es producto de un “evento fortuito” el que determina que un grupo de sujetos esté bajo tratamiento y otro no. Definir con claridad este evento que imprime la distinción de los grupos es esencial para definir cuánto es necesario controlar o no las características de ambos grupos, y, en definitiva, qué tipo de método de evaluación es posible utilizar. El caso de la aplicación de una política pública de forma generalizada puede considerarse un experimento natural, siempre y cuando su aplicación no dependa de la voluntad de los actores afectados, para evitar así el sesgo de selección<sup>2</sup>. Aunque la ley que crea el Nuevo Sistema de Educación Pública tiene características que se asemejan a las condiciones de un experimento natural, no es factible llevar a cabo este diseño en su definición pura, puesto que es indispensable crear una muestra para el grupo de control para levantar datos de proceso que den cuenta de la teoría de acción de la política y, de esta forma, entender posibles factores que podrían estar interviniendo en la “puesta en práctica” de la iniciativa.

Para resolver este último problema, se puede aplicar otra técnica: los métodos de emparejamiento, que se basan en la idea de unir individuos similares pertenecientes al grupo de control y tratamiento, controlando diferencias en los grupos que podrían sesgar los resultados de una intervención. En la estimación del efecto mediante emparejamiento se incluyen estas características observables como variables de control, de manera de aislar todo sesgo de selección de los resultados de la intervención (Bernal y Peña, 2012; Rosenbaum y Rubin, 1983). Existen diversos modos de estimar similitud entre individuos de grupos de tratamiento y control cuyos promedios en características observables difieren entre sí (Conniffe, Gash y O’Connell, 2000). Dos de los más comunes y que han sido utilizados previamente para la evaluación de políticas educativas (Treviño, 2007; Bellei, 2013; Godoy, Romo y Treviño, 2017) son el Emparejamiento

---

<sup>2</sup> A modo de ejemplo, investigadores como Susan Dynarski (1999) en Estados Unidos, y Cristián Bellei (2009) en Chile, han utilizado esta metodología para evaluar políticas de financiamiento a la educación superior y la extensión de la Jornada Escolar Completa.

por Puntajes de Propensión ó *Propensity Score Matching* (PSM) y la Expansión por el Inverso de los Puntajes de Propensión o *Inverse Propensity Weighting* (IPW)<sup>3</sup>.

La política de Nueva Educación Pública (NEP), dadas las características de su implementación, puede ser evaluada a través del uso de técnicas de emparejamiento. Esto permitirá identificar ex ante las comunas del grupo de control que, al momento del inicio de la política, son similares en variables observadas a las comunas que integran los primeros Servicios Locales que se implementen. Es necesario, sin embargo, tener especial consideración del cronograma de implementación, pues el universo de comunas para seleccionar el grupo de control debería limitarse a aquellas que ingresarán en los años finales en el cronograma de implementación de la política establecido en la ley. Extenderemos este punto en sección de conclusiones y recomendaciones.

## 2.2. DIMENSIÓN DE EVALUACIÓN DE PROCESOS Y EL SISTEMA DE NUEVA EDUCACIÓN PÚBLICA

Considerar el contexto de implementación es crucial para elaborar hipótesis informadas sobre los efectos probables de una política pública. Sin embargo, la relevancia del entorno no se agota en la etapa del diseño, sino que también debiese ser parte de la evaluación para comprender cómo intervienen—en los procesos y en los resultados—elementos que rodean a la política y que escapan a la agencia de los ejecutores. En este sentido, una evaluación de procesos busca comprender cuáles son, cómo operan y cómo se articulan aquellos ámbitos de funcionamiento que se verán afectados directa o indirectamente por alguna política (Weiss, 1972), en este caso, por la nueva institucionalidad de la Nueva Educación Pública. Así, la evaluación cualitativa tiene la posibilidad de complementar los resultados de la evaluación de impacto con explicaciones sustantivas y contextualizadas (Acharya, Blackwell y Sen, 2016; como se cita en Maldonado, Krönmüller y Gutiérrez, 2016).

En el ámbito de las políticas públicas, las evaluaciones de procesos han sido útiles para identificar, entre otros aspectos, procedimientos, narrativas, trayectorias y expectativas de actores relevantes. En el área educativa en Chile, por ejemplo, se han realizado estudios destinados a indagar cualitativamente en los procesos y condiciones de operación del sistema de financiamiento del sector municipal (CEPPE UC, 2017), o a comprender los elementos clave que confluyen en las transiciones educativas de los estudiantes durante la enseñanza media (CPCE, 2016). En general, estos estudios han combinado el uso de técnicas cualitativas -como entrevistas y grupos focales- con el análisis de fuentes secundarias, cuya triangulación tiene por resultado un entendimiento más acabado de los fenómenos y su desarrollo.

Algunos de los cambios más inmediatos y notorios del Nuevo Sistema de Educación Pública para quienes trabajan en los equipos técnicos de educación serán las modificaciones organizacionales y el cambio de escala. Esto necesariamente modificará una serie de reglamentos de gestión administrativa y financiera, y requerirá de un período de aprendizaje de los actores involucrados en términos prácticos y de conocimiento.

<sup>3</sup> El emparejamiento mediante PSM se basa en la estimación de un índice de probabilidad de participación a partir de características observables que determina la probabilidad a priori de participar en la intervención (Conniffe, Gash y O'Connell, 2000; Heckman, Ichimira y Todd, 1998). Por otro lado, el Inverse Propensity Weighting (IPW) también se basa en la construcción de un índice de probabilidad de participación en el programa, permitiendo ponderar cada observación para hacer que ambos grupos, en promedio, sean estadísticamente parecidos y comparables. Por lo tanto, IPW es un método especialmente ventajoso para controlar diferencias en grupos de comparación disímiles y que además son altamente heterogéneos.

Dado que la evaluación de impacto de la Nueva Educación Pública requerirá de un considerable período de maduración previo, la evaluación de procesos cumple un rol crucial. A través de ella es posible identificar dinámicas a nivel micro que tal vez no hayan sido cuidadosamente tomadas en cuenta por la política, y que pudiesen influir en su éxito (como problemáticas locales, el nivel de conocimiento o comprensión de la nueva institucionalidad, expectativas y apoyo de los actores involucrados, entre otras).

