

# El pensamiento de diseño y la evaluación de políticas e intervenciones sociales: hacia un diálogo *realista*

Juan David Parra  Instituto de Estudios en Educación, Universidad del Norte, Colombia  
jparrad@uninorte.edu.co

Camilo Vieira  Instituto de Estudios en Educación, Universidad del Norte, Colombia  
cvieira@uninorte.edu.co

Doi: <https://doi.org/10.12804/revistas.urosario.edu.co/desafios/a.12199>

## RESUMEN

Si bien la literatura documenta un interés creciente en las aplicaciones del Pensamiento de Diseño (PDD) en la administración pública en general, existe una brecha en la producción académica sobre su papel en el campo específico de la evaluación de políticas e intervenciones sociales. Este artículo hace un aporte conceptual y metodológico en este último debate. Parte de la reflexión propuesta es de naturaleza meta-teórica, a partir de la cual se identifican puntos en común entre fundamentos epistemológicos del PDD y una escuela de evaluación en particular: la *evaluación pragmática*. En un segundo momento analítico del texto se examinan algunas limitaciones epistemológicas del pragmatismo (como paradigma de investigación) al momento de fundamentar conclusiones causales en el estudio de fenómenos sociales. Este punto no es trivial, en tanto, el análisis causal es la razón de ser de la evaluación de políticas o programas (públicos). Se propone que es posible enmendar algunos de estos problemas en la evaluación *basada en el diseño* a partir de la adopción de principios epistemológicos del Realismo Crítico. Las implicaciones metodológicas de estos argumentos son ilustradas con la discusión de dos ejemplos de evaluaciones de programas educativos en Colombia, lideradas por los autores.

**Palabras clave:** Pensamiento de Diseño; Administración Pública; Pragmatismo; Realismo Crítico; Evaluación Realista; Colombia.

**Para citar este artículo:** Parra, J. D., & Vieira, C. (2023). El pensamiento de diseño y la evaluación de políticas e intervenciones sociales: hacia un diálogo *realista*. *Desafíos*, 35(1), 1-31.  
<https://doi.org/10.12804/revistas.urosario.edu.co/desafios/a.12199>

## Design Thinking and The Evaluation of Social Interventions: Towards a *Realist* Dialogue

### ABSTRACT

While the literature reports increasing interest in the applications of Design Thinking (DT) in the field of public administration in general, a gap persists in academic reflections about the specific role of DT in informing policy and program evaluation. This article makes a conceptual and methodological contribution to this last debate. Part of the proposed reflection is meta-theoretical, allowing the identification of commonalities between the epistemological foundations of the DT and one evaluation school, in particular: *pragmatic evaluation*. In a second analytical moment, the text examines some epistemological limitations of pragmatism (as a research paradigm) to ground causal conclusions in social inquiry. This is not a trivial matter since the very essence of evaluation is, precisely, the identification of causal dynamics unleashed by social interventions. It is possible to amend some of these problems in a DT-based evaluation by adopting some epistemological principles from Critical Realism. A brief discussion of two applied evaluations of education interventions in Colombia led by the authors serves to illustrate some methodological implications of these last arguments.

**Keywords:** Design thinking; public administration; pragmatism; critical realism; realist evaluation; Colombia.

## O pensamento de design e a avaliação de políticas e intervenções sociais: rumo a um diálogo *realista*

### RESUMO

Embora a literatura documente um crescente interesse pelas aplicações do Pensamento de Design (PDD) na administração pública em geral, há uma lacuna na produção acadêmica sobre seu papel no campo específico da avaliação de políticas e intervenções sociais. Este artigo traz uma contribuição conceitual e metodológica a este último debate. Parte da reflexão proposta é de natureza metateórica, a partir da qual se identificam pontos comuns entre os fundamentos epistemológicos do PDD e uma escola de avaliação em particular: *a avaliação pragmática*. Em um segundo momento analítico do texto, examinam-se algumas limitações epistemológicas do pragmatismo (como paradigma de pesquisa) ao fundamentar conclusões causais no estudo dos fenômenos sociais. Esse ponto não é trivial, pois a análise causal é a razão de ser da avaliação de políticas ou programas (públicos). Propõe-se que é possível corrigir alguns desses problemas na avaliação *com base no design* a partir da adoção de princípios epistemológicos do Realismo Crítico. As implicações metodológicas desses argumentos são ilustradas a partir da discussão de dois exemplos de avaliações de programas educacionais na Colômbia, conduzidos pelos autores.

**Palavras-chave:** pensamento de design; administração pública; pragmatismo; realismo crítico; avaliação realista; Colombia.

## Introducción

Autores como Dorst (2011) invitan a entender el Pensamiento de Diseño (PdD) —o *Design Thinking*— como un paradigma —o una forma de pensamiento— al servicio de problemáticas que surgen en diferentes profesiones y disciplinas.<sup>1</sup> El desafío de lidiar con estos problemas abiertos y complejos, escribe, “lleva a un interés particular en las formas en que diseñadores crean marcos de referencia y la forma en que [diferentes] organizaciones [...] hacen uso de dichos marcos de referencia en su campo de acción” (Dorst, 2011, p. 522. Énfasis original).<sup>2</sup> Este ámbito de reflexión tiene orígenes en la ingeniería y en algunas ramas del diseño industrial, y se perfila en la actualidad como un área de estudios en torno a prácticas o procedimientos de razonamiento idóneos para generar conocimiento práctico en diferentes escenarios sociales (Cross, 2011; Dorst, 2011).

Quizás por ese enfoque prescriptivo, la literatura ha documentado un interés creciente en las aplicaciones del PdD en el campo de la administración pública (Hermus et al., 2020; Romme & Meijer, 2020). Parafraseando a Bason (2016), existe un nexo evidente entre el diseño y la gestión estatal, entendiendo que los Gobiernos son, en últimas, los dueños de los problemas públicos, y que las políticas o iniciativas gubernamentales son estrategias para intentar resolverlos. Debido a la complejidad implícita en este tipo de visiones —desde las cuales se entiende que hacer políticas públicas implica abordar intereses y expectativas cambiantes de una diversidad de actores sociales—, algunos autores resaltan la importancia de esquemas de trabajo colaborativo e iterativo entre beneficiarios, el sector privado y/o representantes de oficinas gubernamentales, etc., como requisito para contar con un servicio público eficaz y legítimo (en el sentido que responda a preferencias diversas) entre la ciudadanía (Bason & Austin, 2022; Mintrom & Luetjens, 2016).

La evaluación de intervenciones sociales corresponde a una fase del ciclo de políticas públicas que se incluiría dentro de este marco analítico de resolución de problemas (Lewis et al., 2020). Sin embargo, existe un espacio de reflexión inexplorado sobre cómo explicitar el diálogo entre la evaluación —entendida como un ejercicio sistemático para explicar el cómo y por qué detrás del éxito/fracaso de iniciativas gubernamentales— y el PdD (Olejniczak et al.,

1 Kees Dorst es considerado un referente contemporáneo de la teorización del PdD (Weedon, 2019).

2 Las citas de textos en otro idioma son traducciones libres de los autores.

2020; Adams & Nash, 2016; Villa Álvarez et al., 2022). En principio, la lógica inmersa en el razonamiento de diseñadores sociales coincide con el principio de la denominada evaluación pragmática de poner a prueba soluciones parciales para ir adaptando elementos de un programa o intervención según la retroalimentación de usuarios finales (Gray & Shaw, 2018). Sin embargo, como se profundizará más adelante, dicha lógica de razonamiento (que se denominará, de tipo abductivo) puede subestimar aspectos más estructurales de los entornos institucionales y culturales que condicionan la sostenibilidad política de diferentes propuestas de mejora del sector público.

El presente artículo se adentra en el debate sobre la complementariedad que puede existir entre los campos de estudio mencionados. En primer lugar, se exalta el aporte que puede hacer el PdD a formas de evaluación más tradicional (que se denominarán de corte deductivo e inductivo), invitando a una exploración más abductiva (o fundamentada en contrastar formas de entender una problemática) de la implementación de intervenciones sociales. Las implicaciones de dicha reflexión son ilustradas con un ejemplo de una evaluación aplicada (y dirigida por uno de los autores) a una estrategia auspiciada por el sector público para desarrollar habilidades de pensamiento computacional en docentes y estudiantes de instituciones educativas oficiales de Colombia. En este sentido, y ante la escasez relativa de este tipo de reflexiones en la literatura, este trabajo hace una primera contribución en el plano de la praxis sobre cómo aterrizar un diseño de evaluación y hacer uso de herramientas concretas para evaluar una intervención social de gran escala, pensando desde el diseño (abductivo).

En segundo lugar, se avanza en una crítica esbozada desde la filosofía del Realismo Crítico (RC) a los límites epistemológicos de la abducción al momento de teorizar sobre el cómo y/o el por qué, detrás de diferentes fenómenos sociales. Este es un debate fundamental en el campo de la evaluación, donde el reto intelectual (y práctico) consiste en estudiar comportamientos e interacciones humanas que emergen tras intervenciones en diversos contextos (Rigby et al., 2016). El punto central en dicha reflexión consiste en resaltar que un ejercicio de abducción (en la versión propuesta desde el pragmatismo) hace uso de diferentes teorías (ej. sobre el comportamiento humano) que, de ser ciertas, aportarían a resolver un problema de investigación. No obstante, desde el RC se considera que, si bien, dicho conocimiento es útil para pensar creativamente sobre problemáticas sociales, resulta insuficiente para esclarecer mecanismos causales específicos a partir de los cuales puede valorarse la idoneidad —o qué

tan factible es ser llevada a cabo con éxito— de una política o programa en un contexto político, social y/o económico determinado (Ritz, 2020). En línea con esta literatura, se propone que para resolver dicho problema es necesario explorar el campo del pensamiento retroductivo (Jagosh, 2020).

