

## Construyendo explicaciones: el uso de modelos en la sociología

*Mauricio Salgado\**

### RESUMEN

La construcción de modelos es una actividad básica de toda disciplina científica y la sociología no está al margen de esta regla. Sin embargo, existe una escasa literatura sobre los supuestos epistemológicos y metodológicos que subyacen a cada técnica de modelamiento. El presente artículo busca aportar algunas ideas sobre este tópico. Para ello, indaga en el estatuto científico de las técnicas de modelamiento en la disciplina sociológica, especialmente en aquellas expresadas en lenguajes artificiales. Además, destaca la importancia que ellas tienen para la construcción de explicaciones causales y el lugar que le corresponde a la teoría de la sociedad en la especificación de los modelos. Una visión panorámica de las tres técnicas de modelamiento más usadas en el campo de la sociología, es ofrecida en el artículo, a saber, el modelamiento estadístico, el tipo matemático y la simulación computacional, destacando los compromisos teóricos que ellos suponen, el tipo de explicación que podemos construir y, por cierto, las principales limitaciones, tanto prácticas como sustantivas, que ellos presentan para la investigación social.

### Palabras clave

Explicaciones causales • mecanismos generativos • modelación • simulación computacional • teoría de la sociedad

### Building explanations: On the use of models in Sociology

### ABSTRACT

The process of model building is a basic activity for any scientific discipline, and sociology is no exception to this rule. The sociological literature devoted to clarify the epistemological

---

\* Sociólogo Universidad Alberto Hurtado. Estudiante Ph.D. Centre for Research in Social Simulation (CRESS), Department of Sociology, University of Surrey, Surrey, Reino Unido. E-mail: m.salgado@surrey.ac.uk.

and methodological assumptions behind each modelling technique is rather scarce, however. This article intends to provide some ideas on this topic. In order to do so, it tackles onto the scientific status of modelling techniques in sociology, especially those expressed in artificial or formal languages. In addition, it stresses the importance of these techniques in building or validating causal explanations, and of the role played by social theory in the specification of such models. The three most common and widely used modelling techniques in sociological research are considered, namely, statistical modelling, mathematical modelling and social simulation. The article emphasises the theoretical underpinnings of these modelling techniques, the kind of explanation we can draw from them, and the main drawbacks, both practical and substantive, they present for social research.

### Keywords

Causal explanation • generative mechanisms • modelling • computer simulation • social theory

## Introducción

El desarrollo de toda práctica científica está ligado a la elaboración de modelos que permiten representar o –si el concepto resulta muy cargado de supuestos ontológicos– reproducir internamente los fenómenos que a la práctica científica le interesan. Ninguna ciencia se ha consolidado sin elaborar, expresar, probar y refutar modelos. Las ciencias sociales, y en particular la sociología, no es una excepción a esta regla. Gran parte de sus avances interpretativos y de intervención se deben a la presencia de modelos relativamente consolidados, a las disputas que ellos abren y a las comunidades que los defienden. Sin embargo, pese a que existe una amplia discusión en el campo de la filosofía de las ciencias sobre el estatuto epistemológico de la modelación (Giere 2004, 2006; Godfrey-Smith 2006; Knuutila 2005, 2009), la literatura al respecto en el campo de la sociología es más bien escasa. Esto, no obstante los enormes avances que la sociología ha mostrado en el uso y desarrollo de técnicas de modelamiento, especialmente de aquellas expresadas en lenguajes artificiales, tales como el matemático o el computacional. Este artículo busca, al menos en parte, plantear algunas ideas sobre este tópico.

Iniciaremos este trabajo con algunas reflexiones generales sobre el papel que desempeñan las técnicas de modelamiento en la práctica científica, destacando su importancia en la construcción de explicaciones causales (sección 2). En segundo lugar, haremos un recorrido por las tres principales técnicas de modelamiento que las ciencias sociales han empleado como herramientas de investigación (i.e., de tipo estadístico, matemático y computacional), mostrando sus posibilidades de aplicación, los compromisos teóricos que suponen y los límites que presentan para la comprensión del fenómeno

social (secciones 2, 3 y 4). Finalizaremos el artículo con una breve comparación entre las técnicas de modelamiento presentadas y algunas reflexiones generales sobre la necesaria compatibilidad entre teoría social y técnicas de modelamiento.

## Modelos para armar

El uso de modelos es una forma estándar de comprender el mundo, algo que hacemos intuitivamente, pero que la práctica científica ha formalizado. Puesto que entre el observador y el mundo que debe distinguir existe una gradiente de complejidad en favor del segundo (Nowotny 2005, Luhmann 1998), todo observador, especialmente el científico, debe proceder selectivamente si desea trazar distinciones que puedan ser tratables; esto es, distinciones que se transformen en temas para la comunicación. Los modelos son precisamente herramientas que ayudan a reducir la complejidad del mundo con el fin de hacerlo observable. En términos gruesos, ellos son simplificaciones selectivas –más pequeñas en escala, más simples en sus mecanismos o en las variables que los componen– de alguna otra estructura o fenómeno que buscamos comprender (Gilbert y Troitzsch 2005). Ellos siempre omiten elementos que sabemos son o pueden ser importantes. Para usar una metáfora tradicional, los modelos poseen las mismas propiedades que tienen los mapas: son útiles en cuanto contienen los detalles del territorio que más nos interesan, ignorando los demás. Al igual que los mapas, los modelos representan el mundo, pero las representaciones que ellos nos proveen son convencionales y siempre parciales. Sin embargo, a pesar de su miopía, o quizás como resultado de ella, las técnicas de modelamiento son muy útiles en todas las disciplinas científicas. Esto, por varias razones.

En primer lugar, los modelos nos permiten realizar un doble ejercicio que es de suma importancia para la práctica científica: (a) la *simplificación*, es decir, construyendo modelos reducimos la complejidad del fenómeno de interés, distinguiendo sus principales elementos y haciendo explícita la lógica o mecanismo que lo produce; y (b) la *inferencia*, es decir, nos permiten experimentar, hacer juicios y extraer conclusiones sobre los efectos de cada uno de esos elementos y los mecanismos causales que producen el fenómeno a explicar (White 2000). El reconocimiento de este doble ejercicio fue un logro evolutivo básico para la consolidación de la ciencia moderna, encontrando su primera expresión en la tradición inaugurada por Descartes.

En segundo lugar, mediante el uso de modelos los científicos pueden formalizar las teorías, haciéndolas de paso fácilmente comunicables. Un buen modelo exige definir precisa y claramente los procesos o mecanismos que producen el fenómeno que nos interesa explicar (Taber y Timpone 1996). Por supuesto, los objetos que son comúnmente llamados modelos constituyen una clase muy heterogénea, incluyendo modelos físicos, modelos a escala, modelos análogos y modelos expresados en lenguajes artificiales. El modelo de doble hélice de la estructura molecular del ADN hecho en cartón y alambre

por Watson y Crick, el modelo de segregación residencial realizado con fichas de dos colores sobre un tablero de ajedrez creado por Schelling o el modelo de predador-presa basado en ecuaciones diferenciales establecido por Lotka y Volterra son ejemplos de esta heterogeneidad. En cada caso se trata de formalizar una explicación que dé cuenta de un fenómeno señalando sus causas, como es el proceso de replicación y transmisión de la información genética, la persistencia de la segregación residencial en ciudades multirraciales o la evolución de un ecosistema. Los modelos formales no sólo se restringen al uso de los lenguajes artificiales (la lógica, la matemática o los programas computacionales), aunque son estos últimos los que presentan mayores ventajas para la práctica científica, principalmente debido a su precisión conceptual, la facilidad para decidir sobre su validez lógica o interna, la transparencia de los supuestos usados y la aptitud que ellos revisten para una comunicación libre de ambigüedades.

En tercer lugar, los modelos nos ayudan a probar nuestras teorías. En efecto, los modelos son objetos abstractos contruidos en conformidad con principios o leyes generales que son establecidos y articulados en teorías científicas (todo modelo es siempre una función derivada de alguna teoría). Estos principios o leyes, como los principios de variación, selección y transmisión en la teoría evolutiva, las tres leyes del movimiento en la mecánica clásica o los principios de equilibrio en la moderna teoría económica, tienen un valor heurístico para la construcción de modelos, y de esos principios derivan modelos como los basados en algoritmos genéticos, un oscilador armónico simple o un modelo de equilibrio de precios en un mercado en particular. Principios o leyes generales como aquellos no pueden, por definición, ser presentados a la manera de enunciados empíricamente verificables; antes bien, como sostiene Giere (2004), ellos pueden ser más o menos útiles para la construcción de modelos que posean un buen ajuste con la 'realidad'. Siguiendo a este autor, "[u]no no puede directamente probar principios por medios empíricos. Uno solamente puede probar el ajuste al mundo de modelos particulares que incorporan esos principios" (Giere 2004:751).

En síntesis, en la práctica científica los modelos permiten simplificar los fenómenos de interés a sus principales mecanismos causales; hacen posible la inferencia de juicios y conclusiones mediante una forma *indirecta de experimentación* (basada en la manipulación de ciertos elementos en entornos controlados);<sup>1</sup> formalizan las teorías que están a la base de los modelos (los principios heurísticos abstractos usados para su construcción) y, finalmente, establecen un modo verificable para probar esos supuestos teóricos.