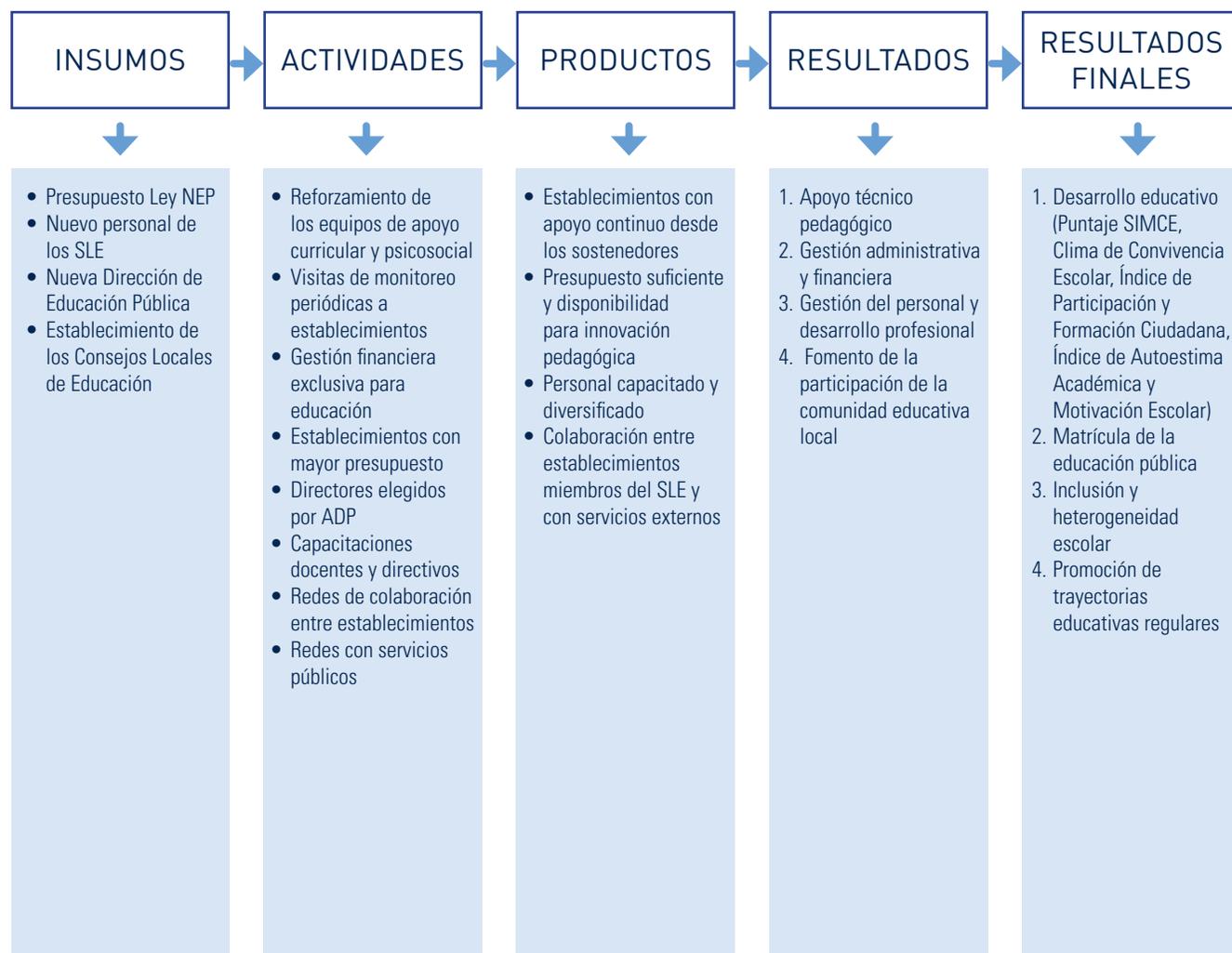
Este conocimiento, sumado al monitoreo de los indicadores de resultados intermedios, son cruciales para hacer seguimiento a la política durante sus primeros años. Acorde con esta finalidad, instrumentos como la entrevista y el grupo focal son apropiados para acceder a narraciones que den cuenta del contexto inmediatamente anterior a la puesta en marcha de la ley, así como de la evolución de la implementación, ajuste y aprendizaje de los sostenedores y las escuelas en el tiempo.

### 3. TEORÍA DEL CAMBIO DE LA NEP QUE ENMARCA SU EVALUACIÓN: PROPUESTA CONCEPTUAL PARA EL DISEÑO DE UNA EVALUACIÓN DE RESULTADOS, IMPACTO Y PROCESOS.

Dada la discusión sobre la evaluación de procesos y de impacto descrita anteriormente, se procedió a generar una propuesta conceptual de la evaluación, que diera cuenta de decisiones críticas del diseño de cualquier evaluación: qué evaluar, a quiénes hacerlo y qué indicadores utilizar (Khandker, Koolwal y Samad, 2010). Para ello, se utilizó la metodología de la “teoría de cambio”. Una teoría de cambio es una descripción lógica y argumentada sobre cómo se espera que el programa alcance el impacto esperado en la población (Gertler et al., 2011). De esta forma, establecer una teoría de cambio al principio de cualquier evaluación ayuda a clarificar cual es el flujo lógico de eventos y de actores involucrados en el programa. Para facilitar la elaboración de la teoría de cambio, y sobre todo para integrar los supuestos que hay detrás del modelo de impacto del programa y con ello lograr un plan de evaluación objetivo y realista, es de gran utilidad elaborar una “cadena de resultados” (Gertler et al., 2011). Una cadena de resultados es una descripción gráfica, lógica y temporal de los resultados esperados por el programa, y que incluye los insumos (todo tipo de recursos disponibles por el programa), las actividades (cómo estos recursos se despliegan), los productos (bienes y servicios que ofrece el programa), los resultados (cambios a corto plazo) e impacto o resultados finales (cambios a largo plazo).

La literatura destaca tres componentes principales de la cadena de resultados. Primero, la *implementación* del programa, que considera los insumos, actividades y productos del programa. Así, la implementación corresponde a todo lo que es desplegado deliberadamente por el programa para alcanzar sus objetivos. El segundo componente son los *resultados* y el *impacto*. Los resultados corresponden a los mediadores que deben generarse a partir de la política y que son precursores del impacto. El impacto, por su parte, se refiere a los cambios perdurables en el tiempo, generalmente asociados a los propósitos últimos de la política. Finalmente, el tercer componente de la cadena de resultados considera las *limitaciones* de la intervención, condicionantes que escapan a la agencia del programa y que podrían incidir en los resultados esperados. Considerando estos elementos, la Figura 1 muestra la cadena de resultados que este proyecto ha construido para conceptualizar la operación del Sistema de Educación Pública y su posterior evaluación. Como se puede observar, la teoría construida detalla cuatro dimensiones de resultados y cuatro dimensiones de resultados finales.

Figura 1. Teoría del Cambio y cadena de Resultados. Sistema de Nueva Educación Pública



Fuente: Elaboración propia CEPPE UC & DESUC.

#### 4. DISEÑO Y METODOLOGÍA DE EVALUACIÓN DE LA LÍNEA DE BASE DE LA NEP. TÉCNICAS, METODOS Y PROCESO DE RECOLECCIÓN DE INFORMACIÓN.

Después de revisar las características del Sistema de Nueva Educación Pública, el marco de referencia de las evaluaciones educativas y la teoría del cambio diseñada, el presente apartado presenta la metodología utilizada para la elaboración de la línea de base para la evaluación cuantitativa de resultados e impacto de la NEP, así como la evaluación cualitativa de procesos.

Debido a la extensión y heterogeneidad del actual sistema municipal, el estudio empleó métodos de investigación mixtos, combinando la recolección y análisis de datos cuantitativos y cualitativos para caracterizar de manera holística el momento inmediatamente anterior a la implementación de la NEP. Este diseño permite obtener datos más ricos, mediante la multiplicidad de observaciones y fuentes de información (Hernández Sampieri, Fernández Collado y Baptista Lucio, 2010). Además, posibilita una mayor riqueza interpretativa, combinando el potencial de generalización de los datos cuantitativos con la comprensión profunda que aportan las técnicas cualitativas (Creswell, 2014). La Tabla 1 resume los distintos abordajes utilizados, describiendo el tipo de metodología utilizada, los datos generados o usados y las fuentes de información.