Para ilustrar las implicaciones en la práctica de esta segunda reflexión se discuten algunos elementos clave de una estrategia abductiva y retroductiva que se utilizó para llevar a cabo la evaluación de la implementación del programa Jornada Única —una iniciativa gubernamental para incrementar la permanencia escolar— en instituciones educativas públicas en Colombia. El diseño de investigación de dicho proyecto (el cual lideró uno de los autores de este artículo) se inspiró en los principios de la denominada Evaluación Realista (ER), a partir de lo cual se argumentará que este último enfoque es una opción plausible para ser presentado como una propuesta de evaluación basada en PDD que solventa (al menos en gran parte) los problemas epistemológicos de alternativas existentes.

La última sección sintetiza los aportes teóricos explorados en el artículo. Se hace también una mención al potencial de la institucionalización del PDD en el campo de la política pública a través de los denominados Laboratorios de Diseño de Política y los aportes que se pueden hacer desde estos espacios al fortalecimiento de las prácticas de monitoreo y evaluación del sector público.

## Hacia una evaluación más creativa e iterativa

### La abducción como principio básico del PDD y la evaluación

El texto de Dorst (2011), citado en la introducción, anuncia desde su título —el núcleo del PDD y sus aplicaciones— la intención de proponer una definición básica para un concepto que es, quizás por naturaleza, difuso. Esto último debido a la incidencia PDD en diferentes disciplinas y, por lo tanto, a la necesidad de su adaptabilidad a distintos lenguajes, teorías y posibles metodologías de trabajo. Señala el autor, como punto de partida, que los niveles de complejidad de los problemas por resolver son un elemento clave al momento de reflexionar sobre el fundamento epistemológico del acto de pensar como lo haría un(a) diseñador(a). Al respecto, vale la pena citar un fragmento del artículo mencionado cuya riqueza conceptual y descriptiva permiten identificar elementos clave para la presente reflexión:

Existen varias formas de incorporar prácticas de diseño para enfrentar [una] situación problemática. (1) Las organizaciones a menudo [...] parten de forma convencional de resolución de problemas (Abducción-1) para crear un nuevo *algo* que salve el día, manteniendo constantes el *cómo*, el *marco de referencia* y el *valor* [...] A menudo, el problema, tal y como se presenta, necesita ser primero *deconstruido* [...] antes de que pueda convertirse susceptible de solución. Entonces procedemos a (2): si el enfoque Abducción-1 de crear un nuevo *qué* no ayuda, es posible que la organización deba entrar en modo Abducción-2, que también les obliga a crear un nuevo *cómo*. La organización podría hacer esto simplemente aplicando uno de los otros marcos de referencia que ya tiene en su repertorio, en su *campo*. (3) Alternativamente, podría contratar un consultor externo que utiliza su experiencia para traer un nuevo marco de referencia a la situación problemática. (Dorst, 2011, p. 530. Énfasis original)

Según puede interpretarse de esta cita, tanto los modos de Abducción-1 y de Abducción-2 representan procesos de razonamiento que toman distancia de aproximaciones comunes de resolución de problemas (organizacionales) —ej. de implementación de estándares a partir de nociones de buenas prácticas.<sup>3</sup> En criterio del autor, el PdD se ve más claramente reflejado, sin embargo, en la segunda de estas categorías (la Abducción-2), la cual invita a trascender estrategias de toma de decisiones que presumen de entrada la existencia de un objetivo claro por alcanzar —o una solución óptima (Dorst, 2006). Es decir, se hace un llamado a reconocer que la definición misma de un problema puede estar determinada por el uso de marcos de referencia específicos para leer una situación (lo que en la literatura especializada se entiende como *framing*), de lo cual se deriva la posibilidad de que al utilizar otros marcos de referencia se modifique también el motivo y el foco de búsqueda de una solución. Este principio permite entender el porqué del énfasis de diseñadores en asuntos públicos en la importancia de la deliberación democrática para ampliar perspectivas (o marcos de referencia) y, por esa vía, buscar acrecentar la legitimidad de los servicios y las intervenciones del Estado (Hermus et al., 2020; Mintrom & Luetjens, 2016).

3 Piénsese, por ejemplo, en los llamados a construir repositorios de evidencias científicas (sobre medidas estandarizadas de política que funcionan), comunes en discursos dominantes sobre la política basada en evidencia (Parra, 2021).

Como sugiere la misma reflexión de Dorst (2011), la fundamentación epistemológica de esta lógica de resolución de problemas sienta sus bases un modo de pensamiento denominado abducción. Los fines teóricos y metodológicos del presente artículo motivan una breve exploración de este concepto y su distinción de otras formas de inferencia clásicas como lo son la deducción y la inducción. La conceptualización del pensamiento abductivo se asocia a la escuela pragmática norteamericana, y en particular al trabajo del filósofo Charles Peirce (Danermark et al., 2019), presentándolo como una forma de razonamiento a través del cual las personas producen generalizaciones para explicar fenómenos sociales haciendo uso de diferentes marcos de referencia (Tavory & Timmermans, 2014). Estos últimos emergen, por ejemplo, de experiencias, ideas y creencias personales y colectivas. Ello se diferencia de un proceso netamente deductivo, en el cual existe un marco de referencia predefinido (ej. una gran teoría) a partir del cual se explica de antemano (o se predice) un resultado, y de la inducción pura, donde se teoriza directamente desde los datos sobre tendencias o diferencias empíricas percibidas por el observador (Wuisman, 2005; Sayer, 2000).

Desde la filosofía del conocimiento se ha propuesto que la abducción es necesaria para resolver problemas complejos, para los cuales la deducción y la inducción (puras) no se encuentran suficientemente equipadas (Douven, 2021). El argumento dicta que para generar nuevas teorías (que son, en sí, las soluciones que se busca proponer)<sup>4</sup> es necesario trascender marcos de referencia predeterminados (o las soluciones fijas *a priori* que propone la deducción), sin por ello desconocer el valor de examinar previamente marcos de referencia existentes (contrario a lo que sugiere la base epistemológica de la inducción clásica)<sup>5</sup> (Tavory & Timmermans, 2014). La alternativa consiste en una especie de punto intermedio entre ambas posturas, inspirado en una lógica de proponer posibles teorías (o explicaciones provisionales) sobre cómo solucionar un problema, para luego recolectar evidencia que permita reflexionar y perfeccionar propuestas de acción.

4 Como lo recuerda el sociólogo Ray Pawson (2003), toda acción de una persona en el mundo es una teoría de lo que esa persona cree que va a generar un resultado. Bajo ese argumento —y ante el hecho inminente de que las personas se equivocan— toda la solución a un problema es realmente una teoría de los elementos y lógicas que podrían solucionarlo.

5 Esto bajo la premisa de que la rigurosidad e imparcialidad del trabajo en campo requiere que investigadores se abstraigan de sus preconcepciones (o prejuicios), de modo que no impongan sus sesgos al recoger información.

La ciencia del diseño llega a esta misma conclusión (sobre el papel epistemológico de la abducción) a partir del examen detallado del PdD aplicado. Así lo refleja la reflexión plasmada en influyente artículo por Dorst y Cross (2001), en el cual analizan cómo un grupo de ingenieros proponen soluciones para introducir sistemas de disposición de basuras en trenes de pasajeros en Holanda, y que corrobora el diálogo natural que existe en el campo del diseño y la filosofía pragmática (Gray & Shaw, 2018). En palabras de los autores:

Parece que el diseño creativo no es una cuestión de tener primero una *interpretación fija* del problema y luego buscar un concepto de solución satisfactorio. El diseño creativo parece ser más una cuestión de *desarrollar y refinar juntos* tanto la formulación de un problema como las ideas para una solución (Dorst & Cross, 2001, p. 434. Énfasis agregado).

Ahora bien, pensando en la evaluación de intervenciones sociales —que es el foco de reflexión de este artículo—, el debate sobre diferentes formas de inferencia permite identificar una primera coincidencia entre los dilemas metodológicos de ese campo disciplinario y las virtudes del PdD para ayudar a resolverlos. Por ejemplo, los métodos (cuasi)experimentales —asociados al paradigma de la evaluación de impacto cuantitativa— tienden a reproducir una lógica deductivista de inferencia, al reducir las explicaciones sobre efectos estadísticos de un programa a axiomas racionalistas, provenientes de la economía neoclásica (Parra, 2021). Por su parte, planteamientos metodológicos que privilegian diseños más cualitativos, como la evaluación constructivista o la denominada teoría fundamentada, tienden a reducir las explicaciones del cómo y el porqué de los resultados de una evaluación a interpretaciones inductivas que hacen investigadores sobre diferentes fuentes de información empírica (Tavory & Timmermans, 2014; Pawson & Tilley, 1997).