<sup>1</sup> Por supuesto, esto no implica que la construcción de modelos se iguale a la experimentación que otras ciencias realizan, por ejemplo, en laboratorios. En primer lugar, dada la naturaleza de la estructura causal del mundo social, el científico social no tiene el mismo grado de control sobre las variables que intervienen en la producción del fenómeno que el que pueden tener otros en las ciencias experimentales. En segundo lugar, quienes diseñan modelos como forma de experimentación sufren otro tipo de restricciones, en comparación con quienes experimentan en laboratorio. Por ejemplo, todo modelo formalizado en lenguaje matemático debe ser manejado por los requerimientos de 'tratabilidad'. Por lo mismo, preferimos referirnos a la construcción de modelos como una manera 'indirecta' de experimentar.

Mediante estas cuatro operaciones, los científicos pueden proponer *explicaciones* causales válidas —algo que es muy diferente de elaborar clases, enumerar tipologías o presentar descripciones. En términos generales, una explicación es aquella que provee un enunciado causal, plausible en relación con un corpus teórico disponible y validado de por qué un evento sucede, por qué algo cambia a través del tiempo o por qué unos estados o eventos covarían en el tiempo o en el espacio (Hedström 2005).<sup>2</sup>

En este nivel, conviene no sólo definir el concepto de explicación científica, sino que además deben ser aclaradas las condiciones bajo las cuales ella se produce y la relación que guarda con la construcción de modelos. El clásico trabajo de Hempel y Oppenheim (1948) cubre estos requerimientos. Para estos autores, una explicación científica consiste de dos componentes: un *explanandum*, u oraciones que describen el fenómeno a ser explicado, y un *explanans*, o el conjunto de oraciones que son invocadas para dilucidar el fenómeno. De acuerdo a Hempel y Oppenheim, la adecuada relación de ambos tipos de oraciones está determinada por la satisfacción de algunos requisitos: los explanans dan cuenta del explanandum si y sólo si cumplen condiciones tanto lógicas como empíricas (Bar 2001). Entre las condiciones lógicas, los autores destacan que, en primer lugar, la explicación debe tomar la forma de un argumento deductivo en el que el explanandum se sigue como una conclusión de las premisas en el explanans y, en segundo lugar, el explanans debe contener 'leyes causales' exigidas para la derivación del explanandum (sin las cuales la derivación del explanandum no sería válida). Por otro lado, entre las condiciones empíricas, los autores establecen que el explanans debe tener contenido verificable, es decir, las proposiciones incluidas en él deben ser posibles de comprobar mediante experimentos u observaciones.

En cuanto abstracciones selectivas, los modelos ofrecen un soporte útil para la construcción de explicaciones científicas, pues por un lado asisten en la definición, delimitación y operacionalización del explanandum mediante la simplificación del fenómeno y la formalización del mecanismo que lo produce, especialmente cuando el mecanismo es formalizado en algún lenguaje artificial, mientras que, por otro lado, los modelos garantizan el proceso de verificación de los supuestos implícitos en el explanans, mediante los juicios y conclusiones que el científico infiere al someter a prueba esos mismos supuestos en experimentos indirectos derivados del modelo empleado. En la práctica científica, tanto el explanandum como el explanans están acoplados mediante la formalización de

<sup>2</sup> Reconocemos que esta afirmación es polémica. Una buena parte de la tradición sociológica, desde Dilthey, Makkreel y Rodi (1991), ha construido sus límites aplicando la distinción entre '*Verstehen*' (comprensión) y '*Erklärung*' (explicación), señalando insistentemente que el dominio propio de la sociología y de las ciencias humanas en general está dado por la primera, siendo la segunda una pretensión exclusiva de las 'ciencias de la naturaleza'. Sus defensores no cuestionan la excesiva centralidad del 'sujeto' que supone esa distinción, ni menos parecen estar enterados de los innumerables cuestionamientos que desde la filosofía, de Heidegger (1997) a Derrida (1985), pasando por Foucault (1997b), se le ha hecho a esta tradición hermenéutica. Frente a esta guerra entre las 'dos culturas' (Snow 1993), un enfoque transdisciplinario como el que permiten las ciencias de la complejidad parece ser una salida más productiva a esta vieja polémica.

modelos que consideran sólo algunos principios teóricos abstractos, trasladándolos al ámbito de la experimentación y de la percepción.

No obstante la utilidad de la propuesta de Hempel y Oppenheim, esta debe ser seguida con precaución. Así, por ejemplo, las condiciones empíricas de inteligibilidad y verificación exigidas al explanans promueven la idea de que el principal rol de la modelación es derivar de las teorías tantas implicancias sujetas a confirmación empírica como sea posible (Ahrweiler y Gilbert 2005). Esta posición, sin embargo, no es más que una ilusión, pues las teorías nunca están determinadas por la información recogida de la observación. Por el contrario, es gracias a las distinciones y supuestos teóricos que toda observación y distinción del mundo es viable. Como afirman Kuhn (1962) y Feyerabend (1984), la observación científica está siempre ‘cargada de teoría’, siendo, de hecho, esta última la que determina los objetos que observamos y el estatuto que les asignamos. Tal como lo sugirió Quine (1977): si es posible construir dos o más teorías incompatibles basadas en el mismo conjunto de información experimental, la elección entre estas teorías no puede depender de ‘hechos empíricos’. De acuerdo a Quine, no existe un procedimiento único para establecer una relación entre teoría y datos de una manera lógica exclusiva y, por lo mismo, no existe un modo único en que podamos examinar la plausibilidad de una hipótesis determinada solamente prestando atención a la ‘realidad’.

Por lo anterior, parece más fructífero entender tanto explanandum y explanans como rendimientos internos de la práctica científica, y en especial de las teorías científicas. La distinción entre explanandum y explanans, en la fórmula de Mascareño, es una operación propia “de la teoría respecto de una temática externa pero que se constituye como temática internamente. Es decir, en la vinculación de explanans y explanandum se define el mundo que la teoría observa” (Mascareño 2008:223). Así, estrictamente hablando, los modelos nunca remiten, al menos directamente, a la ‘realidad exterior’. Ellos son más bien ‘agentes autónomos’, en la fórmula de Aydinonat (2007) que *median* entre la teoría y el mundo, o en la fórmula más constructivista de Knuuttila (2005), los modelos son ‘entidades independientes’ porque no tienen ningún lazo representacional preestablecido con el ‘mundo real’.

Por cierto, desde una perspectiva filosófica se puede poner en cuestión la utilidad epistemológica y el valor científico de unos modelos que no tienen una relación directa con la realidad. En un nivel aun más fundamental, si afirmamos que los modelos nos ayudan a construir explicaciones causales sobre los fenómenos que observamos y buscamos comprender, ¿podemos, sin caer en una paradoja lamentable, al mismo tiempo sostener que ellos no tienen contacto con la realidad? En otras palabras: ¿cómo puede una entidad que no tiene contacto con el mundo servirnos para hacernos una idea (¿científica!) sobre cómo funciona ese mundo? Al revisar con atención lo dicho hasta ahora, tal paradoja es, sin embargo, sólo aparente. De la misma manera que para Foucault un enunciado verdadero debe habitar en un ‘régimen de verdad’ (Foucault 1997a), todo

modelo científico está, por definición, acoplado a las estructuras propias de la ciencia moderna, tales como las teorías establecidas, los métodos estandarizados y las comunidades consolidadas. En definitiva, los modelos responden a las exigencias internas del sistema científico. A través de estos elementos la ciencia define el mundo que observa, y lo hace como una operación interna. Por lo mismo, el valor científico de un modelo se desprende de las comunes (aunque sofisticadas) estructuras de la ciencia y su práctica. Siguiendo a Ahrweiler y Gilbert (2005), la evaluación efectiva de un modelo científico se produce al responder preguntas tales como: ¿aceptan otros científicos los resultados del modelo como coherentes con el conocimiento existente?; ¿otros científicos usan sus resultados (o el modelo mismo) para apoyar su trabajo? ¿El modelo o sus resultados inspiran el trabajo de otros científicos? En síntesis, la construcción de modelos es entonces un elemento de la práctica científica que permite su clausura operativa; con ellos la ciencia formula y prueba las explicaciones causales sobre el mundo que la misma ciencia construye.

## Modelamiento estadístico

En las ciencias sociales existe una larga tradición de modelamiento, siendo las técnicas de tipo estadístico las más usadas. Este tipo de modelamiento ha sido aplicado extensamente en la investigación empírica desde los años cincuenta del siglo pasado (Lazarsfeld 1961). Es en gran medida debido a este legado que la modelación estadística es tan común en la disciplina y hoy cualquier sociólogo en formación es entrenado en su uso. En términos generales, un modelo estadístico toma la forma de una ecuación matemática en la que se asume que los conceptos de interés (operacionalizados y medidos a través de variables) están relacionados de algún modo (Tarling 2008). De esta manera, las explicaciones basadas en modelamiento estadístico identifican uno o más factores que parecen hacer una diferencia en la probabilidad de que el evento a explicar ocurra, por lo que este modelamiento es presentado como ‘análisis causal’ (cfr. Agresti y Finlay 2008). En palabras de uno de los padres de esta tradición: “Si tenemos una relación entre  $x$  e  $y$ ; y si para cualquier factor de prueba antecedente la relación parcial entre  $x$  e  $y$  no desaparece, entonces la relación original debe ser llamada una de tipo causal” (Lazarsfeld 1955:124-125).