Tabla 1. Resumen de metodologías para la elaboración de la línea de base

| <b>Etapa</b>             | <b>Tipo de metodología</b> | <b>Datos</b>   | <b>Fuentes de información</b>  |
|--------------------------|----------------------------|--|--|
| Evaluación de Procesos   | Cualitativa                | Primarios. Entrevistas semi-estructuradas y grupos focales   | Equipos sostenedores municipales<br>Directores de establecimientos             |
| Evaluación de Resultados | Cuantitativa               | Primarios. Encuestas diseñadas   | Directores y docentes de establecimientos municipales                          |
| Evaluación de Impacto    | Cuantitativa               | Secundarios. Datos administrativos y de resultados escolares a nivel de establecimientos y estudiantes | Bases de datos MINEDUC<br>Bases de datos de Agencia de Calidad de la Educación |

*Fuente: Elaboración propia CEPPE UC & DESUC.*

##### 4.1. DISEÑO DE LA EVALUACIÓN CUANTITATIVA DE RESULTADOS E IMPACTO

El diseño de la evaluación cuantitativa de resultados e impacto tiene como finalidad estimar los efectos de la NEP en las variables de resultados (apoyo pedagógico, gestión administrativa y financiera, gestión de recursos humanos y participación) y en los indicadores de impacto definidos en la teoría de cambio (desarrollo educativo, matrícula de la educación pública, inclusión y heterogeneidad escolar, y promoción

de trayectorias educativas regulares). A continuación se presentan los elementos centrales del diseño de la evaluación cuantitativa de resultados e impacto.

#### *a. Diseño muestral para la evaluación cuantitativa de resultados e impacto*

En términos metodológicos, un desafío del estudio fue escoger muestras enfrentando el hecho de que la autoridad pre-definió las comunas de tratamiento que entrarían inicialmente en el proyecto NEP. Dada esta situación, y aprovechando la gradualidad de la implementación para construir grupos de tratamiento y control, se establecieron dos criterios ex ante para la selección de establecimientos que participarían en la línea de base de la Nueva Educación Pública y que permitirían llevar a cabo en el futuro la evaluación de impacto. En primer lugar, se determinó que el universo del grupo de tratamiento estaría compuesto por escuelas que ingresarían al régimen NEP entre los años 2018 y 2019, conformando cinco Servicios Locales de Educación (SLE). En segundo lugar, se definió que el grupo de control estaría conformado por establecimientos de regiones cuyo ingreso está previsto para la segunda etapa de implementación de esta política entre los años 2022 y 2025. Después de definir estos criterios, para asegurar la similitud de los establecimientos que forman parte de la muestra (grupos de tratamiento y control), se utilizó el método de *Propensity Score Matching* (PSM) (Conniffe, Gash, y O’Connell, 2000; Heckman, Ichimira y Todd, 1998), el cual se basa en la estimación de un índice de probabilidad de participar en la intervención a partir de variables observables. Con ese fin, en el contexto de este estudio, se consideraron seis variables de las escuelas (ruralidad, vulnerabilidad socioeconómica, niveles impartidos, tamaño de la escuela y SIMCE lectura y matemática de 4° básico) y 4 variables de las comunas (número de establecimientos municipales, razón entre establecimientos municipales y el total de establecimientos de la comuna, gasto total en educación y deuda municipal). Una vez estimada la probabilidad de participación para todos los establecimientos, se seleccionó primero la muestra del grupo de tratamiento de los Servicios Locales de Educación que entran a la política en 2018 y 2019. Asimismo, se seleccionaron escuelas pares—indistinguibles estadísticamente según el puntaje de propensión—entre aquellas de regiones que ingresan a la política entre 2022 y 2025.

Para el grupo bajo tratamiento (establecimientos que ingresan pioneramente a la NEP) se consideró un diseño muestral probabilístico y estratificado entre los establecimientos previamente seleccionados. Anticipándose a posibles problemas de no contacto o rechazo, se amplió el número de unidades que corresponden a la muestra objetivo original de 200 escuelas a una muestra sobredimensionada<sup>4</sup> de 255 establecimientos. Por otra parte, la selección del grupo de control se realizó en función de los atributos de las escuelas que componen el grupo bajo tratamiento. Así, basados en la estimación de la probabilidad de participación, se utilizó el método de emparejamiento de unidades tratadas con 10 potenciales unidades de control para cada observación. Para seleccionar el par final de estas 10 potenciales unidades de control, se realizó una elección intencionada, procurando seleccionar unidades pertenecientes al menor número de comunas posible y cuya dispersión geográfica de las comunas sea, dentro de lo posible, similar al de comunas pertenecientes al grupo de tratamiento. Esto con la intención de poder evaluar adecuadamente

<sup>4</sup> La sobredimensión de la muestra objetivo es una práctica común utilizada en grandes encuestas y promovida por la Asociación Americana de Opinión Pública (AAPOR), que permite considerar al momento del diseño muestral, las tasas de rechazo y no contacto propias de encuestas similares. De esta manera, el sorteo se realiza una sola vez, y permite claridad en los reportes de desempeño del terreno.

los cambios en los indicadores de resultados relativos a la gestión comunal (apoyo pedagógico, gestión administrativa y financiera, gestión de personal y fomento a la participación) al contar con un mayor número de escuelas por municipio. Para este caso, la muestra sobrestimada correspondió a 236 establecimientos.

*b. Técnicas de recolección y de análisis de datos para la evaluación de resultados e impacto.*

Para la evaluación de resultados e impacto se utilizaron una serie de fuentes de información, combinando recolección de datos primarios y recolección de datos secundarios. La Tabla 2, que se presenta a continuación, resume las fuentes de datos utilizadas, caracterizando su origen y dimensiones

**Tabla 2. Resumen de fuentes de datos para el levantamiento de la línea de base para la evaluación de impacto y de procesos de la NEP**

| Evaluación                         | Tipo de datos | Fuente de datos                       | Datos utilizados    | Unidad de observación                 | N° de casos |
|------------------------------------|---------------|---------------------------------------|---------------------|---------------------------------------|-------------|
| Evaluación de resultados e impacto | Secundarios   | Agencia de la Calidad de la Educación | IDPS 2016           | Establecimientos escolares            | 10.281      |
|                                    |               | Agencia de la Calidad de la Educación | Resultados SIMCE    | Estudiantes                           | 495.684     |
|                                    |               | MINEDUC                               | Rendimiento 2016    | Estudiantes                           | 3.226.943   |
|                                    |               | MINEDUC                               | Matrícula 2016      | Estudiantes                           | 3.550.949   |
|                                    | Primarios     | Levantamiento propio                  | Encuesta Directores | Directores establecimientos escolares | 368         |
|                                    |               | Levantamiento propio                  | Encuesta Docentes   | Docentes establecimientos escolares   | 1.691       |

**Fuente:** *Elaboración propia CEPPE UC & DESUC. Nota:* *Para las fuentes de datos secundarias, el número de casos indica el número total de observaciones incluidos en cada base. Para las fuentes primarias, en cambio, el número de observaciones indica el total de encuestas alcanzadas.*

Para la recolección de datos de la evaluación de resultados e impacto se utilizaron fuentes de datos secundarias y primarias. Las primeras corresponden a las bases de datos del sistema escolar chileno del MINEDUC y SIMCE de la Agencia de la Calidad de la Educación para el último año escolar finalizado a la fecha de la evaluación, el 2016. Para la recolección de datos primarios se levantaron dos encuestas una administrada al director y otra a docentes (en promedio 4) de cada establecimiento. La encuesta, elaborada y piloteada<sup>5</sup> por el equipo evaluador, abordaba las diferentes dimensiones del quehacer administrativo, pedagógico, temas de gestión, toma de decisiones y vínculos externos desarrollados por el establecimiento, siguiendo las directrices de la teoría del cambio esbozada.

<sup>5</sup> El pilotaje de los instrumentos se realizó durante el mes de octubre de 2017 en 10 establecimientos de la Región Metropolitana que no formaban parte de la muestra definitiva.