Otros autores han propuesto que, sin embargo, el pensamiento abductivo tiene también un papel crucial en el campo de la evaluación de intervenciones sociales para solventar los problemas epistemológicos de la deducción y la inducción. Dicho principio se ve reflejado en los fundamentos de la denominada evaluación pragmática, o evaluación durante el desarrollo (Patton, 2010), al punto que se ha propuesto que esta última puede ser catalogada como un ejercicio de PdD aplicado (Gray & Shaw, 2018; Adams & Nash, 2016). En el siguiente fragmento Gray y Shaw (2018) resumen este argumento:

La evaluación durante desarrollo se basa en la premisa de que lograr activamente el resultado es más importante que la certeza epistémica percibida que surge de los enfoques de evaluación convencionales [...] y que las medidas de resultado específicas pueden desarrollarse y adaptarse a lo largo del proceso de intervención (Patton, 2006, 2010) [...] La evaluación durante desarrollo resuena fuertemente con los enfoques orientados al diseño para la implementación y la evaluación. El diseño es fundamentalmente pragmático, orientado a realizar cambios en un servicio o en artefactos físicos, que interactúan con la experiencia de las personas de formas novedosas (Steen, 2013). El diseño de servicios es holístico, centrado en el ser humano, iterativo y reconoce la influencia de los ecosistemas en los que se anidan los servicios. (p. 244)

Dicho lo anterior, cabe advertir que, no obstante, una crítica hecha a la evaluación pragmática es que carece de una propuesta metodológica o de un marco conceptual y/o procedimental claro para guiar prácticas idóneas de recolección y análisis de información (Astbury, 2013; Pawson, 2013; Pawson & Tilley, 1997). Ello implica, por tanto, que sería importante explicitar estrategias por medio de las cuales el PdD —no solo como una forma de razonamiento, sino también como un campo de estudios que ha acumulado conocimiento práctico para pensar abductivamente (en el sentido del modo Abducción-2) sobre la resolución de problemas complejos— puede aportar a la literatura teórica y práctica de la evaluación de intervenciones sociales. Esto, además, teniendo en cuenta que el diálogo epistemológico entre el PdD y de evaluación es uno que hasta ahora se está empezando a dar (Gray & Shaw, 2018; Adams & Nash, 2016). La discusión que se presenta a continuación sobre la evaluación de un programa educativo a gran escala en Colombia da pasos en esa dirección.

### Una estrategia de evaluación basada en diseño: la evaluación del proyecto CFK

El proyecto Programación para Niños y Niñas (*Coding for Kids*, CFK) es una iniciativa del Ministerio de Tecnologías de la Información y Comunicaciones y del Consejo Británico en Colombia, entre otros actores, para formar a docentes de instituciones educativas públicas en la enseñanza del pensamiento computacional (Vieira et al., 2022). Este campo disciplinar comprende un conjunto de conceptos y habilidades derivadas de las ciencias de la computación que pueden ser utilizadas por profesionales de todas las áreas para representar ideas y fenómenos, o resolver problemas complejos. El programa se planteó,

además, con un enfoque de género para promover el interés de niñas y adolescentes mujeres en las áreas STEM (en inglés, Ciencia, Tecnología, Ingeniería y Matemáticas).

La implementación de CFK inició en el primer semestre del 2019 y se puso como meta la formación de 288 docentes (para ese año). El diseño original del programa tuvo como eje central la creación de unas fichas didácticas basadas en teorías educativas sobre el aprendizaje de la programación (ej. teorías de cargas cognitivas) en educación básica y media. En tal sentido, la primera teoría de cambio de la intervención incluía un taller práctico de formación inicial a educadores para el manejo de la ficha guía en clase y la entrega de unos dispositivos tecnológicos llamados *Micro: bit* para ser incorporados en sus actividades didácticas. Un supuesto implícito en la operación de la iniciativa indicaba que la posibilidad para beneficiarios de recibir retroalimentación en sus primeras experiencias de enseñanza de pensamiento computacional era fundamental para el éxito de la estrategia, por lo cual se planearon además cinco visitas para realizar ejercicios de observación en aula.

La evaluación de CFK se fundamentó en la denominada Investigación Basada en Diseño (IBD), la cual puede entenderse como una aplicación del PdD inspirada en principios similares de la evaluación pragmática (Hoadley & Campos, 2022; Schön, 1992). Es decir, parte del reconocimiento de la complejidad de los entornos de aprendizaje y, por tanto, problematiza la utilidad de hallazgos que derivan de estudios psicológicos experimentales (es decir, deductivistas) para informar decisiones encaminadas a la innovación pedagógica en ambientes no controlados (Barab & Squire, 2004). En línea con el planteamiento epistemológico abductivo básico del PdD, la IBD hace uso de procesos iterativos y flexibles de diseño y de implementación y evaluación en contextos reales, involucrando a los participantes para comprender y refinar el funcionamiento de una intervención educativa.

Durante la fase de diseño del estudio, el equipo consultor propuso un marco de referencia propio (y fundamentado en la experiencia académica de sus miembros) para enmarcar la evaluación. La teoría *TPACK* sugiere que la docencia efectiva con tecnología depende de una conjugación de los diferentes saberes que hacen parte del acrónimo; es decir, el conocimiento Tecnológico, Pedagógico y del Contenido de la disciplina en la que se incorporan estos

elementos (Mishra & Koehler, 2006). Esto implicaba la importancia de centrar parte del esfuerzo del trabajo de campo en recolectar información para examinar tres conocimientos específicos en docentes beneficiarios: i) conceptos básicos de pensamiento computacional, ii) estrategias pedagógicas para aplicación de conceptos en aula y iii) su solvencia en el manejo de recursos tecnológicos.

Los resultados de la evaluación del primer año de implementación mostraron una mejora en la comprensión global de educadores sobre el conocimiento de los elementos TPACK para la integración del pensamiento computacional en distintas áreas disciplinares. También se documentó un avance en el aprendizaje de estudiantes participantes al comparar su desempeño en una prueba de conocimientos con el de un grupo control proveniente de instituciones no beneficiarios de las actividades de formación (Espinal et al., 2022; Vieira et al., 2022). Sin embargo, se identificaron elementos por mejorar, como la dificultad de los docentes con el desarrollo de ciertas habilidades técnicas y la necesidad de diseñar estrategias adicionales para el cierre de brechas de género. Los hallazgos también apuntaban a la importancia de revisar algunos materiales de estudio, ya que los contenidos parecían no ser idóneos para el nivel educativo en el que algunos de los docentes participantes enseñaban (educación primaria).

En 2020, tras el inicio del confinamiento por la pandemia COVID-19, fue necesario rediseñar algunos elementos de la estrategia (ej. materiales) y llevarla a una modalidad de formación virtual. Ello también permitió ampliar la población beneficiaria a 8.500 docentes formados en un curso de nivel inicial y a otros 4.250 en un curso de nivel avanzado. Toda esta transformación implicó una revisión a la teoría de cambio y un proceso de *framing* (Schön, 1992), en el que evaluadores y dueños del programa redefinieron algunos problemas por resolver. Específicamente, el Consejo Británico retomó los resultados alrededor de las tres dimensiones de TPACK de la evaluación de 2019 para hacer mayor énfasis en contenidos (ej. actividades prácticas sobre álgebra booleana) y en estrategias pedagógicas específicas (ej. progresión Usa-Modifica-Crea para la integración de ejemplos) dentro del acompañamiento de mentores en la formación en línea. De igual forma, se decidió integrar tanto al equipo consultor técnico como al de evaluación expertas en género que contribuyeran en aterrizar conceptos clave (ej. como la diferencia entre equidad e igualdad), y

en definir acciones afirmativas diferenciales, en busca de fomentar más activamente la participación de estudiantes mujeres en espacios de ciencias y computación.

Dado ese mismo contexto de confinamiento y cierre de instituciones educativas, la evaluación en el segundo año solo se centró en la formación docente. A pesar de que también se identificaron algunos resultados positivos en conocimientos adquiridos en esta fase, el avance era apenas marginal (ej. muy pocos llegaban a niveles altos de ‘experticia’), denotando la necesidad de una mayor retroalimentación de su práctica en aula. La lectura que se dio a este hallazgo por parte de ejecutores de CFK y evaluadores era que existía un riesgo de que, a falta de un mecanismo que permitiese a educadores sentirse apoyados en la implementación de las estrategias que habían aprendido, el programa fuese poco sostenible en el tiempo —es decir, que se abandonara la iniciativa una vez se diera por terminado el acompañamiento externo en instituciones educativas—. Los resultados del análisis sugerían, a su vez, un cumplimiento parcial de metas asociadas a cerrar brechas de género en el fortalecimiento de habilidades STEM. En el lenguaje del PDD, ello implicaba la necesidad de buscar nuevos marcos de referencia, a fin de identificar soluciones oportunas y efectivas a estas dificultades.

Una respuesta del equipo consultor fue acudir nuevamente a la literatura para examinar y proponer nuevos ajustes al diseño de CFK. Para abordar el primer reto surgió la idea de crear una comunidad de práctica para el intercambio de experiencias y solución de dudas, a través de interacciones entre educadores novatos y aquellos con un grado de experticia acumulado en la enseñanza del pensamiento computacional. Con base en referentes como Vásquez (2011), se tomó entonces la determinación, por parte del Consejo Británico, de construir dicho espacio haciendo uso de una red social (*Facebook*) ampliamente usada y reconocida por beneficiarios del programa, y que fuese además de fácil acceso (dado los paquetes de planes y aplicaciones web ofrecidos por empresas de telefonía móvil) entre todos los participantes.