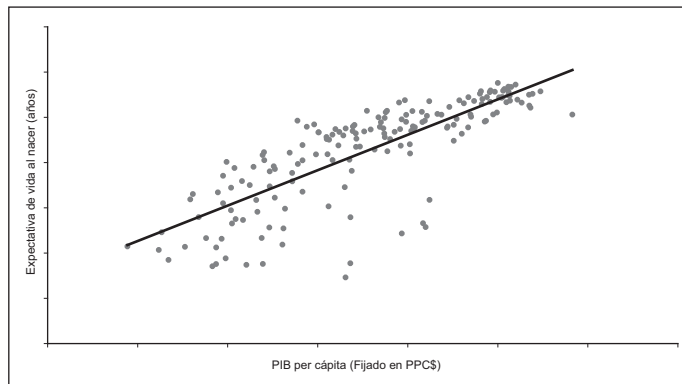
Dentro de los diferentes modelos estadísticos existentes, la *función lineal* es la expresión matemática más simple disponible para modelar la relación entre dos o más variables cuantitativas. Formalmente expresamos esta función del siguiente modo:

$$y_i = b_0 + b_1 x_{1i} \dots + b_k x_{ki} + \epsilon_i \quad (1)$$

En esta función,  $y$  es la variable dependiente que se busca explicar e  $y_i$  es el valor de la variable dependiente para el  $i$ -ésimo sujeto. Los valores  $x_{1i} x_{2i} \dots x_{ki}$  corresponden al

número de variables independientes del modelo (desde 1 a  $k$ ) y donde el valor de  $x_k$  varía entre los  $i$  sujetos. A su vez,  $b_1, b_2, \dots, b_k$  identifican los coeficientes, o parámetros, de las correspondientes variables independientes,  $x_{1i}, x_{2i}, \dots, x_{ki}$ . El término  $\mathbf{E}$  corresponde al error o residual del modelo, el cual recoge todos los factores de la realidad no controlables u observables (y que, por tanto, se asocian con el azar). Este último elemento confiere al modelo su carácter estocástico. Finalmente,  $\mathbf{E}_i$  es el residual para el  $i$ -ésimo sujeto (Tarling 2008). Este modelo, conocido como *modelo lineal general*, nos ayuda a cuantificar la tendencia y la magnitud en que una variable se correlaciona con otra (u otras). Si esta correlación es significativa, el mismo modelo nos permitirá realizar 'predicciones' sobre el valor que tomará la variable dependiente (o  $Y$ ), pues para cada valor de la o las variables independientes, el segundo miembro de la ecuación  $b_0 + b_1 x_{1i} + \dots + b_k x_{ki} + \mathbf{E}_i$  representa el promedio de la distribución condicional de  $Y$  para los sujetos teniendo esos valores.

Figura 1: Modelo de Regresión Lineal  
Expectativa de vida al nacer y PIB per cápita por Paridad de Poder de Compra (PPC\$)  
2007 (173 países)



Supongamos que sólo nos interesa observar la posible relación que existe entre dos variables, digamos, la esperanza de vida al nacer y la renta *per cápita* de un conjunto de países. Supongamos además que, luego de recolectar los datos apropiados, representamos gráficamente estas dos variables. El resultado de este ejercicio sería similar al presentado en la Figura 1. Como vemos, obtendríamos una alta correlación entre las variables que resulta doblemente interesante. En primer lugar, habríamos realizado una *simplificación*, pues nuestras dos dimensiones iniciales (esperanza de vida al nacer y la renta *per cápita*) se habrán reducido a una. Puesto que la correlación es muy elevada, podemos decir que



la línea misma de tendencia (esto es, una sola dimensión) representa prácticamente toda la información que antes aparecía en forma bidimensional. Eventualmente podríamos construir un modelo lineal de cien dimensiones y obtendríamos un resultado similar de reducción de complejidad. Así como nuestro modelo bidimensional pudo ser reducido a una sola dimensión, un modelo de cien medidas podría ser simplificado a un pequeño número de dimensiones. Sin duda, una parte de la información se perdería (del mismo modo que en el gráfico (a) de la Figura 1, al centrarnos en la línea de tendencia perdemos de vista la dispersión de los puntos respecto de la misma), pero estamos dispuestos a aceptar esa pérdida a cambio de una simplificación y las posibilidades que se nos abren para la *interpretación* de los datos. En segundo lugar, ahora podemos hacer una *inferencia* razonable sobre la *causa* de la correlación. En nuestro ejemplo, la correlación entre la esperanza de vida al nacer y la renta *per cápita* es muy intensa porque ambas son medidas parciales de un mismo fenómeno demográfico subyacente: el nivel de desarrollo humano de esos países (United Nations 2004, Sagar y Najam 1998). Este último actúa entonces –de acuerdo a lo que podemos ver en el modelo– como un factor causal y, por tanto, explicativo de la correlación existente entre esas dos variables.

El modelamiento estadístico presenta, no obstante, importantes limitaciones. La primera tiene que ver con la naturaleza misma de los fenómenos sociales. En efecto, la mayor parte de los procesos sociales no son lineales, involucran grados variables de heterogeneidad y describen dinámicas de cambio que es muy difícil entender mediante modelos estadísticos que optimizan soluciones promedio, establecen una regla lineal entre *inputs* y *outputs* y son estáticas (Castellani y Hafferty 2009, Miller y Page 2007).

Existe además un segundo problema, relacionado con el tipo de explicaciones que podemos construir con estos modelos. Las regularidades estadísticas son rara vez –si es que alguna– tan inequívoca y fácilmente interpretables en términos causales, pues la existencia y magnitud de la correlación no proporciona indicación alguna sobre la naturaleza de la causa. De hecho, la mayoría de las correlaciones que podemos encontrar no son causales (así, por ejemplo, la correlación entre la distancia del cometa Halley respecto de la Tierra y el alza del precio de los combustibles será positiva y muy marcada para los últimos treinta años). Existen innumerables casos de aplicación de modelos estadísticos en las ciencias sociales que han confundido correlación con causalidad, mermando el desarrollo de un corpus teórico riguroso (Gould 1999).<sup>3</sup> El modelamiento estadístico nos puede mostrar que una determinada asociación entre variables existe, pero difícilmente podrá decirnos el porqué de esa relación, o cuál es la lógica del proceso analizado (Hedström y Swedberg 1996). En nuestro ejemplo, parece razonable inferir un mecanismo causal que explica la naturaleza de la relación entre esperanza de vida al nacer e ingreso per cápita como una

<sup>3</sup> Como señala Hedström, “[a]unque la mayor parte de quienes usan modelamiento estadístico refieren a teorías en sus trabajos, ellos rara vez les prestan atención. La mayoría de las veces simplemente usan las teorías para justificar la inclusión de ciertas variables tomadas desde un conjunto de datos que a menudo es recolectado para fines completamente distintos a los declarados” (Hedström 2005:105).

expresión del desarrollo humano de los países medidos. Pero una afirmación como esta nunca puede basarse sólo en la matemática del modelo lineal general; es imprescindible contar con otros conocimientos teóricos relativos a la naturaleza demográfica, económica y sociológica de las medidas mismas y el mecanismo por el cual interactúan. Es este conocimiento teórico el que operacionalizamos en variables y luego testeamos mediante modelamiento estadístico.

Aunque el modelamiento estadístico ha estado asociado tradicionalmente con la forma estándar de construir y validar explicaciones causales, las explicaciones que podemos obtener del mero uso de correlaciones entre variables, *stricto sensu*, no nos conduce al establecimiento de explicaciones completas, pues estas últimas deben siempre especificar el mecanismo por el cual una variable influencia o produce a otra. En palabras de Kiser y Hetcher, el hacer explícitos los mecanismos que producen los fenómenos observados “es vital para las explicaciones, porque ellos indican qué variables debieran ser controladas con el fin de destacar las relaciones causales existentes. La discusión explícita de los mecanismos hace más difícil construir argumentos ad hoc y a menudo revela contradicciones en los argumentos que no serían aparentes en una lista de relaciones causales” (1991:5).

Todo lo anterior no significa que el modelamiento estadístico carezca de valor científico o resulte nocivo para el desarrollo de un corpus teórico sólido en las ciencias sociales. Muy por el contrario, es uno de los métodos de modelamiento más importantes y extendidos, que ha alcanzado grados de refinamiento en el procesamiento de datos muy sofisticados. Mediante su uso la sociología ha logrado enormes avances en la comprensión del fenómeno social, pues este modelamiento es una herramienta eficaz para la validación de teorías. Sin embargo, de lo que se trata es de situar los elementos en su secuencia correcta: el modelamiento estadístico es una forma de validar explicaciones y probar teorías; nunca es la explicación ni la teoría en sí misma.