Para la elaboración de la línea de base se realizó un análisis descriptivo de diferencia de medias, de proporciones y análisis de regresión *OLS* y *Logit* para explorar diferencias estadísticamente significativas entre los grupos de tratamiento y control en los indicadores de resultado y de impacto. Para los indicadores de resultados, se utilizaron los datos recogidos por el equipo evaluador a través de las encuestas a directores y docentes. Para la construcción de los indicadores se realizó Análisis de Componentes Principales (Jolliffe, 2002). Para los indicadores de impacto, aquellos en los que se espera que la NEP genere un efecto esperado a largo plazo y sostenible en el tiempo, se utilizaron los datos secundarios y se reportan tanto a nivel de estudiantes como de establecimientos. Para testear diferencias entre los grupos de tratamiento y control tanto en indicadores de resultados e impacto, se ponderó cada observación siguiendo el método de *Inverse Propensity Weighting* (IPW), método que permite ponderar cada observación según su probabilidad de participar en la intervención, asegurando la mayor similitud posible de los grupos.

El levantamiento de la línea de base, si bien entrega resultados importantes sobre la similitud inicial de los grupos de tratamiento y control, es, por sobre todo, un insumo relevante para la futura estimación del impacto de esta nueva política, cuyo levantamiento de datos debería ocurrir el año anterior al ingreso de los establecimientos del grupo de control a la política NEP. Para la evaluación de impacto o evaluaciones intermedias que podrían realizarse, el equipo evaluador sugiere utilizar los puntajes del *Inverse Propensity Weighting* (IPW) para ponderar los casos de la evaluación. Posteriormente, se sugiere usar un método de análisis multivariado, según corresponda al tipo de variable dependiente, para estimar el impacto introduciendo una variable dicotómica que identifique a los establecimientos de tratamiento y que controle por el valor que tuvo la variable dependiente en la línea de base. Dado que el IPW controla por diferencias en variables observadas, el análisis de impacto se podría realizar con el método de Diferencias en Diferencias (dif-dif), dado que este contribuye a controlar por diferencias no observadas, las cuales se revelan especialmente en los casos en que hay diferencias significativas en las variables dependientes en la línea de base (Khandker, Koolwal y Samad, 2010).

## 4.2. DISEÑO DE LA EVALUACIÓN CUALITATIVA DE PROCESOS

La evaluación cualitativa de procesos busca recoger las opiniones de los actores respecto de la forma en que se desarrollan los procesos de apoyo a las escuelas en la gestión local. Para ello, se diseñó un estudio cualitativo que es esencial para complementar la evaluación cuantitativa de resultados e impacto.

### a. Diseño muestral de la evaluación cualitativa de procesos

A diferencia de lo que ocurre en la evaluación cuantitativa de resultados e impacto, la unidad de selección para la evaluación de procesos corresponde a las unidades sostenedoras municipales: DAEM o Corporaciones. Para definir la muestra se utilizó el procedimiento de “muestreo intencionado” (Patton, 1987) o también conocido como “selección basada en criterios” (Goetz y Le Compte, 1988), que consiste en un ejercicio deliberado de atender a un conjunto de características que son “representativas” de la población de interés. Para esta evaluación, el universo de comunas correspondió a aquellas que forman parte de la muestra para la evaluación de impacto, manteniendo la diferenciación entre comunas de tratamiento y de control. El principal criterio para seleccionar comunas fue la representación de todos los SLE

que ingresan (dentro del grupo de tratamiento) o potenciales (grupo de control), seleccionando dentro de ellos dos comunas, una predominantemente urbana y otra rural. Las unidades de observación corresponden a la unidad sostenedora y directivos de escuela de cada comuna. En total, se seleccionaron 10 comunas del grupo de tratamiento y 10 comunas del grupo de control.

#### *b. Técnicas de recolección y de análisis de datos para la evaluación cualitativa de procesos*

Para la evaluación de procesos, por su parte, se utilizaron entrevistas semi-estructuradas individuales y grupales, elaboradas por el equipo evaluador y que abordan aspectos referidos a la estructura organizacional y de gestión (financiera y administrativa) de las unidades sostenedoras, y apoyo pedagógico brindados por la unidad a las escuelas. Para la sistematización de los datos recogidos se utilizó el *software DEDOOSE* de análisis de datos cualitativos.

Basado en una estructura de códigos previamente elaborada más la inclusión de códigos que emergieron inductivamente, se realizó un proceso de análisis en tres niveles. En un primer nivel, se realizó una codificación abierta, la cual permitió reducir los discursos de los participantes a códigos o unidades de sentido en torno a la implementación de la NEP. En un segundo nivel, de mayor condensación y abstracción que el anterior, se realizó una codificación axial para identificar los grandes ejes temáticos que respondan a los objetivos del estudio. Finalmente, en un tercer nivel, se realizó una codificación selectiva, mediante la cual se generaron relaciones intercategoriales parsimoniosas entre las categorías de análisis.

De esta manera, se buscó estudiar la puesta en práctica existente a la fecha de ciertas temáticas críticas para el desarrollo del Sistema de Educación Pública, de tal manera que pudiera comprender los procesos críticos necesarios de entender para la implementación de esta política.

## 5. HALLAZGOS DE LA EVALUACIÓN DE LÍNEA DE BASE.

A continuación se presenta un resumen de los principales hallazgos de la evaluación de la línea de base, distinguiendo entre aquellos provenientes de la evaluación cuantitativa de resultados e impacto y los de la evaluación cualitativa de procesos.

### 5.1. HALLAZGOS DE LA EVALUACIÓN CUANTITATIVA DE RESULTADOS E IMPACTO

En el presente apartado se presentan los resultados de la línea de base para la evaluación cuantitativa de resultados e impacto de la NEP. Estos hallazgos, que se dividen entre indicadores de resultados y de impacto (corto y largo plazo, respectivamente) y ofrecen un panorama general de la situación de las escuelas y sostenedores antes de la puesta en marcha del Nuevo Sistema de Educación Pública (NEP).

#### *a. Indicadores de resultados*

Como se mencionó, los indicadores de resultados fueron elaborados a partir de las encuestas aplicadas a directores y docentes de los establecimientos educacionales de los grupos de tratamiento y control. Para la mayoría de los índices se realizó un

Análisis de Componentes Principales, de manera de reducir una batería de datos correlacionados provenientes del cuestionario en componentes o nuevas variables. En total, se generaron 19 indicadores agrupados en cuatro dimensiones en las que la Nueva Educación Pública explicita actividades e insumos para generar cambios positivos: i) apoyo pedagógico; ii) gestión administrativa y financiera, iii) gestión del personal y desarrollo profesional y; iv) participación de la comunidad educativa y local.

Los hallazgos muestran que en 15 de los 19 indicadores construidos no se encontraron diferencias estadísticamente significativas entre grupo de tratamiento y grupo de control. Los cuatro indicadores que mostraron diferencias fueron: i) *eficiencia en la toma de decisiones financieras*; ii) *nivel declarado de déficit financiero*; iii) *participación de padres y apoderados en actividades escolares* y; iv) *número de capacitación de instituciones privadas*. En los tres primeros casos, las diferencias fueron a favor del grupo bajo tratamiento. Respecto de la *eficiencia* en la toma de decisiones financieras y el *nivel declarado* de déficit, los resultados podrían tener relación con el trabajo de regularización administrativa que llevó a cabo el Ministerio de Educación con los municipios en el año 2017 como parte de la preparación para la puesta en marcha de los primeros SLE. En el caso del indicador de *capacitaciones privadas*, esta diferencia podría explicarse, al menos en parte, por la inclusión de las comunas que actualmente pertenecen al Servicio Local de Barrancas, en la Región Metropolitana. Dado que es en esta región donde se concentra una mayor oferta de servicios privados de capacitación en el país, lo que podría haber sido un factor para generar esta diferencia.