Frente al segundo reto, se puso sobre la mesa (y en diálogo entre el equipo consultor y el equipo interventor) la experiencia de una aplicación móvil (*UK-Bebras*), previamente utilizada en Inglaterra para trabajar en habilidades precursoras al pensamiento computacional. El fundamento conceptual en este caso provino, entre otras fuentes especializadas, de La Organización de

las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO), en un informe dedicado específicamente al reto de generar estrategias de aprendizaje STEM enfocadas en reducir brechas de género (UNESCO, 2017). Como resultado de ese ejercicio se propuso el diseño de la aplicación *GreenTIC*, bajo el supuesto de que la generación de espacios donde estudiantes pudieran tener experiencias interactivas enmarcadas en ecosistemas locales y en problemas con comunidades y minorías étnicas, contribuiría a que las niñas (en particular) se vieran más interesadas en aprender conceptos y aplicaciones básicas de la programación.

En suma, y desde la discusión teórica del presente artículo, la evaluación de CFK es una experiencia significativa del PdD en acción, aplicado a la evaluación de intervenciones sociales a gran escala. En la descripción de las diferentes decisiones que se iban tomando en la revisión y el ajuste periódico de la teoría de cambio del programa se hizo evidente la puesta en marcha de un proceso abductivo (concordante con la Abducción-2) de contrastar marcos de referencia (muchos aportados desde afuera) para re enmarcar preguntas de investigación y definir, con base en ello, ajustes a estrategias y metas de la intervención. Cabe mencionar que dichas iteraciones analíticas (propias de la IBD) se dieron en medio de un escenario de un trabajo escalonado de implementación, evaluación, ajuste y nuevamente implementación, posible gracias a cierta apertura de la entidad financiadora del estudio de forjar una alianza de mediano plazo con los consultores externos. Este escenario no es tan común.

El proceso de evaluación de CFK sigue en curso, siempre abierto a incorporar enseñanzas derivadas del estudio cercano de su implementación y del aporte de marcos de referencia adicionales para abordar nuevos problemas emergentes en su diseño. Por ejemplo, el seguimiento a la interacción de educadores y estudiantes con la aplicación *GreenTIC* ha mostrado que esta ha tenido, en efecto, un recibimiento positivo, especialmente entre estudiantes mujeres. Sin embargo, surgen nuevas demandas por mayor acompañamiento para que sea utilizada con los fines pedagógicos esperados. Las dificultades económicas que experimentan algunas familias para financiar factores de soporte al uso de dicha tecnología (ej. el acceso a un dispositivo móvil inteligente) se encuentra también en la lista de nuevos factores/problemas a solucionar.

## Hacia una evaluación basada en diseño (realista)

Llegado a este punto de la discusión, el mensaje es concreto: dado su fundamento en el pensamiento abductivo (pragmático), y en medio de la acogida de los principios del pragmatismo por parte de algunos teóricos de la evaluación de intervenciones sociales, se propuso que el PdD, desde su planteamiento abductivo, puede hacer una contribución metodológica directa en este segundo campo disciplinar. La discusión del primer ejemplo de una evaluación aplicada —el programa CFK— ilustró, por tanto, una posible ruta metodológica para llevar a cabo el análisis informado por la lógica de razonamiento del diseño, orientada a buscar soluciones a problemas sociales complejos.

Como se anticipó en la introducción, en esta segunda parte del artículo se propone hacer un salto argumentativo, esta vez orientado a examinar las limitaciones del pragmatismo para informar, de forma decidida, la deliberación frente a acciones (o búsqueda de soluciones a problemas sociales) de la administración pública. Esto, en línea con críticas al pensamiento abductivo esbozadas desde la literatura y los principios del RC. Este segundo es un paradigma de investigación que cuenta en la actualidad con una influencia significativa en el desarrollo teórico y conceptual de disciplinas como la sociología y la ciencia política (Danermark et al., 2019; Porpora, 2015), y en campos de investigación aplicada como, precisamente, la evaluación de intervenciones sociales (Fontaine, 2020; Brousselle & Buregeya, 2018; Parra, 2021; Jagosh, 2020; Pawson, 2013).

En particular, se pone en consideración la visión realista sobre las limitaciones intrínsecas del paradigma pragmático al momento de resolver interrogantes causales en el estudio de fenómenos sociales (Pawson & Tilley, 1997; Ritz, 2020). Este punto no es trivial; si se parte del principio de que evaluar equivale a buscar evidencia sobre relaciones causa-efecto (ej. entre intervenciones educativas e indicadores de desempeño escolar), es ciertamente deseable que una estrategia de evaluación esté expresamente fundamentada en argumentos epistemológicos que permitan resolver interrogantes causales (Parra, 2021).<sup>6</sup> Y dado que, como lo ilustrará el ejemplo al final de la sección,

6 Esta aseveración tiene una carga epistemológica importante, en tanto existen diferentes posturas en la literatura sobre los objetivos de la evaluación (ej. la evaluación constructivista se distancia de hacer ciertos análisis causales). No obstante, Brousselle y Buregeya (2018) documentan un acuerdo emergente

desde el paradigma realista ya existe una propuesta explícita de evaluación —la denominada Evaluación Realista (ER)—, en lugar de avanzar con un desarrollo propio de una forma específica de evaluar —como se buscó a través en el primer ejemplo discutido—, la intención en este caso consiste en proponer que la ER representa, en sí misma, una aplicación del PdD (con algunas adendas) en el contexto de la evaluación de intervenciones sociales.

### Realismo Crítico y PdD: un diálogo desde (y más allá) de la abducción

Para entablar un diálogo entre el PdD y el RC es importante aclarar que, similar al primero, el segundo es un paradigma que comprende una serie de principios ontológicos para informar diseños de investigación. Es decir, pese a que sus proponentes buscan incidir en debates metodológicos, el RC no debe ser equiparado con una metodología o método específico de recolección o análisis de información (Sayer, 2000). El origen de dicha tradición meta-teórica está asociado al trabajo del filósofo Roy Bhaskar, planteándose como alternativa a paradigmas positivistas y constructivistas dominantes en el estudio de la sociedad (Bhaskar & Hartwig, 2010). Lo anterior, quizás debido al reconocimiento de las bondades del paradigma realista para dar respuesta a dilemas clásicos, y que tienden a suscitar controversias significativas, en las ciencias sociales (ej. objetividad vs subjetividad, dualismo vs monismo, estructura vs agencia, materialismo vs idealismo) (Parra, 2016; Sayer, 2000).

En el plano epistemológico —es decir, frente a la reflexión sobre cómo estudiar y lograr comprender fenómenos sociales— es relevante iniciar mencionando que el RC comparte la visión pragmática de problematizar lógicas netamente inductivas y deductivas subyacentes en diseños de investigación (Danermark et al., 2019; Jagosh, 2020; Maxwell, 2012; Parra, 2016; Sayer, 2000; Wuisman, 2005). Lo anterior se ve reflejado en planteamientos típicos de metodólogos realistas, como el siguiente:

Un objetivo general en la investigación de las ciencias sociales es explicar eventos y procesos. Explicar algo implica (desde la perspectiva del realismo crítico) primero describir y conceptualizar las propiedades y los mecanismos causales que generan y permiten eventos, haciendo que las cosas

---

entre expertos frente a lo que denominan una nueva generación de la evaluación orientada, necesariamente, a la explicación causal de fenómenos evaluados a partir de técnicas de investigación no experimentales.

sucedan [...] y luego describir cómo los diferentes mecanismos se manifiestan bajo condiciones específicas. Este tipo de investigación requiere un enfoque metodológico basado en la *abducción* y en la retroducción, rompiendo con el llamado modelo [deductivo] Popper-Hempel de explicaciones científicas. (Danermark et al., 2019, p. 97. Énfasis agregado)

Como lo delinear Danermark et al. (2019) en su cita, similar al espíritu del PdD, las metodologías de investigación permeadas por principios del RC se suscriben a la lógica abductiva de exploración de posibles teorías para explicar potencialmente un fenómeno observable previo (o como prerrequisito) a la construcción misma del argumento explicativo. Similar al caso del tipo de Abducción-2 (discutido más atrás), se propone, por ende, ir más allá de las técnicas de resolución de problemas más tradicionales (ej. pensar en posibles respuestas a partir de un único marco de referencia), para dar peso epistemológico a la lógica de explorar ideas y percepciones sobre una problemática dada a partir de ejercicios de re contextualización y reinterpretación de preguntas iniciales de investigación (Jagosh, 2020; Maxwell, 2012; Wuisman, 2005).

Pese lo anterior, diferentes autores han puesto también de manifiesto diferencias ontológicas significativas entre la filosofía pragmática —la cual, valga reiterarlo, converge con los fundamentos del PdD— y una visión realista de la investigación (Elder-Vass, 2022; Pawson & Tilley, 1997; Ritz, 2020; Zotzmann et al., 2022). Parafraseando a Elder-Vass (2022), quizás uno de los contrastes más significativos entre ambas tradiciones está en que los teóricos pragmáticos suelen hacer caso omiso al papel de estructuras sociales externas a (o que operan y existen más allá de) los individuos en la generación de explicaciones causales (p. 262). Esto se traduce, por ejemplo, en una tendencia a encontrar en la investigación pragmática un enfoque demasiado supeditado al nivel micro —o centrado exclusivamente en el estudio de casos específicos—, a expensas de un análisis de factores más macroestructurales, o que den cuenta del surgimiento mismo de dicho fenómeno en un contexto determinado y su interacción con elementos de dicho contexto.