## Modelamiento matemático

Aunque menos extendido, el modelamiento de tipo matemático tiene también una larga tradición en sociología (Coleman 1964, Sorensen 1978, Gill 2006, McElreath y Boyd 2007). Ampliamente usado en las ciencias físicas, biológicas, ingenierías y economía, es extremadamente útil para el modelamiento de sistemas dinámicos que, por definición, no describen comportamientos lineales. Esto ha permitido a la sociología ponerse en contacto con los desarrollos de esas ciencias y utilizar sus avances para el estudio de su propio objeto.<sup>4</sup> En comparación con los modelos estadísticos, el modelamiento matemático presenta diferencias fundamentales (Brown 2007). En primer lugar, los modelos estadísticos

<sup>4</sup> Muchos modelos provenientes desde la epidemiología, la biología y la física son crecientemente utilizados en el análisis de los problemas sociales. Para una visión sobre el estado actual de esta comunicación entre el modelamiento de fenómenos naturales y el de fenómenos sociales, cfr. Cho (2009).

son estocásticos, mientras que los modelos matemáticos son deterministas (aunque los procesos estocásticos pueden ser incorporados en modelos matemáticos tales como ecuaciones diferenciales estocásticas). En segundo lugar, los modelos estadísticos asisten al científico social en la prueba de teorías, mientras que los modelos matemáticos ayudan al investigador a hacer exploraciones teóricas y formalizar nuevas teorías (de hecho, para Sorensen la construcción de modelos que implementan razonamientos teóricos debiera ser la actividad central de la 'sociología matemática'; ver Sorensen 1978). Finalmente, los modelos estadísticos involucran la reducción de los datos y la información, mientras que los modelos matemáticos implican una expansión del conocimiento, por ejemplo, prediciendo patrones de comportamiento a partir de unas pocas condiciones iniciales.

Dentro de las técnicas de modelamiento matemático existentes, una de las más usadas es la basada en ecuaciones diferenciales; pilar del modelamiento de sistemas dinámicos (Brown 2007). Mediante esta técnica, un sistema, con sus propiedades y dinámica de cambio, es modelado usando un conjunto de ecuaciones con las que podemos derivar su comportamiento futuro desde su estado actual. Además, estas técnicas de modelamiento presentan una gran ventaja: son muy apropiadas para el modelamiento de explicaciones que involucran una distinción entre individuo y sociedad (o agencia y estructura) (Lane 1999). En efecto, mediante modelos de ecuaciones diferenciales podemos, como sostiene Hedström, "analizar el vínculo entre el individuo y la sociedad pues ellos exactamente describen los patrones sociales que son consecuencia de los supuestos a nivel individual sobre los cuales el modelo está basado" (2005:89).

Existen innumerables estudios de modelos de ecuaciones diferenciales aplicadas al comportamiento social. Uno de los más conocidos es el realizado por Coleman, Katz y Menzel (1957) para analizar la adopción de una nueva droga entre un grupo de médicos, estudio que marcó un hito en el campo sobre los procesos de 'difusión de innovaciones' (Burt 1987). Destacan también las formalizaciones matemáticas de teorías sociales, como el modelo de Land (1970) sobre la división social del trabajo de Durkheim y el de Martínez Coll (1986) sobre la emergencia del altruismo desde un 'estado de naturaleza' hobbesiano. En los últimos años, uno de los modelos basados en ecuaciones diferenciales más interesantes es el propuesto por Åberg (2000), el cual revisaremos en más detalle.

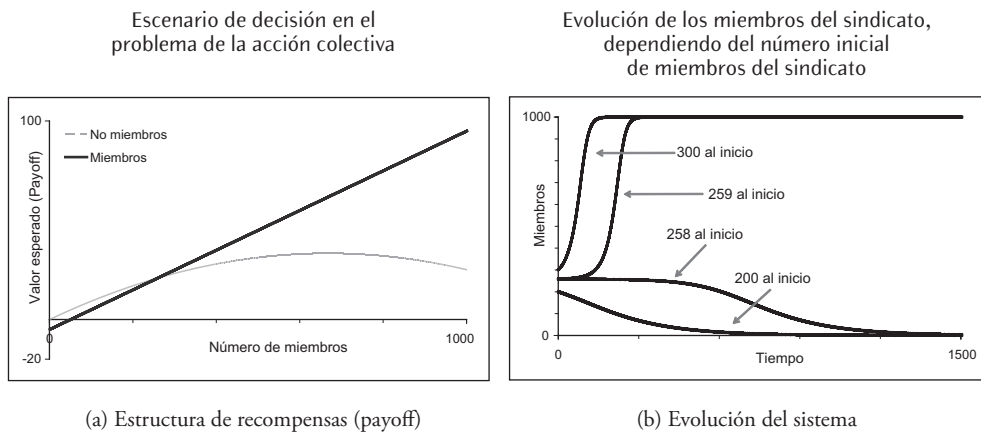
Basado en el pionero trabajo de Schelling (1978) sobre decisiones binarias y los llamados 'puntos clave' (de acuerdo a la traducción española del inglés *tipping-points*, Gladwell 2007), Åberg (2000) desarrolló un sencillo modelo de ecuaciones diferenciales para el análisis de cualquier situación caracterizada por un conjunto de actores que están permanentemente confrontados con una elección entre dos cursos de acción diferentes. El modelo de Åberg, para su correcta aplicación, supone, por un lado, que los individuos deciden la alternativa de acción que les resulta más beneficiosa y, por otro lado, que el fenómeno analizado satisface tres condiciones: (a) el conjunto de oportunidades para actuar que los individuos tienen consiste de dos alternativas; (b) el valor que para un individuo tiene elegir una acción depende de lo que los otros individuos eligen, y (c) la

velocidad con que los individuos cambian su acción depende de cuán deseable perciben cada alternativa. Existen innumerables fenómenos sociales en los que estas condiciones se cumplen, desde reportar un caso de corrupción en la oficina o hacer la vista gorda, hasta la decisión de unirse a un movimiento social o beneficiarse de sus resultados permaneciendo como un *free rider*.

Para estudiar estas decisiones y las dinámicas que ellas generan debemos establecer, en primer lugar, la estructura y relación de los beneficios esperados (o *payoffs*) para los dos cursos de acción y, en segundo lugar, debemos precisar el modo en que evoluciona el sistema social que emerge a partir de esa estructura de recompensas. Tomemos por ejemplo la decisión que enfrentan los individuos pertenecientes a una organización de mil trabajadores cuando deben decidir si unirse o no al sindicato.

El gráfico (a) de la Figura 2 establece la primera función, en la que el valor esperado de ser o no ser miembro del sindicato depende del número de los otros individuos que actualmente son miembros del mismo.<sup>5</sup> Como podemos ver, cuando el sindicato tiene pocos miembros es mejor no unirse a él, pues el beneficio esperado (o *payoff*) de quienes no

Figura 2: Modelo de Ecuaciones Diferenciales



<sup>5</sup> Por simplicidad, asumimos que el beneficio de los miembros del sindicato se incrementa linealmente y es igual a  $u_M(n(t)) = K \times n(t) - D$ , donde  $u_M(n(t))$  corresponde a la recompensa de los miembros del sindicato, como una función del número de miembros en el tiempo  $t$ ;  $K$  es un coeficiente que representa el cambio en la recompensa debido a cada miembro adicional;  $n(t)$  es el número de miembros del sindicato en el tiempo  $t$  y  $D$  es igual al costo de pertenecer al sindicato. La simulación numérica presentada en el gráfico (a) de la Figura 2 se basa en un  $K = 0,1$  y un  $D = 5$ . Por otro lado, la función de recompensa de los no sindicalizados considera un elemento de presión social de los sindicalizados y es asumida por ser igual a  $u_F(n(t)) = K \times n(t) - S \times n^2(t)$ , donde  $u_F(n(t))$  corresponde al beneficio como una función del número de miembros en el tiempo  $t$ ; y  $S$  es el coeficiente de presión social de los sindicalizados que se incrementa por el cuadrado del número de miembros sindicalizados. La simulación numérica presentada en el gráfico (a) de la Figura 2 usa un valor de  $S = 0,000075$ .

son miembros es mayor que el de quienes están sindicalizados. Sin embargo, tan pronto como el tamaño del sindicato excede un umbral crítico –el *punto clave*– entonces se hace más ventajoso unirse al sindicato antes que permanecer fuera de él. En el gráfico (a), el llamado ‘punto clave’ es donde ambas curvas se intersectan. Si este punto es alcanzado, entonces el beneficio esperado de ser miembro comienza a ser mayor que la alternativa, por lo que el número de sindicalizados aumenta. Por el contrario, si ese punto no es alcanzado, entonces el sindicato no tiene éxito y el número de sindicalizados disminuye, pues el beneficio de estar fuera del sindicato es mayor que el de estar adentro.

Sin embargo, no parece razonable asumir que todas las personas en la empresa se sindicalizarán de inmediato si el punto clave es alcanzado. Más bien, la magnitud de la diferencia en las recompensas establece cuán rápidamente un individuo se convierte a la alternativa más ventajosa. Así, en el modelo descrito, “la tasa de crecimiento del sindicato dependerá en parte del número de no sindicalizados que pueden ser reclutados y en parte de la diferencia en las recompensas” (Åberg 2000:196). Lo contrario es cierto cuando la diferencia en las recompensas es negativa. Desde luego, tanto las recompensas como el número de miembros y no miembros del sindicato cambiarán continuamente durante el proceso y alterarán la velocidad de cambio del sistema. Por tanto, el cambio en el número de los sindicalizados por unidad de tiempo es una función tanto de la diferencia en las recompensas como también del número de individuos que son susceptibles para cambiar su estado. En el modelo de Åberg todo esto es establecido por las siguientes ecuaciones diferenciales:

$$\frac{dn(t)}{dt} = \beta \times (u_M(n(t)) - u_F(n(t))) \times n(t) \quad (2a)$$

$$\text{si } (u_M(n(t)) - u_F(n(t))) \leq 0$$

$$\frac{dn(t)}{dt} = \beta \times (u_M(n(t)) - u_F(n(t))) \times (N - n(t)) \quad (2b)$$

$$\text{si } (u_M(n(t)) - u_F(n(t))) > 0$$

Donde:

- $n(t)$  = número de sindicalizados en el tiempo  $t$ ;
- $\beta$  = coeficiente de cambio que indica cuán rápido los individuos reaccionan a una diferencia en las recompensas;
- $u_M(n(t))$  = la recompensa asociada con ser miembro, dependiendo del número de miembros en el tiempo  $t$ ;

$$\begin{aligned}
 u_F(n(t)) &= \text{la recompensa asociada con no ser miembro, dependiendo del número} \\
 &\quad \text{de miembros en el tiempo } t; \\
 N &= \text{el número total de individuos, y} \\
 \frac{dn(t)}{dt} &= \text{el cambio en el número de miembros sindicalizados por unidad de tiempo.}
 \end{aligned}$$

Resolviendo las ecuaciones 2a y 2b numéricamente modelamos la evolución del grado de sindicalización a través del tiempo, pues mediante la solución matemática de estas ecuaciones estimamos el número de miembros sindicalizados como una función del tiempo. La ecuación 2a establece la salida de miembros del sindicato (en el tramo en que la diferencia de las recompensas es negativa, donde la mejor alternativa es estar fuera del sindicato), mientras que la ecuación 2b determina la entrada de los no sindicalizados al movimiento (en el tramo en que la diferencia de las recompensas es positiva, donde la mejor alternativa es estar sindicalizado). En el gráfico (b) de la Figura 2 podemos ver cómo el sindicato evolucionará dadas diferentes condiciones iniciales.<sup>6</sup> El umbral crítico que determina si el sindicato perdura o se extingue (con el conjunto de parámetros hipotéticos que estamos usando) está entre los 258 y 259 miembros iniciales. Tal como lo mencionamos, en ese umbral se produce el llamado ‘punto clave’, donde la acción de un solo individuo puede ser decisiva en el destino de un grupo social completo. En nuestro ejemplo, si el grupo inicial es de 258 miembros o menos, el sindicato no crecerá y después de un tiempo se extinguirá. Sin embargo, si logra reclutar al miembro 259, el sindicato podrá desarrollarse hasta incorporar a la totalidad de los trabajadores de la empresa.

El ejemplo analizado lo seleccionamos no debido a su realismo, sino porque la lógica modelada ocurre frecuentemente y porque destaca una forma de no-linealidad presente en muchos procesos sociales que puede ser modelada mediante ecuaciones diferenciales. Ellas son especialmente útiles para el modelamiento de procesos sociales que dependen del tiempo y son muy eficaces como herramientas analíticas para comprender cómo las acciones en el nivel micro producen diferentes resultados en el nivel macro. Pero su utilidad no es sólo teórica, pues una gran ventaja de los modelos de ecuaciones diferenciales es que existen métodos estandarizados para vincularlos con información empírica (ver Press et al. 2007), por ejemplo, mediante el uso de series de tiempo y un algoritmo de optimización adecuado para encontrar los parámetros de la ecuación que mejor se ajustan a los datos.

Sin embargo, el modelamiento de tipo matemático presenta también algunas limitaciones. La principal de ellas tiene que ver con los grados de heterogeneidad que pueden

<sup>6</sup> La simulación numérica del gráfico (b) de la Figura 2 se basa en un coeficiente de cambio  $\beta = 0,001$  y un  $N = 1.000$ . Las condiciones iniciales de los sindicalizados, o  $n(t)$ , son presentadas en el mismo gráfico como 300, 259, 258 y 200 miembros.

ser representados, pues los modelos de ecuaciones diferenciales asumen el sistema como un todo indiferenciado. Como en nuestro ejemplo, estos modelos suponen que todos los individuos realizan el mismo cálculo de recompensas y todos tienen acceso a la misma información (en este sentido, se trata de un modelo en donde los individuos gozan de información perfecta). Más importante en términos sociológicos es que en la mayoría de estos modelos los actores interactúan con otros en un modo completamente aleatorio y todos son igualmente influyentes o igualmente susceptibles de ser influenciados. Por cierto, es posible flexibilizar estos supuestos (ver Grunbaum et al. 2008), pero a condición de complejizar el modelo mismo en magnitudes que los hacen muy arcanos para quienes no tienen un dominio matemático sofisticado. Y aun para los más expertos el lenguaje matemático puede significar una camisa de fuerza que obliga a simplificar a tal extremo el modelo que muchas veces este termina siendo 'irreconocible' (Fiorina 1975). De agregar a los modelos matemáticos mayores grados de heterogeneidad, corremos el riesgo de volverlos intratables, tanto analítica como numéricamente. De hecho, los requerimientos de *tratabilidad* matemática agregan mayores restricciones a este tipo de modelos, poniendo en cuestión el objetivo general de los mismos como simplificaciones selectivas que aíslan y reproducen los principales mecanismos causales de los fenómenos estudiados, dado que los problemas que enfrentan los científicos sociales en el diseño de sus modelos se relacionan no sólo con lo 'irreal' de sus principios teóricos, sino que también con la 'modelo-dependencia' de los resultados.

Por lo mismo, los modelos matemáticos parecen ser útiles en un rango limitado de casos. A pesar de su mayor versatilidad en comparación con los modelos estadísticos, especialmente en el tratamiento de sistemas no-lineales, el modelamiento matemático tiene límites de flexibilidad, y esos límites resultan ser demasiado estrechos para el análisis de ciertos fenómenos sociales caracterizados por su alta complejidad.

Ninguna de las anteriores limitaciones debiera, no obstante, llevarnos a rechazar el modelamiento matemático. Por el contrario, estas técnicas son muy útiles en la formalización y validación de teorías. El uso del riguroso lenguaje matemático, con las posibilidades que abre para la demostración de teoremas y el establecimiento de soluciones analíticas a problemas con variables infinitas o en tiempo continuo, ha permitido la clausura operativa de la mayor parte de las ciencias de la naturaleza, y las ciencias sociales no son (ni debieran ser) la excepción. Los modelos matemáticos y las técnicas disponibles para analizarlos constituyen uno de los mejores 'laboratorios' que los investigadores tienen para probar sus teorías, y, si un determinado problema de investigación puede ser abordado satisfactoriamente mediante el uso de ecuaciones matemáticas, entonces técnicas de modelamiento alternativas pueden ser innecesarias. Sin embargo, cuando el fenómeno estudiado exige la construcción de modelos más complejos, la simulación computacional<sup>7</sup> resulta ser la mejor alternativa.

<sup>7</sup> Aunque en rigor pueden corresponder a cuestiones distintas, a lo largo de este artículo utilizaremos indistintamente, como sinónimos, simulación computacional, modelamiento computacional o simulación.



## Modelamiento computacional

Al igual que el modelamiento matemático, el de tipo computacional permite formalizar teorías mediante modelos de cambio en tiempo continuo y el análisis causal de patrones emergentes, aunque empleando para ello el lenguaje, no de la matemática, sino que de la programación. Como sostienen Taber y Timpone (1996), los modelos computacionales son teorías traducidas al lenguaje de las computadoras. Frente a las de tipo matemático, las técnicas de modelamiento computacional presentan importantes ventajas (Taber y Timpone 1996, Gilbert y Troitzsch 2005). En primer lugar, el lenguaje de programación es más expresivo, más versátil y menos abstracto que la mayoría de las técnicas de modelamiento matemático, por lo que conceptos teóricos que sería muy difícil expresar mediante ecuaciones matemáticas pueden ser representados computacionalmente de un modo directo. En segundo lugar, los programas computacionales permiten tratar fácilmente procesos paralelos sin la detallada especificación del orden que los sistemas de ecuaciones requieren. En tercer lugar, los programas computacionales son modulares, de modo que modificaciones mayores pueden ser hechas en una parte del código sin la necesidad de cambiar otras partes del mismo. Finalmente, es relativamente sencillo construir simulaciones que incluyen agentes heterogéneos, con diferentes recursos y capacidades, mientras que esto es muy difícil hacerlo mediante modelos matemáticos.

Pese a lo anterior, las diferencias entre modelamiento matemático y modelamiento computacional no son fundamentales; se trata de ventajas prácticas en favor de la segunda que resultan ser importantes en términos de construcción de teorías y validación de explicaciones causales de fenómenos complejos. En principio, para todo algoritmo o procedimiento computacional efectivo existe una ecuación matemática a la cual aquel puede ser traducido, pues, de acuerdo a la tesis Church-Turing (Church 1936, Turing 1936), toda computación puede ser implementada por una Máquina de Turing. Y es un resultado central de la lógica y de la computación que para cada Máquina de Turing existe una correspondiente Función Recursiva Parcial que la expresa (Rogers 1987). Siguiendo este razonamiento se puede afirmar que cualquier simulación computacional *podría* ser expresada en un conjunto explícito de ecuaciones matemáticas. La pregunta es: ¿puede alguien expresar y resolver analíticamente esas ecuaciones, o siquiera encontrar numéricamente su equilibrio (si este existe)? La respuesta no es clara, puesto que si las ecuaciones deben representar grandes poblaciones de agentes discretos y heterogéneos coevolucionando sobre un espacio con el cual interactúan, no resulta obvio cómo formular dichas ecuaciones ni tampoco, de ser formuladas, cómo solucionarlas. Como sostiene Epstein, “para cierta clase de problemas [...] se puede probar rigurosamente que la simulación computacional es lo mejor que uno puede hacer en principio” (1999:51). En cualquier caso, el asunto principal no es si efectivamente existen ecuaciones que expresan una simulación computacional, sino qué forma de representación (ecuaciones matemáticas o programas computacionales) resulta ser más práctica y explicativa. Para



todos, incluidos los más adeptos matemáticos, una representación mediante funciones recursivas sería irreconocible como modelo de un fenómeno social, mientras que su equivalente computacional resulta inmediatamente inteligible como tal.