Ahora bien, respecto de los indicadores de *apoyo pedagógico*, los hallazgos muestran que los establecimientos realizan y desarrollan ciertas prácticas de apoyo para mejorar la calidad de los aprendizajes y para la atención a la diversidad, no obstante hay espacio para diversificarlas. Lo mismo ocurre con las prácticas e iniciativas para el apoyo pedagógico que son entregadas desde la unidad sostenedora. Por otra parte, los docentes reconocen la existencia de instancias de colaboración entre docentes al interior de los establecimientos que apunten a la mejora de la calidad del aprendizaje, como la realización de actividades en conjunto, integrando a distintos cursos o retroalimentando la labor pedagógica entre docentes del mismo establecimientos, aunque estas actividades suelen ser esporádicas. Finalmente, los docentes declaran contar con altas competencias profesionales para lograr que estudiantes de distinto origen cultural, económico, o con dificultades emocionales, de comportamiento o aprendizaje aprendan.

En lo referente a la dimensión *gestión administrativa y financiera del establecimiento*, los directivos declaran que el establecimiento es parte de la toma de decisiones sobre el uso de recursos orientados a la mejora en los aprendizajes. En lo que respecta a la toma de decisiones financieras, la mayoría de los directivos están conformes con la eficiencia en los pagos de remuneraciones y gastos operacionales por parte del sostenedor, así como también sobre la disponibilidad de este último para resolver urgencias económicas de las escuelas. Finalmente, el 40% de las escuelas declara que sus ingresos no alcanzan a cubrir los costos del establecimiento.

La dimensión *gestión del personal y desarrollo profesional* entrega información sobre la capacidad de las escuelas y sus sostenedores de ofrecer oportunidades de capacitación para los equipos directivos y docentes, así como también de integrar a estos últimos en la toma de decisiones de la escuela. Los indicadores de esta dimensión muestran que poco más de la mitad de las escuelas declara recibir capacitaciones docentes desde

agencias gubernamentales y cerca de un 35% recibe de instituciones privadas, con una frecuencia de 2,5 a 1,8 al año, respectivamente. Además, los directores señalan que la inclusión de los docentes en la toma de decisiones de la escuela es muy recurrente.

Finalmente, los indicadores de la dimensión *participación de la comunidad educativa y local* muestran que, a juicio de los directores, la mayoría de los apoderados participa de las actividades promovidas por el establecimiento. Además, indican que existe un alto nivel de influencia de los equipos directivos en la toma de decisiones de la escuela, y un poco menor nivel de influencia de otros actores escolares. En cuanto al establecimiento de redes de colaboración, se destaca que cerca de un 60% declara establecer redes de colaboración con equipos docentes o directivos de otros establecimientos, y una proporción similar establece redes de colaboración con otros organismos.

### *b. Indicadores de impacto*

Los indicadores de impacto, por su parte, son en su mayoría indicadores de datos administrativos -esto es, bases de datos de acceso público de MINEDUC y SIMCE. El uso de estos datos permite indagar sobre las diferencias en los grupos de tratamiento y control en variables relativas a los estudiantes usando datos censales, en vez de utilizar únicamente datos muestrales como ocurre con las encuestas aplicadas a directores y docentes en la evaluación de resultados. Los indicadores de impacto se dividen en aspectos en donde se espera que la NEP genera un efecto a largo plazo, agrupándose en cuatro dimensiones: i) desarrollo educativo; ii) matrícula de la educación pública; iii) inclusión y heterogeneidad escolar y; iv) trayectorias educativas regulares.

En total, la evaluación contempló el análisis de 20 indicadores (de 11 variables) agrupados en las cuatro dimensiones mencionadas. En total, no se encontraron diferencias significativas entre grupo de tratamiento y grupo de control en 15 indicadores, existiendo diferencias en solo 5 de estos: i) *clima de convivencia escolar en 4° básico*, ii) *clima de convivencia escolar en 2° medio*; iii) *autoestima académica y motivación escolar en 2° medio*; iv) *participación y formación ciudadana en 2° medio* y; v) *porcentaje de matrícula de educación pública*. En todos estos casos, la diferencia es en favor del grupo de tratados. En los cuatro primeros, las diferencias podrían explicarse por la inclusión de indicadores de desarrollo educativo que no fueron controlados durante la selección de la muestra, por la reciente creación de estos indicadores por parte de la Agencia de la Calidad. Finalmente, la diferencia en el indicador de la dimensión de matrícula en la educación pública, podría explicarse por ser un indicador a nivel comunal, a diferencia del resto de los indicadores (que son a nivel de establecimientos o de estudiantes).

La dimensión de *desarrollo educativo* engloba el desempeño académico en las evaluaciones SIMCE de lectura y matemática de 4° básico y 2° medio y algunas de las variables que forman parte de los Indicadores de Desarrollo Personal y Social (IDPS) promovidos por la Agencia de la Calidad de la Educación. A nivel de establecimientos, los puntajes SIMCE de ambos grupos son similares entre sí y similares al promedio nacional de establecimientos municipales. Por su parte, para los tres Indicadores de Desarrollo Personal y Social que son parte de esta dimensión (*clima de convivencia escolar*, *autoestima académica y motivación escolar* y *participación y formación ciudadana*), los promedios son similares a la media nacional de establecimientos municipales o levemente inferior en el

grupo de control, en los casos en que las diferencias son estadísticamente significativas. La dimensión de *matrícula en la educación pública* entrega información sobre el porcentaje de matrícula escolar municipal de las comunas en las que se encuentran los establecimientos. Dado que uno de los objetivos centrales de la NEP es fortalecer la educación pública, es esperable que este indicador aumente con el paso de los años. En su medición previa al inicio de la Nueva Educación Pública, la matrícula pública de las comunas del grupo de tratamiento es de un 57%, mientras que en las comunas del grupo de control el 48%. Esta diferencia es estadísticamente significativa, lo que puede deberse a que el emparejamiento se hizo a nivel de escuela, considerando variables municipales. Por su parte, el indicador de asistencia escolar anual es similar y bastante alto para ambos grupos (90% de asistencia anual), siendo superior a la media de los establecimientos escolares municipales del país (84%).

La dimensión de *inclusión y heterogeneidad* incorpora indicadores de heterogeneidad socioeconómica y académica en las escuelas. Ambos índices se construyeron utilizando información del SIMCE: el índice de heterogeneidad académica utiliza información socioeconómica reportada en los cuestionarios respondidos por los padres, y el índice de heterogeneidad académica se construye a partir de la dispersión de resultados SIMCE de una misma escuela. Los dos se calculan para 4° básico y 2° medio. Los hallazgos muestran que ambos grupos están compuestos por escuelas con similar nivel de heterogeneidad socioeconómica y académica, la que, en promedio, es levemente mayor que el resto de los establecimientos municipales del país.

Finalmente, la dimensión de *trayectorias regulares* presenta indicadores que permiten monitorear la magnitud de eventos disruptivos en las trayectorias escolares, como repitencia y deserción, y de transición, como la titulación en la modalidad Técnico Profesional. Los datos muestran que menos de un 5% de los estudiantes de los establecimientos que forman parte de la muestra repitió en 2016, y cerca de un 2% desertó del sistema.

## 5.2. HALLAZGOS DE LA EVALUACIÓN CUALITATIVA DE PROCESOS

Los resultados dan cuenta de las percepciones de los actores en relación a los procesos, recogidas a través de entrevistas a sostenedores y directores, que se desarrollan actualmente en la administración de educación pública y se organizan en las cuatro dimensiones destacadas en la teoría del cambio (gestión administrativa y financiera; gestión del personal y desarrollo profesional; apoyo técnico pedagógico y participación de la comunidad educativa y local), más una última sección sobre las percepciones de los actores sobre el sistema municipal actual y percepción de los actores sobre la Nueva Educación Pública.