Volviendo al debate sobre la idoneidad de estrategias de investigación pragmática —es decir, aquellas delineadas para resolver problemas a partir del contraste de diferentes marcos de referencia—, lo anterior implicaría entablar una crítica a los límites epistemológicos del pensamiento abductivo. En concreto, desde el RC es posible sostener que, si bien un ejercicio de reflexión

orientado por la abducción es útil para pensar creativamente sobre una problemática determinada (por ejemplo, la orientación que debería tomar el diseño de una política educativa), se cae en el riesgo de omitir la consideración del poder que tienen las estructuras sociales preexistentes (ej. distribución de recursos, preferencias sociales y/o nociones culturales imperantes sobre el significado de lo que implica educar) de incidir en qué tan factible es que la solución propuesta sea aplicable en un contexto socio-político determinado (Archer, 2013b). Tal vez de ahí se derive el dictamen de Peters (2015) señalando que “[los] mejores diseñadores serán incapaces de producir políticas efectivas si los factores fundamentales dentro de sus acuerdos [y/o relaciones] de gobernanza no conducen al éxito” (p. 273).

Es interesante anotar que, según algunos autores, este tipo de reflexiones (epistemológicas) no son del todo ajenas en el debate sobre el PdD aplicado a la administración pública (Carlsson, 2010; Di Russo, 2016; Fontaine, 2020). De hecho, para Di Russo (2016), la ontología del RC puede completar, e incluso potenciar, el poder analítico del PdD. Para ella, es importante que el enfoque de la literatura de esta segunda tradición trascienda del énfasis en la documentación de casos y experiencias de diseñadores, para “profundizar sobre el proceso del diseño por medio de métodos prácticos y estratégicos, de modo que las ideas innovadoras *puedan ser realizadas e implementadas con éxito*” (p. 275. Énfasis agregado). En su trabajo hace referencia específica al papel del razonamiento realista en responder a este objetivo metodológico, que, parafraseando la autora, implica hacer un llamado a que los ejercicios de diseño creativo estén siempre acompañados de una ‘verificación de la realidad’.

### El papel de la retroducción en el estudio, y en la evaluación, de ‘la realidad’

En el mismo fragmento citado al comienzo de la última subsección, Danermark et al. (2019) mencionan un segundo concepto clave para explorar problemáticas sociales complejas; el análisis retroductivo. En el argot del RC, la retroducción hace alusión a “[un] modo de inferencia en el que los eventos [observados] se explican postulando (e identificando) mecanismos que son capaces de producirlos” (Sayer, 1992, p. 107). Puesto en esos términos, el pensamiento retroductivo busca depurar posibles explicaciones que surgen de un ejercicio de razonamiento abductivo hasta identificar, dada la evidencia existente, la mejor explicación posible a un evento social determinado (Jagosh, 2020; Wuisman, 2005).

En el plano epistemológico, el razonamiento retroductivo se propone responder preguntas transfactuales del tipo *¿qué propiedades debe tener el objetivo o fenómeno X para que su existencia, en un contexto determinado, sea posible?* (Bhaskar & Hartwig, 2010; Danermark et al., 2019). En línea con la discusión de más arriba, la búsqueda de respuestas a este tipo de interrogantes no tiene que abandonar el espíritu reflexivo e iterativo de la abducción —en tanto la lógica de indagación realista hace también un llamado a la exploración de posibles teorías, informadas por diferentes marcos de referencia (Wuisman, 2005)—. Sin embargo, en este caso existe un énfasis explícito en comprender factores generadores, o posibilitadores, del objeto social estudiado. Por esta razón, como señala Ritz (2020), mientras que “la conclusión abductiva tiene el carácter de sugerir una hipótesis [que] no es necesariamente verdadera o falsa [...] la conclusión retroductiva [apunta a] qué *debe* ser cierto para que X exista” (p. 462).

La siguiente pregunta es: *¿cómo se aterriza el análisis retroductivo en un diseño de investigación?* La primera pista metodológica para avanzar en esta reflexión se encuentra ya resuelta en mismo planteamiento abductivo. Es decir, similar al procedimiento propuesto desde el PdD de contrastar marcos de referencia relevantes (ej. a partir de indicaciones de la literatura o del conocimiento experiencial de actores involucrados en el proceso de diseño), un primer paso en un ejercicio de retroducción consiste en el planteamiento de explicaciones causales potenciales. La diferencia, sin embargo, es que dicho proceso debe estar orientado no solo por una reflexión sobre diferentes formas de pensar (en el abstracto) en un problema, sino que debe hacer alusión explícita, al menos en parte, al entorno político y social más amplio que incide en la existencia misma (y/o los resultados que se espera observar tras la incidencia) del fenómeno social analizado. Margaret Archer, considerada como precursora en la traducción de la ontología realista al lenguaje de la metodología para investigación social, plantea el siguiente principio general para adelantar ese tipo de discusiones en el marco de una indagación sociológica:

Parece haber un acuerdo general en que para que cualquier proceso merezca ser considerado como generador de cambio social debe necesariamente incorporar relaciones humanas estructuradas (contexto-dependencia), acciones humanas (actividad-dependencia) e ideas humanas (concepto-dependencia). Necesariamente, los tres hacen que la teorización social no sea naturalista [en tanto] la vida social viene en un SAC [o una combinación de Estructura, Agencia y Cultura], siempre y en todas partes. (Archer, 2013a, pp. 4-5)

A partir del planteamiento de esta autora, es posible entonces sostener que toda teoría realista candidata a ser explorada en un ejercicio de indagación retroductivo debe, como mínimo, postular explicaciones sobre posibles interacciones entre estructuras sociales (que pueden ser materiales o culturales) y su incidencia en las acciones de seres humanos en un contexto determinado. Con ello, queda explícita la intención no solo de explorar marcos de referencia que permitan pensar en formas alternativas de aproximarse al fenómeno social, sino que, a su vez, anclen dicha reflexión en una consideración explícita de factores macroestructurales que ayuden a entender la existencia misma del fenómeno explorado. Lo anterior, haciendo alusión a uno de los principios epistemológicos básicos del RC de reconocer en todo intento de explicación causal la existencia de fuerzas estructurales externas a los individuos que, si bien no determinan del todo su comportamiento, sí inciden en la definición de sus posibles marcos de acción (Archer, 2013b).<sup>7</sup>

Llegado este punto, el carácter explicativo (o de búsqueda de explicaciones causales) del análisis retroductivo permite dilucidar una conexión evidente entre las implicaciones metodológicas de un diseño de investigación realista (ej. frente al estudio de relaciones causa-efecto en la sociedad) y el campo de la evaluación de intervenciones sociales. De hecho, en años recientes se ha venido consolidando una tradición denominada Evaluación Realista (ER), presentada por sus múltiples exponentes como una alternativa de escuelas dominantes de la evaluación —incluyendo en este repertorio a la denominada evaluación pragmática— (Parra, 2021; Astbury, 2013; Pawson & Tilley, 1997; Jagosh, 2020) y que se fundamenta en la puesta en marcha de estrategias de indagación informadas por los principios epistemológicos de la retroducción. Así lo releva el siguiente pasaje escrito por uno de los representantes más influyentes de esta tradición:

Las evaluaciones realistas a menudo comienzan con una teoría del programa, es decir, una explicación de cómo, por qué, para quién, en qué contextos y en qué medida se supone que una intervención *funciona* [...] La evaluación realista es una forma de enfoque de evaluación basado en teoría —es decir,

7 Archer (2013b) imprime este razonamiento en su tesis sobre el origen de sistemas educativos europeos: “Los actores [sociales] reaccionan a las situaciones en las que se encuentran; pueden permanecer inconscientes de los factores que moldearon tales situaciones [...] Estos contextos socioeducativos deben investigarse de manera independiente y, al hacerlo, la tarea sociológica no es solo registrar cómo eran vistos por las personas [...] sino también conceptualizar cómo este contexto más amplio estructuró la situación real en la que cada grupo se encontraba” (p. 55).

utiliza datos primarios [...] para confirmar, refutar y refinar la teoría o teorías realistas sobre el fenómeno de interés [...]. Para explicar los resultados (intencionados y no deseados) observados dentro de cualquier [intervención], la evaluación y la revisión realistas tienen un enfoque particular de análisis, capturado sucintamente en el heurístico contexto + mecanismo = resultado (o  $C+M=O$ ).<sup>8</sup> Esta heurística establece una explicación realista de los vínculos entre el contexto y el resultado, a través de mecanismos [...]. En pocas palabras, un mecanismo es una fuerza causal que hace que suceda un resultado. (Wong, 2018, p. 109. *Énfasis original*)

Como se anunció previamente, el ejemplo que se discute a continuación busca ilustrar la puesta en práctica de la combinación de principios abductivos y retroductivos, discutidos en esta segunda parte del artículo. Se reitera que la intención va más allá de mostrar posibles implicaciones de la puesta en marcha de principios del PDD en un ejercicio de evaluación, para proponer que ya existe una opción documentada en la literatura que cumple los requisitos de una aproximación a la evaluación de intervenciones sociales, y que emana principios (con adendas) de la forma de pensar de diseñadores sociales. Desde esa perspectiva, un segundo aporte de este texto consiste en hacer visible (y/o explícita) esa conexión.