Debido a su versatilidad, transparencia, interactividad y las múltiples posibilidades que abre para la 'experimentación teórica' (Edmonds y Hales 2005), la simulación computacional ha gozado de un creciente interés por parte de los científicos sociales en general y los sociólogos en particular. Muy relevante ha sido su potencial para asistir en el descubrimiento y formalización de explicaciones causales, pues los científicos sociales pueden construir modelos que se focalizan en algún fenómeno del 'mundo social' y establecen las consecuencias de sus supuestos y teorías sobre ese fenómeno en una 'sociedad artificial' (Gilbert y Conte 1995, Epstein y Axtell 1995, Sawyer 2003). Con estas técnicas de modelamiento ellos pueden tomar sus teorías, usualmente expresadas en lenguaje natural, y formalizarlas en algoritmos computacionales. Esto asegura que la explicación propuesta no es ambigua y es completa. Tal como sostienen Gilbert y Troitzsch, este "proceso de formalización, que involucra ser preciso sobre lo que la teoría indica y estar seguros de que es completa y coherente, es un campo disciplinar muy importante. En este sentido, la simulación computacional tiene un rol similar en las ciencias sociales al de las matemáticas en las ciencias físicas" (2005:5). Sus resultados han sido de tal importancia que hoy estas técnicas de modelamiento han dado inicio a un nuevo subcampo disciplinar: la sociología computacional (Gilbert y Troitzsch 2005, Castellani y Hafferty 2009), enmarcada en un nuevo programa de investigación conocido como 'ciencia social generativa' (Epstein y Axtell 1995, Cederman 2005). ¿En qué consiste este nuevo programa de investigación y por qué está asociado a la simulación computacional?

Fueron Epstein y Axtell quienes en 1995 propusieron un programa generativo de investigación en ciencias sociales, reconociendo a la simulación computacional como el principal método científico de este nuevo programa. Su principal argumento refería a las posibilidades que la simulación computacional (o, en sus palabras, 'la construcción de sociedades artificiales') ofrece para construir explicaciones científicas. El programa de investigación de la simulación computacional asumió como asunto de primera importancia el problema de la explicación en ciencias sociales, en particular, la pregunta por aquello que en ciencias sociales debe ser asumido como una explicación estándar. Su 'llamado a las armas' fue el siguiente:

¿Qué constituye una explicación de un determinado fenómeno social? Quizás un día las personas interpreten la pregunta '¿puede usted explicarlo?' como '¿puede usted *generarlo*?' El modelamiento de sociedades artificiales nos permite 'generar' las estructuras sociales *in silico*, demostrando que ciertos conjuntos de micro-especificaciones son *suficientes para generar* los macrofenómenos de interés [...] Nosotros podemos, desde luego, usar el modelamiento estadístico para probar el vínculo entre las estructuras observadas y esas que generamos.

Pero la habilidad para generarlas [...] es lo realmente nuevo. De hecho, es esto lo que ofrece la perspectiva de una nueva ciencia social; una de tipo generativa. (Epstein y Axtell 1995:20)

Cinco años después, el mismo Epstein perfiló la noción de una ‘explicación generativa’. De acuerdo a este autor, para el ‘generativista’ explicar la emergencia de una regularidad social macroscópica requiere que uno responda la siguiente pregunta: ¿cómo las interacciones locales de agentes heterogéneos y autónomos pueden generar las regularidades observadas? (Epstein 1999:41). De este modo, dado cierto explanandum macroscópico o emergente, de lo que se trata es de simular y establecer experimentalmente su causa desde el nivel inferior (*from the bottom up*), explicitando lo que algunos llaman una ‘micro-especificación’ (Epstein) o lo que otros denominan ‘mecanismo generativo’ (Hedström). Por tanto, en el programa generativo el valor científico de una explicación está dado por la especificación coherente y efectiva (i.e., computacional) de un mecanismo y la reproducción de un patrón emergente donde aquel mecanismo está involucrado (Cederman 2005). Para Hedström, una teoría sociológica es robusta sólo cuando ofrece explicaciones causales que identifican claramente los mecanismos sociales que están a la base del explanandum, entendiendo por estos mecanismos “una constelación de entidades y actividades que están organizadas de modo tal que ellas regularmente traen consigo un particular tipo de resultado” (Hedström 2005:33). Consecuentemente, agrega, “nosotros explicamos un determinado fenómeno social haciendo referencia al mecanismo social por el cual tal fenómeno es producido” (Hedström 2005:33). Si una teoría no presenta esta microespecificación o mecanismo generativo para los fenómenos sociales que aborda, entonces ella no es robusta, pues no ofrece en sentido estricto explicación causal alguna. Epstein es muy claro en este punto cuando formula el lema del programa generativo: “*if you didn’t grow it, you didn’t explain its emergence*”, o en la notación más formal de la lógica de primer orden:

$$(\forall x)(\neg Gx \supset \neg Ex) \quad (3)$$

La simulación computacional de procesos sociales ofrece diversas técnicas de modelamiento que pueden ser utilizadas para la formalización de explicaciones generativas (para una visión panorámica ver Gilbert y Troitzsch 2005). Entre las más utilizadas destaca la simulación social basada en agentes (*agent-based simulation*). Esta técnica ha tenido un enorme impacto en las ciencias sociales en los últimos diez o quince años (aunque sus orígenes están situados fuera de ella, en el pionero trabajo de Von Neumann 1967) sobre agentes autorreplicantes, en los años cuarenta, y el posterior trabajo de Conway (Gardner 1970) conocido como ‘el juego de la vida’ en los setenta. Su uso se consolidó después de los trabajos de Epstein y Axtell (1995) para estudiar

una amplia gama de patrones sociales en su modelo de agentes *sugarscape* y el trabajo de Axelrod (1997b) para analizar la evolución de la cooperación mediante poblaciones de agentes computacionales. Sin embargo, como método de investigación social, esta técnica tiene como antecedente el trabajo de Schelling (1971) sobre segregación residencial (aunque Schelling utilizó un tablero de ajedrez y fichas para modelar sus agentes; no un computador).

En términos generales, la simulación basada en agentes es un método computacional que permite al investigador crear, analizar y establecer experimentos indirectos para probar supuestos teóricos en poblaciones de agentes autónomos que interactúan localmente en un entorno (Gilbert 2007). Este método analítico resulta especialmente adecuado para modelar la emergencia de patrones estables o dinámicos desde el nivel inferior o de los individuos. En este caso, el procedimiento experimental consiste en situar una población de agentes autónomos y heterogéneos en un entorno espacial relevante, permitiéndoles interactuar de acuerdo a reglas locales simples que producen los patrones emergentes (Epstein 2007). Un agente es la representación simulada de una unidad de análisis autónoma (un individuo, una organización o un sistema social) con capacidad de percibir y reaccionar a su entorno (Pavón et al. 2008). Por tanto, en el caso de la simulación basada en agentes, el mecanismo generativo es expresado como un conjunto de reglas de interacción a nivel local que produce, en el transcurso del tiempo simulado, la emergencia del fenómeno a explicar. Existen innumerables ejemplos de simulaciones sociales basadas en agentes. En este artículo revisaremos una muy simple, construida en conjunto con mis colegas del Centre for Research in Social Simulation, University of Surrey (Marchione et al. 2009).

Este modelo fue desarrollado con el fin de entender el proceso de generación y difusión de léxicos simples (asociaciones entre una palabra y un significado) en una población de agentes autónomos que interactúan localmente en un entorno compuesto de objetos. El modelo tomó como punto de referencia algunas hipótesis establecidas en la teoría comunicacional de Luhmann (1998,1990) pues, como pocas, ella ofrece un detallado mecanismo generativo.

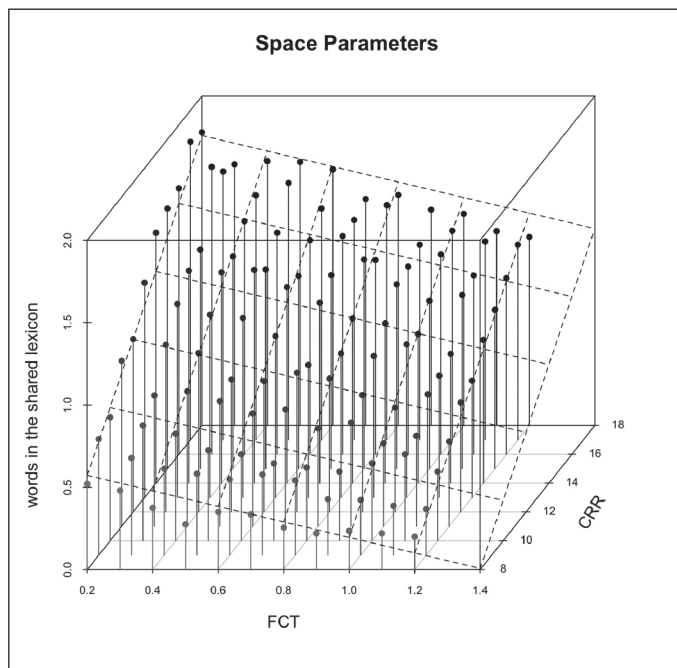
En este experimento, el mecanismo comunicacional que identifica Luhmann (i.e., las tres cifras de todo evento comunicativo) fue replicado en una comunidad de agentes autónomos situados en un entorno relevante. Dada la complejidad del proceso involucrado en la emergencia de los medios de comunicación simbólicamente generalizados (i.e., diferentes constelaciones de selectividad coordinada), nuestro modelo dejó de lado su análisis y se concentró en los dos primeros logros evolutivos de la comunicación: el lenguaje y los medios de difusión. El modelo tiene dos medidas de análisis: por un lado, la emergencia de léxicos o símbolos compartidos (i.e., que probabilizan la comprensión entre los agentes) es operacionalizada como asociaciones entre los objetos en el mundo y las 'palabras' que los agentes pueden crear para nombrarlos, las que emergen y evolucionan endógenamente en la población de agentes; (2) por otro lado, la lógica de los medios

de difusión (i.e., que expanden los límites de la comunicación más allá de los presentes) es operacionalizada como la capacidad de los agentes para alcanzar más ‘destinatarios’ en cada una de sus conductas de notificación. El objetivo de la simulación es analizar la configuración de agentes, establecida mediante las dos medidas de análisis descritas, que hace más probable la emergencia y estabilización de un tipo de comunicación simbólica entre los agentes.

