

# La problemática de la causalidad en la Sociología.

## Algunas reflexiones.

Mónica Gerber Plüss<sup>161\*</sup>

---

### I. Introducción.

Lograr establecer relaciones causales en sociología dista bastante de ser un problema trivial. El estudio de lo social se distancia de otras ciencias en términos de la dificultad –y en muchos casos, imposibilidad– de intervenir su unidad de estudio. En sociología resulta prácticamente imposible realizar estudios de laboratorio o experimentos, y en general, el método que predomina es el de los estudios observacionales, especialmente la encuesta. De esta manera, en la mayoría de los casos, la causalidad termina siendo definida utilizando métodos estadísticos, como la correlación o modelos de regresión. Otros métodos más avanzados, como los modelos de ecuaciones estructurales, el análisis factorial, entre otros, han permitido desarrollar la estadística y mejorar los modelos explicativos en la sociología. Sin embargo, como afirman Pötter y Blossfeld<sup>162</sup>, esta sofisticación de la estadística también ha tendido a separar la estadística del resto de la investigación social. Existe consenso entre sociólogos con respecto a que la causalidad no puede derivarse simplemente de datos empíricos, sino que debe estar siempre vinculada a la teoría.

Si se observa la gran mayoría de la investigación sociológica de los últimos años, se puede ver que estas dificultades han sido más bien obviadas y no solucionadas. En muchas investigaciones se puede ver un uso bastante laxo de la palabra causalidad. En muchos casos, se sacan conclusiones con respecto a relaciones causales entre variables, aunque todo lo que se ha hecho es una correlación. Según Rutter<sup>163</sup>, en muchos casos, los autores evitan hablar de causalidad pero terminan de todas formas sacando conclusiones de este tipo.

En otros casos, se utilizan técnicas estadísticas sofisticadas para hacer “fit” (ajuste) de modelos, sin tener un real modelo teórico e hipótesis asociadas. Un claro ejemplo de esto son las funciones “Stepwise” de los programas estadísticos, los cuales agregan y eliminan variables al modelo según criterios estadísticos. El investigador puede introducir todas las variables que quiera y obtener un modelo a la medida, omitiendo cualquier alusión teórica o sustantiva.

La pregunta que surge de esta problemática es la siguiente: ¿Cómo es posible establecer relaciones causales, asumiendo y no obviando las dificultades que presenta la sociología? A continuación se presentarán distintas nociones de causalidad y se discutirá el acercamiento a la causalidad que debiera ser adoptado por la sociología.

---

<sup>161\*</sup> Socióloga, Universidad de Chile. Estudiante Magíster en Métodos de Investigación Social, London School of Economics and Political Sciences. Correo electrónico: [mmgerber@gmail.com](mailto:mmgerber@gmail.com)

<sup>162</sup> Ulrich Pötter y Hans Peter Blossfeld, “Causal Inference From Series of Events” en *European Sociological Review*, vol. 17, N° 1, 2001, pp. 21-32.

<sup>163</sup> Michael Rutter, “Proceeding from Observed Correlation to Causal Inference” en *Perspectives on Psychological Science*, vol. 2, N° 4, 2007, pp. 377-395.

## II. Nociones de Causalidad.

John Stuart Mill definió en 1843 las tres condiciones fundamentales para la causalidad: a) la causa debe preceder temporalmente al efecto; b) la causa debe estar estadísticamente asociada al efecto y c) no debe existir una explicación alternativa<sup>164</sup>. Estos principios son aceptados hasta el día de hoy. Sin embargo, debieran ser asumidos como necesarios pero no suficientes.

Cox y Wermuth<sup>165</sup> y Goldthorpe<sup>166</sup> afirman que existen tres nociones de causalidad: la causalidad como una asociación estable, la causalidad como el efecto de una intervención y la causalidad como la explicación de un proceso<sup>167</sup>. A continuación se procederá a describir estas tres nociones.