### La evaluación realista como un enfoque basado en el PDD: el caso de Jornada Única<sup>9</sup>

El objetivo del programa Jornada Única (JU) —una de las intervenciones educativas más ambiciosas que ha implementado el Gobierno colombiano en años recientes (Radinger et al., 2018)— es incrementar el tiempo de permanencia de estudiantes de educación primaria y secundaria del sector público en Colombia. Para ello se recurre a un modelo de gobernanza descentralizada, en el cual las instituciones educativas realizan sus propios diagnósticos de necesidades de infraestructura, recursos humanos y servicios de alimentación escolar, etc., que luego presentan a autoridades regionales y nacionales a fin de focalizar esfuerzos diferenciales (o, según demanda territorial) de política pública. El éxito de dicho esquema de operación radica en un trabajo coordinado entre

8 En el ejemplo que se discute en la siguiente subsección se amplía sobre el uso de este lenguaje específico.

9 Esta sección se basa en el artículo de Parra (2022), donde se pueden consultar detalles del diseño de la evaluación, así como sobre los resultados ampliados del estudio.

las secretarías subnacionales de educación —las cuales se encargan de acompañar a colegios en la identificación de sus necesidades y, posteriormente, en la aplicación y seguimiento de estrategias de mejoramiento— y el Ministerio de Educación, autoridad rectora del sistema educativo.

Un punto de partida de la evaluación de JU (realizada durante el primer semestre de 2019) fue el reconocimiento por parte del Gobierno nacional —financiador del estudio— de la necesidad de profundizar en la comprensión de los retos de su implementación, en tanto los análisis previos sugerían un impacto limitado de la iniciativa en mejorar indicadores de calidad educativa. Puesto en términos del debate de este artículo, existía una necesidad de identificar brechas entre el diseño y la ejecución de diferentes procesos y actividades asociadas a la teoría de cambio del programa y generar rutas de acción para enmendarlas. Dada la naturaleza retroductiva de estos interrogantes (ej. ¿qué propiedades debe tener esta intervención para que pueda ser exitosa en diferentes regiones de Colombia?) se optó por una metodología de evaluación alineada con la ontología del RC. La cita de Wong (2018) del final de la subsección anterior resume la lógica general del análisis propuesto desde esta tradición (de la ER).

Nótese cómo esta propuesta de iniciar el análisis con una reflexión explícita sobre explicaciones potenciales del éxito de JU, teniendo en cuenta elementos de contexto que condicionan los efectos causales del programa, evoca el análisis retroductivo discutido arriba (Jagosh, 2020). Más aún, la noción de mecanismo —en términos de una descripción sobre cómo los actores sociales *in situ* interpretan y reaccionan ante los recursos que les son entregados por una intervención (Pawson, 2013; Pawson & Tilley, 1997)— pone sobre la mesa un esfuerzo explícito por analizar interacciones entre personas (ej. docentes, familias, servidores públicos) con las condiciones (materiales, culturales) de los entornos macroestructurales en los que habitan, a partir de teorías realistas sujetas a ser inspeccionadas durante etapas del trabajo de campo.

Tabla 1. Teorías *realistas* preliminares y foco del trabajo de campo

Teorías preliminares estudio	¿Contextos para indagar?	Instrumento cualitativo
<b>Hipótesis 2:</b> el acompañamiento directo del MEN a Instituciones Educativas en su proceso de innovación pedagógica es crucial para motivar a actores del sector educativo local ( <u>Mecanismo</u> ) a implementar el programa JU ( <u>Resultado</u> ). Dicho acompañamiento (directo) se hace más relevante cuando las autoridades educativas locales cuentan con una baja capacidad de gestión pedagógica para realizar dicho acompañamiento ( <u>Contexto</u> ).	Colegios en municipios con Gobiernos locales poco eficientes	<ul style="list-style-type: none"> <li>• Entrevista autoridad educativa local</li> <li>• Entrevista rectores</li> </ul>
<b>Hipótesis 6:</b> existe una relación directa entre las habilidades pedagógicas de los docentes y su capacidad de despertar el interés de los estudiantes ( <u>Mecanismo</u> ) y, por tanto, con la motivación de estos segundos frente al estudio ( <u>Resultado</u> ). Dicho factor (habilidad docente) puede ser aún más relevante en contextos de adversidad económica y social de estudiantes ( <u>Contexto</u> ).	Colegios con estudiantes de bajos niveles socioeconómicos (SES)	<ul style="list-style-type: none"> <li>• Entrevistas docentes</li> <li>• Grupos focales estudiantes</li> <li>• Observaciones etnográficas en instituciones educativas</li> </ul>

Fuente: adaptado de Parra (2022).

La primera columna de la tabla 1 presenta, a modo de ejemplo, dos de las ocho teorías realistas iniciales (usando el lenguaje de Contexto-Mecanismo y Resultado, CMO),<sup>10</sup> que orientaron la evaluación.<sup>11</sup> Dichas conjeturas sobre mecanismos asociados a diferentes momentos y contextos de la implementación de

10 Las configuraciones CMO son una heurística utilizada para delinear *teorías realistas* en el espíritu del principio SAC de Archer (2013a). Estas ponen de manifiesto el hecho de que el resultado de cada desenlace específico de una intervención educativa (ej. actividades específicas de JU) dependerá de una relación entre contexto (estructura culturales o materiales) y el mecanismo (asociado a la agencia humana) desatado en diferentes etapas de su implementación.

11 No se entra en mayores detalles de la temática específica del estudio. Simplemente para resaltar el proceso analítico (abductivo/retroductivo) propuesto, se aclara que primera hipótesis propone que la llegada directa del Ministerio de Educación a colegios en regiones con una baja capacidad institucional local es fundamental para empoderar actores educativos a acelerar la implementación de JU. Por su parte, la segunda anticipa la relevancia de las habilidades didácticas de docentes para generar interés por estudiar particularmente entre estudiantes de hogares más vulnerables.

JU surgieron de la conjugación de tres fuentes de consulta: documentación oficial del programa que permitió rastrear los elementos centrales de su teoría de cambio, una revisión de estudios sobre los efectos de iniciativas similares en América Latina y una discusión con funcionarios del Ministerio de Educación Nacional encargados del diseño y el seguimiento al programa (en la cual se unieron algunos funcionarios de secretarías de educación locales). Se materializó con ello un ejercicio abductivo inicial, en el que se usaron diferentes marcos de referencia para realizar una primera interpretación de las dinámicas causales potenciales del programa.

La planeación del trabajo de campo en la siguiente fase del proyecto se hizo en función de la lógica retroductiva de recolectar evidencia necesaria para revisar las conjeturas iniciales y, por esa vía, inspeccionar mecanismos causales en acción. La tabla 1 presenta también información relevante para la recolección de evidencia en función de este último principio metodológico, en tanto enlaza cada teoría realista con indicadores disponibles (ej. en bases oficiales del Gobierno) útiles para determinar contextos específicos en los que puede ser más relevante indagar cada posible explicación causal del programa.<sup>12</sup> Dicho ejercicio de muestreo cualitativo llegó incluso a preseleccionar informantes clave, según la temática de cada hipótesis de trabajo a ser explorada (ver columna 3 de la tabla 1). Al final, se visitaron diez instituciones educativas ubicadas en diferentes regiones del país para conversar en profundidad con rectores, docentes, estudiantes, padres de familia y realizar cerca de 400 horas de observaciones etnográficas. La consulta incluyó a diez funcionarios de secretarías de educación regionales, a fin de inspeccionar los mecanismos de gobernanza entre ese nivel de Gobierno y las autoridades educativas nacionales.

Finalmente, el equipo de la evaluación también previó estrategias específicas de recolección de información cualitativa,<sup>13</sup> bajo la lógica de ir más allá de la documentación de percepciones subjetivas (dispersas) de usuarios sobre el programa evaluado —como lo dictaría una lógica más inductiva de indagación (Emmel, 2013)—. Para ello, se utilizaron referentes metodológicos de la ER con el fin de diseñar protocolos de entrevista y de grupos focales orientados a

12 También se consideraron otros factores que no se reportan acá, como el tiempo de implementación de JU en la institución educativa visitada; así mismo, se consideró que la muestra incluyera zonas rurales y urbanas.