### *a. Gestión administrativa y financiera: foco en rendiciones por sobre foco en la gestión.*

Esta dimensión buscó identificar y entender los procesos de planificación y ejecución financiera tanto en términos de actores, fuentes de financiamiento y procedimientos involucrados, así como balances generales de su estado financiero. En general, los resultados mostraron que los procesos de planificación y ejecución financiera, así como la cantidad y calidad del personal encargado de finanzas son sumamente variables entre las comunas, situación que aparece entre los actores de los grupos de tratamiento y de control entrevistados. A pesar de estas diferencias, en la mayoría de los casos la planificación y gestión financiera y administrativa

suelen enmarcarse dentro de la actualización del Plan Anual de Desarrollo de la Educación Municipal (PADEM), que da coherencia al gasto de recursos en función de planes de acción definidos y coordina la interacción entre municipalidad, DAEM/Corporación y establecimientos educacionales. Adicionalmente, los entrevistados destacaron el bajo margen de acción que tendrían como sostenedores respecto del uso de recursos. En esta línea, indicaron, por ejemplo, que se destinaría entre un 75% y hasta un 90% sólo a remuneraciones, y que parte importante de la gestión administrativa y financiera se focalizaría en rendir los distintos tipos de subvenciones y fondos educacionales (subvención regular, SEP, PIE, FAEP, entre otros), no existiendo mayor gestión adicional a estos procesos.

#### *b. Gestión del personal y desarrollo profesional: límites en los procesos de selección de personal.*

La gestión del personal y desarrollo fue entendida como el conjunto de procesos de contratación y dotación de personal, agregando, en el caso de los establecimientos, una caracterización de las instancias de desarrollo profesional docente. El tema más recurrente respecto a esta dimensión fueron los procesos de contratación, donde los actores destacaron que existiría una falta de estructura y protocolos formales o estandarizados, tanto a nivel de DAEM/Corporación, como de establecimientos. A nivel de establecimientos, la distancia geográfica y la escasez de oferta de profesionales serían factores que influirían fuertemente en comunas rurales, donde la percepción es que los procesos de búsqueda y selección de profesionales deben adaptarse a la oferta disponible, lo que implicaría una menor exigencia en ellos. Por otra parte, los sostenedores y sus equipos destacaron que existiría una influencia política en algunos casos de procesos de contratación de personal a nivel de DAEM/Corporación y también de establecimientos, especialmente por la influencia de los alcaldes. Respecto a las instancias de desarrollo profesional docente, se identificaron múltiples iniciativas, que surgirían tanto desde los establecimientos —donde el DAEM/Corporación cumple la función de gestionar el correcto desarrollo de éstas— como instancias que son iniciativas del DAEM. En todos estos aspectos se apreciaron similitudes en los patrones de las comunas de los grupos de tratamiento y de control.

#### *c. Apoyo técnico-pedagógico: disparidad respecto de su rol como sostenedor.*

La dimensión de apoyo técnico-pedagógico buscó describir la forma en que se desarrollan procesos de planificación, las percepciones respecto del apoyo técnico pedagógico que otorga el sostenedor u otras instituciones, y las redes que se forman entre establecimientos. Además, también se abarcaron algunas iniciativas de innovación pedagógica. En términos generales, se observa que la planificación pedagógica se ciñe a los lineamientos planteados desde el equipo sostenedor (DAEM o Corporación) a través del Plan Anual de Desarrollo de la Educación Municipal (PADEM). A su vez, cada establecimiento dice adaptar estos lineamientos para generar su planificación como escuela, a través de sus Planes de Mejoramiento Educativo (PME). El resultado más llamativo de esta dimensión fue la discrepancia que existiría entre los propios sostenedores municipales respecto de su rol en este ámbito. Así, mientras algunos sostenedores aseguran que legalmente su función se limita a la administración financiera, otros sostenedores han asumido también labores de soporte pedagógico (aunque la posibilidad de hacer esto depende, en gran medida, de sus recursos humanos y financieros) lo que implica que las iniciativas de ca-

pacitación docente, colaboración entre escuelas y asistencia pedagógica por parte del sostenedor difieren enormemente en cada caso. En este ámbito se encuentran situaciones similares entre todas las comunas estudiadas.

#### *d. Participación de la comunidad educativa y local: participación más formal que sustantiva.*

Esta dimensión buscó caracterizar la comunidad escolar (incluyendo el clima escolar y el involucramiento de las familias), la oferta comunal de escuelas y las relaciones entre ellas, y los espacios de participación de la comunidad educativa. Tres grandes resultados se desprenden de esta dimensión. En primer lugar, y de acuerdo a los distintos actores entrevistados, habría consenso en que la educación municipal recibe estudiantes con un alto nivel de vulnerabilidad social. Por ello, los directores entrevistados indicaron que los padres y apoderados podrían hacer un aporte acotado a nivel pedagógico, y que, por este motivo, se les intenta involucrar en actividades abiertas y comunitarias que les abrieran otra forma de participación en la educación de sus hijos. Aun así, existiría un relativo consenso en que la participación de los padres es más bien baja y difícil de mejorar. En segundo lugar, los establecimientos indican que existe colaboración entre las escuelas de la comuna, con una cierta cantidad de actividades conjuntas para eventos que involucran a la comunidad local y donde el vínculo se mantendría principalmente a través de los directores, aunque también se destacó la existencia de las redes pedagógicas (por nivel, asignatura o zona geográfica). Finalmente, y respecto a las instancias de organización internas, como Centro de Padres y Apoderados, Centro de Estudiantes y Consejos Escolares, las entrevistas mostraron que, si bien la mayoría de los colegios declaró estar cumpliendo con la normativa a este respecto, en la práctica estas instancias serían vistas en general como un requisito a cumplir, pero no como un espacio para que la comunidad educativa pueda incidir en las decisiones de la escuela. Vale la pena indicar que las tendencias son similares entre todas las comunas analizadas.

#### *e. Evaluación del actual sistema municipal y la Nueva Educación Pública: elementos para su transformación gradual*

En esta dimensión se buscó conocer las percepciones y opiniones de los actores entrevistados respecto al sistema actual, su nivel de conocimiento de la ley de Nueva Educación Pública, y sus expectativas respecto al cambio de institucionalidad. Respecto del sistema actual, el primer hallazgo es que parte importante de los entrevistados indicó que el principal nudo crítico de la educación pública sería el sistema de financiamiento, pues muchos DAEM/Corporación sostuvieron que el pago de subvenciones por asistencia generaría incertidumbre y déficit estructurales. En segundo lugar, los entrevistados indicaron que una falencia del sistema municipal sería su énfasis administrativo-financiero en desmedro de la labor técnico-pedagógica. Finalmente, los entrevistados destacaron que el sistema municipal actual entrega facilidades para una fluida relación entre los sostenedores y las escuelas, lo que permitiría que los sostenedores tuvieran mayor conocimiento de la realidad local.

Ahora bien, respecto al Sistema de Nueva Educación Pública, es posible establecer dos grandes hallazgos. Por una parte, existe un bajo nivel generalizado de conocimiento sobre la nueva ley. De esta manera, sólo se conocerían aspectos gruesos de la ley, como el traspaso de los establecimientos municipales al Estado y la creación de los Servicios Locales, lo que produciría una sensación de incertidumbre dentro

de los actores educativos respecto al futuro de la educación pública, su papel profesional en ella y los aspectos prácticos del traspaso. En segundo término, y respecto a las expectativas que tienen los actores sobre la NEP, las opiniones son más bien disímiles. Así, y aunque existiría un relativo acuerdo con los principios generales de la ley, los entrevistados mencionaron una serie de preocupaciones más bien prácticas respecto a la implementación relativas a aspectos como cuándo ingresarían las comunas, cómo se realizaría el traspaso (p.e., en términos laborales), así como dudas respecto a la efectividad de la nueva institucionalidad para mejorar la calidad de la educación municipal.