### 1. Causalidad como una asociación estable.

Esta noción corresponde en términos prácticos al uso de la estadística para determinar una relación causal. Si bien la estadística sólo nos muestra si dos variables se encuentran asociadas, es posible, por distintos medios, analizar si esta asociación puede ser considerada como una relación causal. Si dos variables se encuentran asociadas, y es posible demostrar que esta asociación es asimétrica (es decir, que la variable independiente (X) influye en la dependiente (Y) y no al revés), y que no existen otras posibles explicaciones a esta asociación, es posible agregar evidencia a favor de la existencia de una relación causal<sup>168</sup>. Para Goldthorpe<sup>169</sup>, para que la asociación pueda ser considerada como una relación causal, debe ser robusta, esto es, no debe desaparecer al agregar nuevas variables.

Si bien esta es la noción más utilizada en la sociología, presenta una serie de dificultades, dado que en principio, una asociación no necesariamente representa una relación causal. Para evitar sacar conclusiones erróneas, y analizar si efectivamente la asociación es robusta, es necesario analizar la presencia de alguna de las siguientes situaciones:

- a. Relaciones espurias. Puede existir una asociación entre una variable independiente (X) y una dependiente (Y) y, sin embargo, puede que esta relación se deba a que existe una tercera variable (Z) que causa tanto X como Y. De esta manera, la correlación encontrada entre X e Y no implica una relación causal entre ambas variables.
- b. Variables intervinientes. Es posible que existan variables que intervienen la relación entre X e Y. Si existen variables intervinientes, la relación entre X e Y cambia al momento de introducir una variable Z en la ecuación. Por ejemplo, es posible que para distintos niveles de una variable interviniente Z, la relación entre X e Y sea distinta o hasta desaparezca.
- c. Relaciones en cadena. Si bien X e Y se encuentran asociadas, es posible que esto se deba a que X causa Z, mientras que Z causa Y. De esta manera, la presencia de una relación en cadena puede hacer parecer que X efectivamente causa Y, cuando esto en rigor no es así.
- d. Múltiples causas. En sociología es poco frecuente encontrar variables independientes que por sí solas sean capaces de explicar una variable dependiente. En general, es necesario tomar en cuenta distintas posibles variables que pueden actuar en conjunto<sup>170</sup>.

---

<sup>164</sup> En Michael Rutter, Op. Cit.

<sup>165</sup> D.R. Cox y Nanny Wermuth, "Some Statistical Aspects of Causality" en *European Sociological Review*, vol. 17, N° 1, 2001, pp. 65-74.

<sup>166</sup> John Goldthorpe, "Causation, Statistics and Sociology" en *European Sociological Review*, vol. 17, N° 1, 2001, pp. 1-20.

<sup>167</sup> Según la nomenclatura utilizada por Cox y Wermuth, aunque el principio es el mismo que en Goldthorpe.

<sup>168</sup> Véase D.R. Cox y Nanny Wermuth, Op. Cit.

<sup>169</sup> John Goldthorpe, Op. Cit.

<sup>170</sup> Ver D.R. Cox y Nanny Wermuth, Op. Cit. y Michael Rutter, Op. Cit.

La presencia de este tipo de relaciones no es siempre analizada de forma completa. Bhrolcháin<sup>171</sup> presenta un buen ejemplo de esta situación. La autora analiza estudios que se han llevado a cabo para investigar los efectos a largo plazo que tiene en niños el divorcio de sus padres. Este tipo de preguntas de investigación, muy comunes en la sociología, presenta una serie de dificultades. En términos generales, resulta muy difícil sacar conclusiones causales con respecto a si efectivamente el divorcio de los padres generó determinados resultados en los hijos. Por ejemplo, ¿cómo saber si resultados escolares deficientes en hijos de padres divorciados se deben al divorcio mismo, o a algún factor común que puede haber llevado a los padres a divorciarse, al mismo tiempo que a los hijos a tener resultados escolares deficientes? En este caso, se trataría de una relación espuria entre los resultados de los hijos y el hecho que los padres se hayan divorciado o no. De hecho, es posible pensar en un gran número de variables que podrían encontrarse a la base de esta relación espuria: incapacidad de los padres para mantener relaciones interpersonales satisfactorias a largo plazo, deficiente control de ira, frustración, entre otros.