La sociedad artificial está compuesta por cuarenta agentes, divididos en cuatro subgrupos de diez agentes cada uno. Cada agente tiene la capacidad de observar el entorno y su contenido, desplazarse a través de él y comunicar a otros agentes los objetos que observa en el entorno. Por otro lado, en el entorno están posicionados cuatro objetos diferentes. Los agentes tienen como objetivo construir un léxico común que asigna un nombre o palabra a cada uno de los objetos del mundo. Para ello, los agentes cuentan con un algoritmo de aprendizaje y son capaces de establecer conductas comunicativas recurrentes denominadas ‘juegos de lenguaje’, formalizadas en el campo de la lingüística computacional (Steels 1996, Steels y Kaplan 2002, Vogt y Coumans 2003). En nuestro modelo, un ‘juego de lenguaje’ supone un grupo de dos o más agentes, un emisor o *speaker* y uno o más destinatarios, quienes establecen ‘atención conjunta’ (Carpenter et al. 1998) sobre un mismo objeto –que pasa a ser así el tópico de la conducta comunicativa– y aprenden las asociaciones entre palabras y objetos mediante la llamada ‘Regla de Hebb’ (Song et al. 2000). El modelo se construyó usando el software NetLogo 4.0.4 (Wilensky 2009), mientras que todos los análisis estadísticos fueron implementados en R (R Development Core Team 2008).

La Figura 3 muestra un análisis de regresión lineal múltiple, usando mínimos cuadrados ordinarios, con la información obtenida luego de 121 arranques del programa. En todos los arranques, un mismo grupo de agentes sirvió como grupo experimental, al cual se le modificaron las medidas de análisis, mientras que los otros tres grupos fueron asumidos como de control, siempre con medidas fijas. En la Figura 3, FCT corresponde a una medida de análisis construida para estudiar la dinámica de los léxicos que emergieron durante las simulaciones, a saber, la frecuencia con que los agentes cambian el tópico de su comunicación a otro. Por otro lado, CRR es una medida de la capacidad que los agentes tienen para alcanzar destinatarios. La variable independiente del modelo de regresión fue el número de palabras que cada uno de los cuatro grupos de agentes logró imponer en el léxico compartido que emergió al final de cada arranque de la simulación. Como podemos ver en la Figura 3, la pendiente del plano es negativa para la variable FCT, mostrando que a mayor frecuencia en que los agentes cambian de objeto en sus comunicaciones, menor el número de palabras que impone en el léxico. La pendiente positiva para la variable CRR indica que a mayor capacidad de alcanzar otros agentes en cada comunicación, mayor el número de palabras que pueden imponer en el léxico compartido. Es claro entonces que la capacidad para alcanzar un mayor número de destinatarios y el establecer reiteradamente comunicaciones sobre un mismo objeto

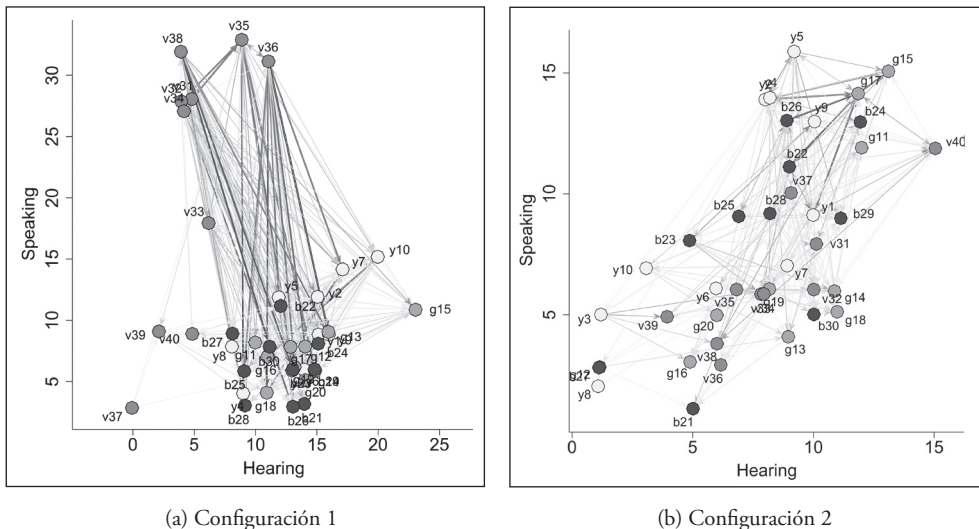
Figura 3: Resultado de las simulaciones



aumenta la probabilidad para un grupo en particular de generalizar un sistema de comunicación compartido entre toda la población de agentes.

Si la relación anterior es cierta, entonces un grupo de agentes dotados con capacidad para alcanzar muchos destinatarios (CRR) y entablar muchas conductas comunicativas sobre los mismos objetos (FCT) podría generalizar su propio léxico al resto de los agentes. Con una configuración como esta, obtenemos un resultado a estas alturas trivial: la comunidad de agentes que tiene más probabilidades de generar un léxico común es aquella con un alto CRR y un bajo FCT. Sin embargo, para comprender la dinámica del modelo analizamos la red de 'inter-conductas comunicativas' entre los agentes. La Figura 4 muestra el análisis de dos configuraciones experimentales para una sola asociación palabra-objeto.

Figura 4: Red de conductas comunicativas entre los agentes



La primera configuración, en el gráfico (a) de la Figura 4, muestra el resultado de un experimento donde la situación más favorable para generar un léxico está representada por los diez agentes de color violeta (señalados aquí en el circulillo relleno de gris oscuro junto a la abreviación 'v'), los cuales estaban dotados con un alto CRR y un bajo FCT en comparación con el resto de los agentes de la sociedad artificial, los que mantuvieron valores estándar iguales entre sí. En el gráfico, cada *link* es un 'juego de lenguaje' que conecta a un emisor con uno o varios destinatarios. El color del vínculo está en una escala de gris; un color más oscuro indica un mayor número de comunicaciones entre los agentes. Los ejes del gráfico señalan el número de emisores que cada agente ha escuchado (eje de abscisas) y el número de destinatarios al que cada agente ha comunicado (eje de ordenadas). Como podemos ver, en primer lugar, los agentes con características favorables para crear un sistema comunicacional escuchan a muy pocos otros agentes pero, por el otro lado, tienen una capacidad de comunicar a muchos agentes. En segundo lugar, el grupo violeta se mantuvo muy agrupado entre sí, pues los vínculos que conectan a los agentes de este grupo son más oscuros. Esto nos lleva a una interesante conclusión: los agentes con las características más favorables para producir un sistema de comunicación simbólica son quienes (1) prestan menos atención a las comunicaciones de agentes que no pertenecen a su grupo, (2) pueden alcanzar muchos destinatarios y (3) mantienen una intensa relación entre ellos. Expresar formalmente esta triple relación nos permite identificar la probabilidad que cada uno de los grupos tiene para producir y extender entre la población de agentes su propio léxico, lo cual formalizamos como:

$$p_g = \frac{(v_g * S_g)}{H_g} \quad (4)$$

Donde  $v_g$  corresponde a una medida de relación intragrupo (el número de conductas comunicativas entre los miembros del mismo grupo),  $S_g$  es la capacidad de cada grupo para alcanzar destinatarios de otros grupos y  $H_g$  es la capacidad de un grupo para ‘escuchar’ y aprender los léxicos de los agentes de otros grupos. Para el caso de la configuración experimental expresada en el gráfico (a) de la Figura 4, la probabilidad del grupo violeta para generalizar su propio sistema comunicacional fue muy alta,  $p_g = 98,42\%$ . La pregunta es si el modelo seguiría siendo efectivo para las configuraciones donde no hay un grupo con características favorables preestablecidas. Eso es lo que expresamos en el gráfico (b) de la Figura 4, donde todos los agentes de la sociedad artificial tenían las mismas características. En este caso, como podemos ver, no existe un patrón identificable y más bien todas las interacciones resultan caóticas. Sin embargo, a medida que las interacciones comunicacionales se van sucediendo, todas ellas aleatorias, la ecuación 4 nos permite determinar las probabilidades que cada grupo tiene para generalizar sus propio léxico entre la población de agentes. Las interacciones del gráfico (b) corresponden a un solo objeto. En ellas fue el grupo amarillo (aquí el circulillo en plomo, junto a la abreviación ‘y’) el que logró generalizar su asociación entre una palabra y ese objeto, siendo su probabilidad la más alta de los cuatro grupos ( $p_g = 56,69\%$ ).