## 6. CONCLUSIONES Y RECOMENDACIONES: MODELO DE EVALUACIÓN Y ASPECTOS PARA EL DESARROLLO DE LA NUEVA EDUCACIÓN PÚBLICA (NEP).

Considerando los múltiples objetivos de la investigación, a continuación se presentan una serie de conclusiones y recomendaciones del estudio, agrupadas en dos ámbitos: i) recomendaciones respecto del proceso de evaluación de la Nueva Educación Pública y; ii) recomendaciones respecto del proceso de implementación de la Nueva Educación Pública en el corto y mediano plazo. En su conjunto, aspiran a informar la toma de decisiones de la Dirección Nacional NEP, MINEDUC, Consejo Evaluación, Directores de SLE, entre otros.

### 6.1 CONCLUSIONES Y RECOMENDACIONES RESPECTO DEL PROCESO DE EVALUACIÓN DE LA NUEVA EDUCACIÓN PÚBLICA.

- **Estabilidad del diseño metodológico y la muestra generada en la línea de base:** Uno de los resultados más relevantes de la investigación es que los distintos resultados son coincidentes en demostrar que, a grandes rasgos, los grupos de tratamiento y de control son comparables en los distintos niveles (estudiantes, establecimientos y comunas). Esto es central, pues estaría indicando que el ejercicio de emparejamiento realizado parece cumplir con las expectativas de una línea de base. En este sentido, se *recomienda a futuro utilizar, en lo posible, tanto el diseño metodológico generado como la misma muestra utilizada en la línea de base*. Lo central en este punto es considerar que ambos aspectos se encuentran altamente encadenados, ya que la muestra construida está desarrollada para la aplicación de una evaluación de impacto específica. Por lo mismo, el desarrollo de una evaluación de impacto posterior debe dar cuenta, por ejemplo, de los posibles problemas de atrición de la muestra (por ejemplo, posible cierre de escuelas) (Bernal y Peña, 2012) para, a partir de este análisis, desarrollar la evaluación del impacto. De esta forma, es relevante que el Estado (y varios actores claves) consideren la importancia de la estabilidad del diseño evaluativo con una perspectiva de largo plazo.
- **Ley de Nueva Educación Pública y ajuste de la teoría del cambio:** Uno de los logros del estudio fue la construcción de la teoría del cambio de la Nueva Educación Pública (NEP). Esto significó la creación de una cadena de resultados, que pudiera dar cuenta con claridad de los resultados e impacto de la intervención. Para lograr este objetivo, el equipo de investigación se basó fundamentalmente en los principios y características de la Ley sobre Nueva Educación Pública, buscando así interpretar el sentido y espíritu de la Ley. A pesar de que esta es una fortaleza de la evaluación realizada, es bastante probable que

exista una distancia entre la norma (tratamiento) sancionada y la forma en cómo esta norma (tratamiento) se implementa finalmente, lo que en la literatura se conoce como adherencia o fidelidad de la implementación (Durlak, 2010; Carroll et al., 2007). Por ello, se considera relevante *recomendar que en las próximas evaluaciones de la propuesta (evaluación intermedia y evaluación final) se realice un análisis empírico de la teoría de cambio, poniendo especial énfasis en la cadena de resultados propuesta y la distancia de esta cadena con la implementación de la NEP*. De esta manera, parece relevante adecuar la teoría de cambio a lo que la Nueva Educación Pública ha generado efectivamente durante su desarrollo, a partir de un análisis de la fidelidad con que la NEP se ha implementado, es decir, analizando la brecha entre lo indicado en la ley y lo realmente generado durante los primeros años de implementación de esta iniciativa.

- **Priorización de Indicadores de resultados y de impacto:** Relacionado con lo anterior, es necesario destacar que la ley de Nueva Educación Pública y, por lo tanto, también la teoría del cambio generada, ha señalado que la NEP traería múltiples transformaciones, impactando, entre otros, en resultados educativos y socioemocionales del estudiante y aspectos organizacionales, financieros y pedagógicos de la escuela. De esta forma, la propuesta generada ha señalado que la Nueva Educación Pública impactaría en un total de 16 resultados de corto plazo (agrupados en 4 dimensiones) y de 11 resultados de impacto (también agrupados en cuatro dimensiones). Esto quiere decir que la NEP impactaría directamente en 27 indicadores, lo que, evidentemente, entrega altas expectativas a la intervención. En la discusión sobre evaluaciones de impacto, esta gran cantidad de indicadores no es común, debido a que, en términos teóricos, se espera que una intervención transforme determinados aspectos de la realidad social (Angrist y Pischke, 2008). Por ello, *se recomienda que, en virtud de la forma en cómo se desarrolle la implementación de la NEP y las prioridades que esta forma de implementación tome, se genere una discusión (y, eventualmente, una priorización) respecto del conjunto de los indicadores (de resultados y de impacto) relevados en la línea de base*. Esto no implica necesariamente descartar o limitar los indicadores, sino que otorgar mayor o menor peso a ciertos indicadores, en virtud de la centralidad que estos hayan tomado para el desarrollo efectivo de la implementación de la Nueva Educación Pública.
- **Tiempos y momentos de la evaluación de la NEP:** Como puede desprenderse del análisis de anteriores evaluaciones de impacto en términos de su diseño, dimensión y alcance (p.e., la Jornada Escolar Completa o el Programa Progresá), los procesos del impacto de una iniciativa pueden tardar varios años en generar los efectos esperados. Aunque no existe un consenso respecto de cuál es exactamente el tiempo de maduración de una intervención educativa, una guía de trabajo del *Building Evidence in Education* (2012) ha indicado, por ejemplo, que es necesario a lo menos dos ciclos educativos completos bajo intervención para poder esperar impacto de una determinada intervención con foco en estudiantes. En esta misma línea, distintos investigadores de evaluaciones sociales (King y Behrman, 2009) y de evaluaciones educativas (Borman et al., 2003) han indicado que, para evaluar el momento en que debe desarrollarse una evaluación de impacto, deben considerarse a los menos dos aspectos. Por una parte, es necesario estudiar las dificultades propias de cualquier implementación (por ejemplo, deficiencias de infraestructura, capacidades de recursos humanos, sinergias institucionales, etc.) y la forma en cómo estas dificultades

podrían impactar en el desarrollo de la teoría del cambio propuesta (Bamberger, Rugh y Mabry, 2012). Por otra parte, y especialmente en el caso de implementaciones progresivas (es decir, a las que se van sumando nuevos espacios o unidades en el tiempo), las dimensiones de escalabilidad de las iniciativas o los plazos de implementación son aspectos críticos que deben evaluarse para indicar cuando deberían realizarse las evaluaciones de cualquier programa (Coburn, 2003; Dede, Honan y Peters, 2005; Datnow, 2005). Para la NEP deben tomarse en cuenta estas dos dimensiones. Por una parte, es necesario considerar una evaluación que permita que los distintos actores involucrados (desde la recién creada Dirección Nacional hasta las escuelas) se adecue a nuevas dinámicas, objetivos y sentidos de su quehacer. De manera complementaria, es necesario considerar que la forma propuesta de escalabilidad (anual y progresiva) también presentará nuevos desafíos, especialmente en aquellos espacios más densamente poblados y/o en aquellos espacios donde las autoridades comunales tienen un trabajo más activo en el área de educación. Por todo lo anterior, se recomienda considerar que el horizonte temporal para la evaluación de impacto probablemente excede una o dos administraciones de gobierno.