Ante estas dificultades, resulta muy difícil sacar conclusiones causales de la relación entre ambas variables. Sin embargo, Bhrolcháin<sup>172</sup> encuentra que en la mayor parte de los estudios al respecto se concluye la existencia de relaciones causales, sin justificación suficiente. En general, los estudios utilizan métodos multivariados para controlar una serie de variables que podrían explicar la relación encontrada. Sin embargo, no son exhaustivos en las variables analizadas. La autora afirma que una serie de variables, tales como, el comportamiento y características personales de los padres, interacción entre los padres, interacciones padre-hijo y aspectos situacionales del ambiente familiar, no son controladas por los estudios. A su vez, la autora afirma que todos los estudios utilizan resultados correlacionales para obtener conclusiones causales, cuando evidentemente correlación no implica causalidad.

Este artículo es un buen ejemplo para demostrar la relevancia de analizar posibles variables que no están siendo consideradas y que pueden cambiar la relación observada en un principio. De esta manera, analizar cada una de estas situaciones resulta crítico al momento de eliminar explicaciones alternativas.

Sin embargo, Goldthorpe<sup>173</sup>, critica la opción de ir agregando variables para analizar si la relación se mantiene. Según el autor, el uso de estas técnicas corresponde al pasado y no al futuro de la sociología.

## **2. Causalidad como el efecto de una intervención.**

Para muchos autores, la causalidad como asociación estable es un criterio que genera bastantes dudas<sup>174</sup>. En general, los autores plantean que el uso de diseños experimentales, donde se manipula una variable dejando constante todo lo demás, es un método mucho más certero para determinar causalidad. Sin embargo, para poder comparar dos o más intervenciones, es imprescindible que la asignación de los casos a los grupos experimentales sea aleatoria<sup>175</sup>. Al hacer esto, es posible asegurar que inicialmente los grupos sean equivalentes en promedio. De esta manera, se controla que los resultados de la variable dependiente se deban efectivamente a la diferencia en los tratamientos y no a diferencias pre-existentes entre los grupos.

Sin embargo, para que una variable pueda ser considerada como causa, esta debiera ser manipulable<sup>176</sup>. Con este requisito, los autores pretenden dejar fuera como posibles causas a todas las variables que

<sup>171</sup> Máire Ní Bhrolcháin, "Divorce Effects and Causality in the Social Sciences" en *European Sociological Review*, vol. 17, N° 1, 2001, pp. 33-57.

<sup>172</sup> Máire Ní Bhrolcháin, Op. Cit.

<sup>173</sup> John Goldthorpe, Op. Cit.

<sup>174</sup> Véase D.R. Cox y Nanny Wermuth, Op. Cit. y John Goldthorpe, Op. Cit.

<sup>175</sup> Véase Paul Holland, "Statistics and Causal Inference" en *Journal of the American Statistical Association*, vol. 81, N° 396, 1986, pp. 945-960; John Goldthorpe, Op. Cit. y William Shadish, Thomas Cook y Donald Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston, Houghton Mifflin Company, 2002.

<sup>176</sup> Véase D.R. Cox y Nanny Wermuth, Op. Cit.; John Goldthorpe, Op. Cit. y Paul Holland, Op. Cit.



de alguna manera representan atributos de las personas. De esta manera, por ejemplo, para evaluar el rendimiento en una prueba, una causa puede ser la técnica de estudio utilizada con distintos niños. Sin embargo, una causa no podría ser el sexo o nivel socioeconómico de los niños, ya que estas variables son estructurales y no pueden ser manipuladas. Este requisito no se debe solamente al hecho de que los atributos no pueden ser analizados por medio de un experimento, sino también a que en sí mismos no constituyen una acción y por lo tanto, no podrían generar un efecto. Según Holland<sup>177</sup>, no tiene sentido hablar de cómo un atributo determina otro atributo. De hecho, Holland llega a afirmar como imperativo que no existe causación sin manipulación.