13 El diseño de la evaluación incluyó también la recolección de información cuantitativa. Sin embargo, esos detalles no son relevantes para efectos de la presente discusión.

fomentar conversaciones semiestructuradas con actores educativos regionales. El objetivo era invitar a estos segundos a ir más allá de sus primeras impresiones sobre JU, motivándolos a opinar y contrastar sus experiencias personales, con las teorías realistas aportadas (implícitamente)<sup>14</sup> por el equipo evaluador. Un ejemplo del resultado de esta dinámica en acción se refleja en el siguiente comentario que surge en una entrevista en una Secretaría de Educación regional, a propósito de la reflexión en campo sobre la primera hipótesis de la tabla 1. A partir de este tipo de evidencia (y una vez realizado el ejercicio de triangulación con otras fuentes de información) fue posible empezar a refinar el entendimiento de factores estructurales, más allá de las capacidades locales, condicionantes de los avances en el proceso de implementación:

Si el Ministerio [de Educación] no tiene la posibilidad de hacer el acompañamiento permanente [...] nosotros sí tenemos aquí cómo, a partir de unos criterios, claros y de un trabajo estructurado, articulado. Nosotros sí podemos hacerle el seguimiento. Nosotros sí podemos coadyuvar para que se haga bien. Entonces lo que hay es el afán de cumplir metas, pero no de articular un trabajo de calidad. (UT Econometría-SEI, 2019, p. 57)<sup>15</sup>

Ahora bien, por la naturaleza de este estudio —que a diferencia de la experiencia discutida con el programa CFK, no permitió hacer seguimiento, o incluso plantear nuevas etapas de indagación— no se reporta un ejercicio explícito de revisión de la teoría de cambio inicial de JU en función de futuras fases de implementación. No obstante, el punto que ha querido resaltarse al presentarlo como ejemplo de una metodología de evaluación alineada con el PdD (potenciado por el RC) es que la combinación de estrategias de deliberación con usuarios finales, mediadas por previsiones metodológicas para examinar mecanismos causales en campo, puede proveer información contextual relevante para rediseñar soluciones que puedan resultar viables, no solo en lo técnico, sino también desde una perspectiva de legitimidad social y política. El siguiente fragmento, que condensa los hallazgos centrales de la presente evaluación, soporta esta última conclusión:

14 Parra (2022) presenta en mayor detalle esta estrategia. Esta implicó, por ejemplo, traducir las teorías realistas de la evaluación a un lenguaje menos abstracto.

15 Entrevista de un representante de una Entidad Territorial Certificada, ETC, tomada del informe oficial de la evaluación de operaciones del programa Jornada Única.

Al leer [los] resultados generales [de la evaluación] en el contexto de algunas discusiones teóricas de educación comparada más generales, es difícil esperar efectos significativos de esta intervención en el funcionamiento de las escuelas públicas colombianas, al menos como lo anticipó el diseño del programa. [El] impacto potencial de variables específicas en la transformación de las prácticas educativas locales parece limitado por un entorno institucional hostil, producto de problemas de coordinación y contradicciones entre las autoridades nacionales y sub nacionales. Además, nuestros hallazgos sugieren que la implementación de JU no solo se ve afectada por estas circunstancias institucionales, sino que, en ocasiones, se refuerzan cuando el Gobierno nacional decide deliberadamente no colaborar con los [Gobiernos locales] para apoyar a las escuelas. (Parra, 2022, p. 9)

Como se resalta en el documento citado, uno de los factores que inspiró el diseño de la evaluación del programa JU fue una crítica metodológica implícita a evaluaciones previamente realizadas que examinaban intervenciones de política pública similares. Muchos de esos análisis —propios de campos como el del análisis económico de la educación— tienden a ser más de corte deductivo, en el sentido que proponen soluciones predefinidas a problemas educativos complejos (ej. generar incentivos al desempeño docente para mejorar rendimiento escolar de estudiantes) y centran sus esfuerzos empíricos en verificar correlaciones estadísticas entre indicadores de insumo y resultado (Tikly, 2015). La cita de arriba exalta, no obstante, elementos contextuales (o de la realidad de la operación del programa) más amplios que fueron capturados por la estrategia reductiva del estudio. Estos elementos le permitieron al equipo evaluador indagar sobre el entorno institucional —más asociado a factores macroestructurales, dentro del cual se despliegan estrategias de mejoramiento escolar (ej. la formación docente)— que, de ser ignorados, pueden limitar la efectividad de futuras decisiones de política pública para fortalecer la permanencia y la calidad escolar en Colombia.

## **Discusión: hacia la institucionalización del PdD realista**

Este artículo plantea una reflexión sobre los puntos de convergencia y el potencial diálogo entre el PdD y el RC en el campo de la evaluación de intervenciones sociales. En síntesis, se argumenta que la lógica abductiva (pragmática) inherente al PdD representa un avance frente a escuelas tradicionales de

evaluación, en cuanto a su capacidad de identificar posibles soluciones a retos de implementación de las intervenciones analizadas. No obstante, se sugiere que los fundamentos ontológicos del pragmatismo limitan su alcance para indagar por mecanismos causales asociados al éxito/fracaso de diferentes acciones en escenarios sociales y políticos determinados. El razonamiento retroductivo, asociado al RC, puede entonces complementar estrategias de evaluación basadas en PdD para identificar opciones de diseño de políticas o programas que se ajusten mejor a la realidad de los contextos en que van a ser implementadas.

Sin la intención de entregar un manual de cómo aplicar estos principios, los estudios de caso presentados ayudan a pensar en posibles implicaciones en la práctica —y en los beneficios epistemológicos— del razonamiento abductivo y retroductivo al servicio de una evaluación basada en diseño. De esta manera, el artículo hace un aporte decidido a la literatura del PdD aplicado a la administración pública en uno de sus frentes menos explorados. En concreto, más que delinear una innovación metodológica específica, el análisis propuesto —el cual puede entenderse como un ejercicio de desconstrucción del fundamento epistemológico del PdD— llegó a la conclusión de que ya existe una tradición en la disciplina de la evaluación que recoge de manera explícita sus fundamentos de investigación para abordar problemas complejos, además de enmendar sus posibles limitaciones con el fin de pensar en diseños de política mejor equipados para responder a las particularidades de diferentes contextos socio-políticos: la evaluación realista (ER).

Una forma de cerrar esta reflexión es preguntarse cómo generar estrategias, o incluso arreglos institucionales, que fomenten ejercicios de evaluación de programas y políticas públicas orientados por lógicas abductivas y retroductivas de razonamiento. Los denominados Laboratorios de Diseño de Política —que han ido adquiriendo un papel importante como escenarios de deliberación entre diseñadores (sociales) al servicio de la generación de valor para la ciudadanía (Bason & Austin, 2022; Olejniczak et al., 2020; Villa Álvarez et al., 2022)— tienen el potencial de suplir dicha necesidad. Sin embargo, como apunta un reciente estudio comparativo, la experiencia en el funcionamiento de cerca de 150 laboratorios (en su mayoría en Europa y Estados Unidos) sugiere que estos aún tienen una capacidad limitada en reconocer la complejidad detrás de mecanismos sociales que dan viabilidad a diferentes soluciones (creativas) de política (Olejniczak et al., 2020).

A lo mejor, la incorporación de los preceptos realistas a las estrategias de diálogo colaborativo en estos centros es un aporte adicional que puede hacer el RC al PdD aplicado. Por ejemplo, se podría pensar en lógicas de trabajo en las que, a la par que se establecen propuestas de diseño de programas o intervenciones sociales, se planteen reflexiones más de carácter retroductivo —lideradas, por ejemplo, por profesionales de áreas de ciencias sociales— para planear evaluaciones realistas que permitan evaluarlas. Los resultados de esas evaluaciones pueden luego contribuir a generar nuevas ideas y marcos de referencia que alimenten futuras reflexiones interdisciplinarias para fundamentar nuevos diseños de soluciones potenciales a retos del sector público.

## Agradecimientos

Los autores agradecen a ambos revisores pares por sus excelentes recomendaciones para mejorar el artículo. Se deja también constancia que el estudio de CFK fue financiado parcialmente por el Ministerio de TIC (Colombia), el Ministerio de Educación, y el British Council Colombia y que las referencias que se hacen al mismo son únicamente responsabilidad de los autores y no representan necesariamente la visión oficial de los entes financiadores.

## Referencias

- Adams, C., & Nash, J. B. (2016). Exploring Design Thinking Practices in Evaluation. *Journal of MultiDisciplinary Evaluation*, 12(26), 12-17. [https://journals.sfu.ca/jmde/index.php/jmde\\_1/article/view/434](https://journals.sfu.ca/jmde/index.php/jmde_1/article/view/434)
- Archer, M. (2013a). Social Morphogenesis and the Prospects of Morphogenic Society. En M. Archer (Ed.), *Social Morphogenesis* (pp. 1-22). Springer.
- Archer, M. (2013b). *Social origins of educational systems*. Routledge.
- Astbury, B. (2013). Some reflections on Pawson's Science of Evaluation: A Realist Manifesto. *Evaluation*, 19(4), 383-401. <https://doi.org/10.1177/1356389013505039>
- Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences*, 13(1), 1-14. [https://doi.org/10.1207/s15327809jls1301\\_1](https://doi.org/10.1207/s15327809jls1301_1)
- Bason, C. (2016). Introduction: The Design for Policy Nexus. En C. Bason (Ed.), *Design for Policy* (pp. 1-8). Routledge.
- Bason, C., & Austin, R. D. (2022). Design in the public sector: Toward a human centred model of public governance. *Public Management Review*, 24(11), 1727-1757. <https://doi.org/10.1080/14719037.2021.1919186>
- Bhaskar, R., & Hartwig, M. (2010). *The Formation of Critical Realism: A Personal Perspective*. Routledge.