- **Considerar posible efecto contagio:** Como se ha podido observar, los resultados del ejercicio de construcción del grupo de control y de tratamiento realizado pueden considerarse como buenos puntos de comparación para la evaluación de la Nueva Educación Pública. Sin embargo, las características intrínsecas de esta iniciativa (que sea una ley general, que esté considerada una implementación gradual en todo el país, que considere importantes recursos adicionales) hacen que, con una alta probabilidad, exista un efecto contagio (Sinclair, McConnell y Green, 2012). Esto quiere decir que, considerando la dinámica de la implementación de los Servicios Locales de Educación (SLE), alguna proporción de las comunas y establecimientos considerados como controles comiencen a desarrollar prácticas y/o estrategias imitativas respecto del funcionamiento de los establecimientos que están en NEP (en términos organizacionales, financieros y/o pedagógicos). Aunque esto puede ser evaluado positivamente en términos de la política pública, en términos de la evaluación *es necesario considerar y, en la medida de lo posible, corregir este aspecto en la evaluación intermedia y final del impacto*. Aunque la investigación sobre el efecto contagio y contaminación está en pleno desarrollo, ya existen algunas técnicas estadísticas (como, por ejemplo, el cálculo del *Complier Average Causal Effect* o CACE, que implica análisis de sensibilidad del efecto del tratamiento incorporando nuevos parámetros como el tiempo o intensidad del contagio) (Keogh-Brown et al., 2007) o técnicas de corrección de diseño en las evaluaciones finales, como la selección aleatoria de sub-muestras del grupo de control y de tratamiento (Angelucci y Maro, 2010), que permiten hacer frente a este aspecto. De esta manera, es necesario que las próximas evaluaciones realicen un análisis detallado sobre la probabilidad, naturaleza y proceso de contaminación que se podría generar.
- **Evaluación de procesos relevante y necesaria:** Las evaluaciones de procesos son relevantes y necesarias, particularmente cuando se busca medir el impacto de políticas y programas educativos, dado que el impacto es altamente dependiente de las condiciones de la implementación. Así, las evaluaciones de proceso entregan información valiosa sobre aspectos actitudinales, la implementación, las prácticas y las formas en que se desarrollan elementos críticos del sistema educativo que deberían tener influencia sobre el impacto de la NEP.

Por tal motivo, se recomienda que las próximas evaluaciones (intermedias y finales) también incorporen una evaluación de procesos (idealmente cualitativa) que permita analizar, entender y profundizar en los resultados del impacto encontrados, aportando así en el entendimiento del fenómeno a evaluar y aportando en posibles políticas educativas para su mejora. Recientes guías metodológicas sobre la evaluación en el campo educativo (Building Evidence in Education, 2012) han destacado la importancia de la información cualitativa, como un alimentador de los resultados de impacto y como una relevante guía para el mejor entendimiento de los avances y dificultades de cualquier intervención.

## 6.2 CONCLUSIONES Y RECOMENDACIONES RESPECTO DE LA IMPLEMENTACIÓN DE LA NUEVA EDUCACIÓN PÚBLICA EN EL CORTO Y MEDIANO PLAZO.

- **Gestión del cambio y relación con los aspectos actitudinales:** Uno de los resultados más interesantes de la evaluación de procesos, bastante en línea con otras evaluaciones de procesos (ver resumen de evaluaciones a este respecto desarrolladas en el informe, especialmente la evaluación CEPPE UC sobre funcionamiento de la subvención escolar) es que los aspectos actitudinales de los actores (percepciones, expectativas, juicios y prejuicios) son centrales para el éxito de la implementación de cualquier iniciativa, pues movilizan sentidos, prácticas y capacidades de acuerdo a estos aspectos. En el caso de la Nueva Educación Pública (NEP), estos aspectos son quizás aún más cruciales, pues el proyecto implica una serie de nuevos desafíos, por ejemplo, una nueva institucionalidad para el sector público, nuevos roles para los sostenedores y reconfiguraciones de los roles (históricos) que distintos actores educativos han jugado, que, en definitiva, implican una gestión del cambio que debe dirigirse y analizarse. Evidencia de investigación sobre este tópico (Levin y Datnow, 2012; Spillane, 2015) ha mostrado que el liderazgo en los distintos niveles (encargados nacionales, sostenedores, directores de escuela) es central para la gestión del cambio y transformación de las formas (muchas veces, naturalizadas) de funcionamiento del sistema escolar. Considerando estos elementos, *nos parece central que los encargados de implementar la Nueva Educación Pública aborden y potencien el trabajo en esta dimensión con todos los actores involucrados*, con un especial énfasis en los directivos y docentes, verdaderos engranajes de esta iniciativa.
- **Indicadores de resultados y de impacto como guía para el desarrollo de la política sobre educación pública:** Considerando que los procesos de implementación de cualquier política suelen estar asociados a momentos de incertidumbre, lo que es necesario reconocer y estudiar para el desarrollo de la evaluación educativa (Bando, 2013), *recomendamos considerar los resultados de los indicadores de resultado e impacto como una guía para la implementación de la política en el sector*. En el caso de la Nueva Educación Pública, la incertidumbre está relacionada con la creación de una nueva institucionalidad, lo que implica desafíos políticos, organizacionales, jurídicos y pedagógicos respecto del sentido, orientación y acciones de la Dirección Nacional de Educación Pública, los Servicios Locales de Educación, los Comités Directivo Locales y los Consejos Locales de Educación Pública. Considerando la reciente historia de implementación de nueva institucionalidad educativa en Chile (por ejemplo, la Agencia de la Calidad de la Educación, la Superintendencia de la Educación o la Subsecretaría de la Educación Parvularia) creemos que los tomadores de decisiones y la institucionalidad escolar posee un *know-how* relevante respecto de este aspecto,

que debe ponerse a disposición de la NEP. Para ello, se deben considerar los resultados para profundizar el trabajo que permita mejorar las dimensiones más débiles (p.e., la dimensión de apoyo pedagógico) pero también para avanzar en la transformación cualitativa de aquellos indicadores que presentaron altos niveles de cumplimiento (p.e., aquellos referidos a la participación formal o respecto de trayectorias regulares) considerando los avances y limitaciones de experiencias anteriores. Esto permitirá a los implementadores de la política utilizar esta línea de base para hacer avanzar la política sobre la educación pública, generando así mejoras en el conjunto de los establecimientos del sistema.

- **Análisis de externalidades positivas y negativas no anticipadas:** Un aspecto a considerar por cualquier intervención es la existencia y cuantificación de las posibles externalidades (Bernal y Peña, 2012). Aunque este objetivo es un aspecto que excedía los términos del presente estudio, es necesario considerarlo en las próximas evaluaciones. Ahora bien, considerando las dimensiones de la NEP (aplicación en todo el territorio nacional) y su profundidad en el espacio educativo (reconfiguración de la forma de entendimiento de lo público) es necesario analizar las externalidades que el proyecto podría generar. En este sentido, y considerando que la teoría del cambio indicó que los indicadores de impacto debían centrarse en aspectos de transformación del estudiante y las escuelas, *parece lógico pensar que la NEP también cambiará aspectos de los profesores (p.e., satisfacción laboral, niveles de burnout, formas de colaboración, etc), de los padres y apoderados (sentido de pertenencia a la escuela, nivel de compromiso escolar) o de los actores municipales (recursos disponibles, relación con la oferta comunal, entre otros) que posiblemente se verán transformados con la implementación de la Nueva Educación Pública.* Esto es relevante para dar cuenta de una mejor forma de las implicancias y transformaciones que podría estar generando la NEP en el conjunto del sistema educativo y en sus actores.