Otro tema que debe ser considerado es el nivel en el que el experimento logra alcanzar validez interna y externa. En términos de validez interna, no siempre es posible controlar todas las variables, especialmente porque no siempre es posible trabajar en condiciones de laboratorio. Y, en caso de poder usar condiciones

de laboratorio, resulta cada vez más difícil asegurar la validez externa del estudio. ¿Es realmente posible generalizar un estudio de laboratorio a situaciones fuera del mismo? En general, mientras más controlado sea el experimento, y, por lo tanto, tenga mayor validez interna, más difícil será aplicar este experimento a una muestra representativa de la población.

Debido a todos estos aspectos, resulta muy complicado utilizar este tipo de estudios en sociología. Por un lado, es muy difícil lograr manipular variables o asignar casos a distintas intervenciones. A su vez, en los casos en los que esto sí es posible, resulta muy difícil controlar todas las variables que puedan estar afectando en el aspecto a medir. También en términos de validez externa, los sociólogos han preferido utilizar estudios que sean representativos de una muestra mayor, tales como las encuestas por muestreo. Al mismo tiempo, en sociología se tienden a utilizar frecuentemente variables estructurales (atributos) como variables independientes, mientras que estas no pueden ser analizadas por medio de experimentos.

Según algunos autores, como Campbell y Stanley<sup>178</sup> y Sobel<sup>179</sup>, la sociología debiera de todas formas hacer experimentos o cuasi experimentos, en la medida de lo posible. Una opción para hacer experimentos a pesar de estas limitantes son los experimentos naturales, los cuales aprovechan cierto suceso que ha ocurrido sin la intervención del investigador (desastre natural, intervención de una política pública, entre otros) para medir el efecto que esto ha tenido en las personas. Sin embargo, estos métodos presentan serios problemas de validez interna. Para empezar, es difícil que se den las condiciones para que los distintos grupos (beneficiarios de la política pública versus no beneficiarios, por ejemplo) sean equivalentes en todas las otras variables, lo que complica mucho la posibilidad de sacar conclusiones acerca de si efectivamente fue el tratamiento el que generó los resultados y no algún sesgo en la asignación de los sujetos a los grupos. A su vez, en muchos casos, la medición se lleva a cabo cierto tiempo después de la intervención (pueden hasta ser años). En este tiempo, las personas pueden atravesar situaciones muy distintas, las cuales no pueden ser controladas.

<sup>177</sup> Paul Holland, Op. Cit.

<sup>178</sup> Donald Campbell y Julian Stanley, *Experimental and Quasi-Experimental Designs for Research*, Chicago, Rand McNally & Company, 1963.

<sup>179</sup> En John Goldthorpe, Op. Cit.

Un buen ejemplo de este tipo de experimentos naturales es el estudio llevado a cabo por Gould, Lavy y Paserman<sup>180</sup>, para analizar el efecto de la calidad de la educación básica (Elementary School) en los resultados en la educación secundaria (High School). Dado que en general resulta muy difícil separar los efectos de la escuela misma de otros aspectos del medio ambiente y de la familia, es rara vez posible sacar conclusiones causales referidas al efecto de la escuela en resultados a largo plazo. ¿Cómo saber si los buenos resultados de los niños se deben a que asistieron a una escuela de buen nivel, o a que estudiaron con otros niños que se encontraban motivados o debido a que sus padres los motivaron a estudiar?.

Los autores aprovecharon la inmigración masiva de 15 mil etíopes a Israel en 1991 para hacer un experimento natural. Luego de esta inmigración, las familias fueron ubicadas de manera supuestamente aleatoria a distintas zonas de la ciudad y sus hijos fueron asignados a diferentes escuelas básicas, sin importar la familia de la cual provenían. De acuerdo a los autores, esta fue una oportunidad única para analizar el efecto de la escuela básica en los resultados de la educación secundaria, separando el efecto del ambiente familiar. Los autores afirman que usar este experimento natural les permite estudiar el efecto de la calidad de la escuela de una manera lo más parecida posible a un experimento de laboratorio controlado y que, debido a esto, los resultados pueden ser interpretados como causales. De hecho, a continuación concluyen que efectivamente existe una relación causal entre la calidad de la educación básica y los resultados en educación secundaria. Sin embargo, este estudio presenta una serie de deficiencias que impiden sacar conclusiones de este tipo.