- Brousselle, A., & Buregeya, J.-M. (2018). Theory-based evaluations: Framing the existence of a new theory in evaluation and the rise of the 5th generation. *Evaluation*, 24(2), 153-168. <https://doi.org/10.1177/1356389018765487>
- Carlsson, S. A. (2010). Design Science Research in Information Systems: A Critical Realist Approach. En A. Hevner & S. Chatterjee (Eds.), *Design Research in Information Systems: Theory and Practice* (pp. 209-233). Springer. [https://doi.org/10.1007/978-1-4419-5653-8\\_15](https://doi.org/10.1007/978-1-4419-5653-8_15)
- Cross, N. (2011). *Design thinking: Understanding how designers think and work*. Berg.
- Danermark, B., Ekström, M., & Karlsson, J. C. (2019). *Explaining Society: Critical Realism in the Social Sciences* (2.ª ed.). Routledge. <https://doi.org/10.4324/9781351017831>
- Di Russo, S. (2016). *Understanding the behaviour of design thinking in complex environments* (Tesis de doctorado no publicada). Swinburne University.
- Dorst, K. (2006). Design Problems and Design Paradoxes. *Design Issues*, 22(3), 4-17. <https://doi.org/10.1162/desi.2006.22.3.4>
- Dorst, K. (2011). The core of 'design thinking' and its application. *Design Studies*, 32(6), 521-532. <https://doi.org/10.1016/j.destud.2011.07.006>
- Dorst, K., & Cross, N. (2001). Creativity in the design process: Co-evolution of problem-solution. *Design Studies*, 22(5), 425-437. [https://doi.org/10.1016/S0142-694X\(01\)00009-6](https://doi.org/10.1016/S0142-694X(01)00009-6)
- Douven, I. (2021). Abduction. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 ed.). Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2021/entries/abduction/>
- Elder-Vass, D. (2022). Pragmatism, critical realism and the study of value. *Journal of Critical Realism*, 21(3), 261-287. <https://doi.org/10.1080/14767430.2022.2049088>
- Emmel, N. (2013). *Sampling and choosing cases in qualitative research: A realist approach*. Sage.
- Espinal, A., Vieira, C., & Guerrero-Bequis, V. (2022). Student ability and difficulties with transfer from a block-based programming language into other programming languages: A case study in Colombia. *Computer Science Education*, 1-33. <https://doi.org/10.1080/08993408.2022.2079867>
- Fontaine, G. (2020). The contribution of policy design to realist evaluation. *Evaluation*, 26(3), 296-314. <https://doi.org/10.1177/1356389020902496>
- Gray, C., & Shaw, J. (2018). From summative to developmental: Incorporating design-thinking into evaluations of complex interventions. *Journal of Integrated Care*, 27(3), 241-248. <https://doi.org/10.1108/JICA-07-2018-0053>
- Hermus, M., Buuren, A. van, & Bekkers, V. (2020). Applying design in public administration: A literature review to explore the state of the art. *Policy & Politics*, 48(1), 21-48. <https://doi.org/10.1332/030557319X15579230420126>

- Hoadley, C., & Campos, F. C. (2022). Design-based research: What it is and why it matters to studying online learning. *Educational Psychologist*, 57(3), 207-220. <https://doi.org/10.1080/00461520.2022.2079128>
- Jagosh, J. (2020). Retroductive theorizing in Pawson and Tilley's applied scientific realism. *Journal of Critical Realism*, 19(2), 121-130. <https://doi.org/10.1080/14767430.2020.1723301>
- Lewis, J. M., McGann, M., & Blomkamp, E. (2020). When design meets power: Design thinking, public sector innovation and the politics of policymaking. *Policy & Politics*, 48(1), 111-130. <https://doi.org/10.1332/030557319X15579230420081>
- Maxwell, J. A. (2012). *A realist approach for qualitative research*. Sage.
- Mintrom, M., & Luetjens, J. (2016). Design Thinking in Policymaking Processes: Opportunities and Challenges. *Australian Journal of Public Administration*, 75(3), 391-402. <https://doi.org/10.1111/1467-8500.12211>
- Mishra, P., & Koehler, M. J. (2006). Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge. *Teachers College Record*, 108(6), 1017-1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Olejniczak, K., Borkowska-Waszak, S., Domaradzka-Widła, A., & Park, Y. (2020). Policy labs: The next frontier of policy design and evaluation? *Policy & Politics*, 48(1), 89-110. <https://doi.org/10.1332/030557319X15579230420108>
- Parra, J. D. (2016). Realismo crítico: Una alternativa en el análisis social. *Sociedad y Economía*, (31), 215-238. <https://doi.org/10.25100/sye.v0i31.3895>
- Parra, J. D. (2021). Sobre el uso de la evidencia y la validez externa en la evaluación de intervenciones sociales: Una mirada crítica. *Colombia Internacional*, (105), 175-198. <https://doi.org/10.7440/colombiaint105.2021.07>
- Parra, J. D. (2022). Decentralisation and school-based management in Colombia: An exploration (using systems thinking) of the Full-Day Schooling programme. *International Journal of Educational Development*, 91. <https://doi.org/10.1016/j.ijedudev.2022.102579>
- Patton, M. Q. (2010). *Developmental evaluation: Applying complexity concepts to enhance innovation and use*. Guilford press.
- Pawson, R. (2003). Nothing as Practical as a Good Theory. *Evaluation*, 9(4), 471-490. <https://doi.org/10.1177/135638900300900407>
- Pawson, R. (2013). *The Science of Evaluation: A Realist Manifesto*. Sage.
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. Sage.
- Peters, B. G. (2015). State failure, governance failure and policy failure: Exploring the linkages. *Public Policy and Administration*, 30(3-4), 261-276. <https://doi.org/10.1177/0952076715581540>
- Porpora, D. V. (2015). *Reconstructing sociology: The critical realist approach*. Cambridge University Press.

- Radinger, T., Echazarra, A., Guerrero, G., & Valenzuela, J. P. (2018). *OECD reviews of school resources: Colombia 2018*. OECD.
- Rigby, J. G., Woulfin, S. L., & März, V. (2016). Understanding How Structure and Agency Influence Education Policy Implementation and Organizational Change. *American Journal of Education*, 122(3). <https://doi.org/10.1086/685849>
- Ritz, B. (2020). Comparing abduction and retroduction in Peircean pragmatism and critical realism. *Journal of Critical Realism*, 19(5), 456-465. <https://doi.org/10.1080/14767430.2020.1831817>
- Romme, A. G. L., & Meijer, A. (2020). Applying design science in public policy and administration research. *Policy & Politics*, 48(1), 149-165. <https://doi.org/10.1332/030557319X15613699981234>
- Sayer, A. (1992). *Method in Social Science* (2.<sup>a</sup> ed.). Routledge.
- Sayer, A. (2000). *Realism and social science*. Sage.
- Schön, D. A. (1992). Designing as reflective conversation with the materials of a design situation. *Knowledge-Based Systems*, 5(1), 3-14. [https://doi.org/10.1016/0950-7051\(92\)90020-G](https://doi.org/10.1016/0950-7051(92)90020-G)
- Tavory, I., & Timmermans, S. (2014). *Abductive Analysis: Theorizing Qualitative Research*. University of Chicago Press.
- Tikly, L. (2015). What works, for whom, and in what circumstances? Towards a critical realist understanding of learning in international and comparative education. *International Journal of Educational Development*, 40, 237-249. <https://doi.org/10.1016/j.ijedudev.2014.11.008>
- UNESCO. (2017). *Cracking the code: Girls' and women's education in science, technology, engineering and mathematics (STEM)* [Conferencia]. International Symposium and Policy Forum. UNESCO, Paris, France.
- UT Econometría-SEI. (2019). Evaluación de operaciones y de resultados del programa Jornada Única que permita analizar el proceso de implementación y los resultados generados en los beneficiarios. Departamento Nacional de Planeación; MINEDUCACIÓN. [https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Evaluacion\\_Jornada\\_Unica\\_Producto\\_4\\_Informe\\_Resultados.pdf](https://colaboracion.dnp.gov.co/CDT/Sinergia/Documentos/Evaluacion_Jornada_Unica_Producto_4_Informe_Resultados.pdf)
- Vásquez, S. (2011). Comunidades de práctica. *Educación*, 47(1), 51-68. <https://doi.org/10.5565/rev/educar.71>
- Vieira, C., Gómez, R., Gómez, M., Canu, M., & Duque, M. (2022). Implementing Unplugged CS and Use-Modify-Create to Develop Student Computational Thinking Skills: – A Nationwide Implementation in Colombia. *Educational Technology & Society*, 26(3). [https://www.j-ets.net/collection/forthcoming-articles/26\\_3](https://www.j-ets.net/collection/forthcoming-articles/26_3)
- Villa Álvarez, D. P., Auricchio, V., & Mortati, M. (2022). Mapping design activities and methods of public sector innovation units through the policy cycle model. *Policy Sciences*, 55(1), 89-136. <https://doi.org/10.1007/s11077-022-09448-4>

- Weedon, S. (2019). The Core of Kees Dorst's Design Thinking: A Literature Review. *Journal of Business and Technical Communication*, 33(4), 425-430. <https://doi.org/10.1177/1050651919854077>
- Wong, G. (2018, marzo-abril). Getting to grips with context and complexity - the case for realist approaches. *Gaceta Sanitaria*, 32(2), 109-110. <https://doi.org/10.1016/j.gaceta.2017.05.010>
- Wuisman, J. (2005). The Logic of Scientific Discovery in Critical Realist Social Scientific Research. *Journal of Critical Realism*, 4(2), 366-394. <https://doi.org/10.1558/jocr.v4i2.366>
- Zotzmann, K., Barman, E., Porpora, D. V., Carrigan, M., & Elder-Vass, D. (2022). Round table: Is the common ground between pragmatism and critical realism more important than the differences? *Journal of Critical Realism*, 21(3), 352-364. <https://doi.org/10.1080/14767430.2022.2073674>