Para empezar, el requisito más fundamental para poder sacar esta conclusión es la asignación aleatoria de los niños a distintas escuelas. Si bien en este estudio no se asignaron siguiendo ningún patrón, dado que la asignación no fue hecha por el investigador, resulta difícil asegurar que esta realmente haya sido aleatoria. Es importante no confundir lo que puede ser una asignación accidental con una asignación aleatoria. En este caso, resulta muy factible que las familias extendidas y amigas hayan viajado juntas y que por lo tanto, hayan asignado a sus hijos a las mismas escuelas. A su vez, no toma en cuenta el hecho de que los hermanos también son asignados a las mismas escuelas. Esta situación genera dudas con respecto a que los grupos hayan sido inicialmente equivalentes.

Por otro lado, existe una serie de variables que no pueden ser controladas, tales como el hecho de que los padres cambien a sus hijos de escuela o que estos vivan situaciones diferentes a lo largo de los años. Tampoco es posible controlar el efecto de la familia en los niños a lo largo de estos años. De esta manera, resulta muy difícil concluir que los resultados de educación secundaria (medidos años después de la asignación de los niños a las escuelas), se deban necesariamente a la calidad de la educación básica.

Por último, un estudio de este tipo, que estudia niños con características tan específicas, carece totalmente de validez externa. ¿Es posible asumir que las relaciones encontradas son válidas para alguien más allá de niños etíopes que han atravesado un proceso de inmigración a Israel? Difícilmente. Y sin embargo, los autores finalmente hablan en términos generales de que los resultados de la educación secundaria son causados por la calidad de la educación básica, sin especificar que esto sería válido solamente para esta población específica.

Si bien puede ser cierto que la sociología debiera aprovechar todas las situaciones disponibles para realizar estudios y avanzar en el análisis de relaciones causales entre variables, también es cierto que es indispensable considerar las limitaciones mismas de los estudios que se realizan. Si el experimento utilizado presenta problemas de validez interna y externa, es necesario ser cauteloso con las conclusiones de causalidad que se extraen del estudio y, al mismo tiempo, hay que evitar generalizar los resultados más allá de lo posible.

---

<sup>180</sup> Eric Gould, Victor Lavy y Daniele Paserman, "Immigrating to Opportunity: Estimating the Effect of School Quality Using a natural Experiment on Ethiopians in Israel" en *The Quarterly Journal of Economics*, vol. 119 Issue 2, 2004, pp. 489-526.

### 3. Causalidad como la explicación de un proceso

La última noción de causalidad implica la existencia de un entendimiento del proceso que lleva a que X cause Y. En general, la explicación del proceso proviene de la teoría<sup>181</sup>. Según Goldthorpe<sup>182</sup> este método elimina la problemática de poder o no ocupar atributos como causas.

Goldthorpe propone llevar a cabo los siguientes pasos para explicar un proceso:

1. Establecer el fenómeno a ser explicado
2. Hipotetizar el proceso generativo
3. Testear la hipótesis

El testeo de la hipótesis se debe hacer por partes, esto es, analizar pequeños mecanismos del modelo explicativo, más que todo el modelo al mismo tiempo.

La causalidad como explicación de un proceso se diferencia de la causalidad como asociación estable en el sentido de que se acerca mucho más a la teoría y a la utilización del método científico para la construcción de modelos explicativos. En este mismo sentido, Rogosa<sup>183</sup> distingue entre modelos estadísticos per se y modelos científicos expresados de forma estadística. De esta manera, definir causalidad explicando el proceso, se acercaría mucho más a un modelo científico. Según Goldthorpe, este es el camino que debiera adoptar la sociología para lograr avanzar en el análisis de relaciones causales.

### III. Causalidad y Sociología.

Parece evidente que determinar causalidad según la existencia de una asociación estable o como producto de una intervención presentan grandes deficiencias en la sociología y que la noción de causalidad en la que debiera centrarse la sociología corresponde a la última: causalidad como la explicación de un proceso. De esta manera, la relación entre variables será descrita en términos de un modelo explicativo. Este modelo, a su vez, debe provenir de la teoría. La estadística, por su parte, se convierte así en un soporte al testeo de hipótesis particulares del modelo.

Esta visión permite un mayor acercamiento de la sociología al método científico, el cuál asume un vínculo necesario entre teoría y empiria, al mismo tiempo que utiliza la estadística como una herramienta.

Investigar de esta manera permite evitar lo que en sociología se ha llegado a llamar la “cocinería estadística”, es decir, simplemente ir incorporando y eliminando variables hasta lograr un modelo que ajuste a los datos. A su vez, este enfoque evita que se saquen rápidamente conclusiones sobre la existencia de una relación causal sin tener mayor sustento. Desde esta perspectiva, se debe ser capaz de explicar por qué una variable independiente efectivamente produce la variable dependiente.

Tal vez la conclusión más relevante del modelo propuesto por Goldthorpe<sup>184</sup> tiene que ver con su propuesta de testear hipótesis referidas a partes de un modelo más general. En este sentido, jamás se testearía la hipótesis de que el divorcio de los padres genera determinados resultados escolares en los niños (como se hace normalmente), sino más bien se podría testear cada uno de los pasos en una cadena que lleva del divorcio de los padres a los resultados de los niños.

Por ejemplo, si el modelo explicativo afirma que el divorcio de los padres genera frustración en los niños y que esto, a su vez, genera resultados escolares deficientes, habría que testear si efectivamente

---

<sup>181</sup> Véase D.R. Cox y Nanny Wermuth, Op. Cit.

<sup>182</sup> John Goldthorpe, Op. Cit.

<sup>183</sup> En John Goldthorpe, Op. Cit.

<sup>184</sup> John Goldthorpe, Op. Cit.

el divorcio de los padres genera frustración en los niños y luego analizar si la frustración genera resultados deficientes.

Si bien este es un ejemplo muy burdo, permite ilustrar el razonamiento general: es necesario trabajar con un modelo proveniente de la teoría, hipotetizar con respecto a cada una de las relaciones presentes en el modelo y testear una a una. Si finalmente se observa evidencia para el modelo como un todo, entonces se puede apoyar la existencia de una relación causal entre las variables.

#### **IV. Conclusiones.**

Y entonces, ¿cómo establecemos relaciones causales, asumiendo las dificultades que presenta la Sociología? Como hemos visto, si bien el método más utilizado es el de los modelos estadísticos, en general, esta perspectiva presenta muchas complicaciones. Analizar todas las posibles variables que puedan estar interviniendo la asociación encontrada no es fácil. Además, este tipo de análisis facilita un acercamiento que se basa principalmente en criterios estadísticos y que deja de lado el vínculo con la teoría. A su vez, parece evidente que el uso de experimentos no se adapta fácilmente a la Sociología y que, en los casos en los que es posible utilizarlos, se presentan varias problemáticas que es necesario considerar para no sacar conclusiones sin suficiente fundamento.

Para lograr sacar conclusiones que puedan ser sustentadas con suficiente evidencia, es necesario avanzar en el establecimiento de modelos explicativos basados en la teoría, que luego puedan ser testeados por medio de herramientas estadísticas. La estadística, desvinculada del método científico, no es suficiente para crear conocimiento. Agregar y sacar variables hasta lograr armar un modelo que “funcione” con los datos no nos asegura la calidad del modelo, ni nos permite sacar conclusiones realmente causales.

La visión de Goldthorpe<sup>185</sup> implica un avance muy lento en el establecimiento de modelos: el testeo de un modelo teórico requiere de un conocimiento acabado del tema de estudio y del estado de arte del mismo. A su vez, implica partir por analizar la relación entre cada uno de los componentes del modelo, para luego ser capaces de explicar el modelo como un todo. Sin embargo, este proceso permite sacar conclusiones con un sustento mayor que en el caso de simples ajustes estadísticos. *N*

---

<sup>185</sup> John Goldthorpe, Op. Cit.