En síntesis, la sociedad artificial creada valida alguna de las premisas luhmannianas sobre la emergencia y probabilización de la comunicación, mostrando que efectivamente la capacidad de alcanzar destinatarios es un factor clave para asegurar la continuidad de un medio comunicacional. Al mismo tiempo, muestra que un sistema comunicacional puede emerger endógenamente a partir de una comunidad de agentes autónomos, permitiendo así la coordinación simbólica entre ellos. Pero además muestra que, para hacer viable un léxico por sobre otros, es necesario que la dinámica de difusión la produzcan unos agentes que: a) experimentan una alta experiencia con los mismos objetos y con los mismos signos para referirse a ellos, b) que se muestren más bien inmunes o cerrados al uso de signos externos (esto es, de otros grupos), y c) que tienen una alta interrelación con los agentes de su mismo grupo. Esta dinámica puede no ser cierta para los orígenes del lenguaje, pero sí parece ser extremadamente importante para las comunidades científicas, sobre todo cuando ellas están emergiendo, momento en que la producción de conocimiento de estas comunidades reproduce la misma dinámica que muestra nuestra sociedad artificial: aislación respecto del resto de las comunidades científicas ya consolidadas, una fuerte conexión entre los investigadores al interior del nuevo campo y la habilidad para difundir sus nuevos conocimientos mediante publicaciones científicas,



congresos y otros medios. Otros ejemplos de esta misma dinámica en fenómenos sociales son los grupos juveniles, la emergencia de nuevos movimientos sociales o de protesta y las vanguardias artísticas.

Por cierto, la simulación computacional también presenta algunas limitaciones (cfr. McElreath y Boyd 2007). En primer lugar, y a diferencia de los modelos analíticos formalizados en el lenguaje de la matemática, las simulaciones computacionales no ofrecen pruebas; sólo nos proveen con una enorme colección de ejemplos que debemos analizar estadísticamente o con otras técnicas. En segundo lugar, existen riesgos con el análisis de sensibilidad de estos modelos. En efecto, una de las tareas más complicadas es explorar la sensibilidad de las simulaciones a los cambios en los valores de los parámetros. Los parámetros son cantidades que especifican supuestos del modelo para los diferentes arranques de la simulación —el tamaño de la población, el valor de los recursos, la velocidad de desplazamiento, las tasas de mutación, etc. Mientras en un modelo matemático el efecto de esos cambios puede ser leído directamente de las ecuaciones o usando varias técnicas de análisis, en la simulación computacional no existen expresiones análogas. En vez de eso, el modelador debe arrancar un gran número de simulaciones, variando los parámetros en todas las combinaciones (lo que se conoce como ‘analizar el espacio de los parámetros’). Para un pequeño número de parámetros esto no reviste mayor problema, pero las complicaciones escalan con el número de parámetros empleados. Por ejemplo, si asumimos un modelo de cuatro parámetros, cada uno con sólo 10 valores de interés, entonces para explorar el espacio de los parámetros requerimos 10<sup>4</sup> simulaciones. Si el modelo tiene algún efecto estocástico, entonces requeriríamos cien o mil veces más arranques con el fin de muestrear cada combinación y realizar análisis estadísticos que describan las tendencias generales de la simulación. Para ver los recursos de tiempo que exigiría un modelo como este, asumamos que tenemos un computador que tarda un segundo en arrancar una simulación de este tipo. Si requerimos cien simulaciones para cada combinación (cuatro parámetros con 10 valores de interés), entonces debemos arrancar la simulación 10<sup>6</sup> veces. A un segundo cada arranque, la sola exploración del espacio de los parámetros nos tomaría cerca de 12 días de computación continua. Y el resultado de estos arranques nos dejaría una enorme cantidad de información que aún debemos manejar e interpretar.

La dificultad de los programas computacionales para ser comunicados y replicados es otra limitación importante, pues produce un escollo para el trabajo de validación científica. Finalmente, la facilidad para construir simulaciones tienta muchas veces a los modeladores a agregar un gran número de parámetros a la simulación, llevando a complicados modelos de un mundo ya complicado. Cuando eso sucede, de una gran cantidad de parámetros pueden emerger resultados sorprendentes que, sin embargo, no podemos explicar, llevándonos así a una empresa científica que no se apega al principio básico de parsimonia: queríamos explicar un fenómeno complejo que no entendíamos y terminamos con un modelo sobre ese fenómeno que tampoco comprendemos.



Por supuesto, todos los problemas anteriores son, antes que limitaciones de la simulación computacional, alicientes para preferir siempre simulaciones más bien simples en la cantidad de parámetros empleados (un enfoque que ha sido denominado KISS, por las siglas del inglés *Keep it simple, stupid!*). Por otro lado, ya existe una literatura metodológica que muestra caminos para lidiar con los problemas de construcción, análisis del espacio de los parámetros y validación de las simulaciones (Axelrod 1997a, Axtell et al. 1996, Bianchi et al. 2007, Edmonds y Moss 2005, Gilbert y Terna 2000, Louie y Carley 2008, Marks 2007, Oreskes et al. 1994). Parece ser cuestión de tiempo entonces para que la comunidad de 'simuladores sociales' defina estándares y procedimientos de validación comunes que permitan una práctica más parsimoniosa de la simulación computacional de procesos sociales.

## Conclusión

En este artículo hemos revisado la importancia que tiene la modelación en la práctica científica en general y de las ciencias sociales en particular, especialmente para la elaboración de explicaciones causales a los fenómenos que nos interesa comprender. Con ellos reducimos la complejidad del mundo a sus principales mecanismos causales y podemos realizar experimentos indirectos que son la base de nuevas inferencias y conclusiones. Además, ellos nos permiten formalizar las teorías que les sirven de soporte y establecen un mecanismo de verificación para probar tales teorías. Simplificación, inferencia, formalización y verificación son las características que dotan a los modelos como una poderosa herramienta para construir explicaciones causales, esto es, explicaciones que apuntan a los mecanismos generativos de un fenómeno en particular. Construcción de modelos y elaboración de explicaciones son los dos lados de un mismo proceso de producción científica, un proceso en el que la ciencia no necesariamente captura el mundo tal y como es, sino que más bien lo define y reconstruye (después de todo, tal como sostiene von Glasersfeld, de la realidad sólo podemos saber aquello que no es; ver Watzlawick 1984). Así, los modelos son objetos independientes del mundo exterior, entidades inventadas que soportan una serie de inferencias y que están sometidos a las exigencias estructurales de su campo: son una operación interna de la ciencia.

Lo mismo es cierto para las ciencias sociales, las que también construyen modelos y con ellos definen el mundo que observan. En el caso de la sociología, en este artículo hemos revisado tres técnicas de modelamiento: de tipo estadístico, matemático y computacional. Cada una de ellas ofrece ventajas, implica ciertos compromisos epistemológicos y también padece de algunas limitaciones. El modelamiento estadístico es una de las técnicas más extendidas para construir explicaciones y, sin duda, una de las más exitosas. Sin embargo, buena parte de los modelos de este tipo rara vez prestan atención seria a las teorías sociales sobre los fenómenos de interés. Los modelos estadísticos, por

sí mismos, no entregan una explicación causal (sólo una de tipo descriptiva). Además, estos modelos rara vez se ajustan a la naturaleza dinámica y no lineal de los fenómenos sociales. Las técnicas estadísticas de modelamiento son muy importantes para validar explicaciones causales, aunque no para construirlas.

Por otro lado, las técnicas de modelamiento matemático parecen superar varias de las limitaciones del modelamiento estadístico. Ellas son más flexibles y adecuadas para describir fenómenos que cambian a lo largo del tiempo. Además, emplean el rico y altamente estilizado lenguaje matemático, lo que ofrece una interpretación directa y no-ambigua al mecanismo causal que el científico social expresa mediante un conjunto de ecuaciones; las cuales, con el conocimiento adecuado, pueden, además, ser resueltas analíticamente. La limitación del modelamiento matemático es, sin embargo, práctica: las ecuaciones pueden fácilmente volverse intratables si se les agregan demasiados parámetros, por lo que el científico que emplea esta técnica está siempre obligado a simplificar el fenómeno estudiado, sacrificando con ello la posibilidad de comprar punto por punto los términos de la ecuación con los elementos del fenómeno a explicar. Asimismo, el lenguaje matemático requiere competencias técnicas que muchas veces están más allá del dominio del sociólogo, razón por la cual muchos de los modelos matemáticos sobre fenómenos sociales existentes son realizados principalmente por científicos con un gran entrenamiento previo en la materia. Finalmente, dados todos estos inconvenientes prácticos, el modelamiento matemático no puede lidiar con fenómenos que implican altos grados de heterogeneidad (usualmente, lo propio de los fenómenos sociales). Una ecuación matemática no distingue entre niveles (social e individual) o espacio y todos los agentes son asumidos por tener un comportamiento homogéneo.

La última técnica de modelamiento que revisamos en este artículo, la simulación computacional, ofrece importantes ventajas por sobre el modelamiento matemático. Ventajas prácticas, como la mayor expresividad, versatilidad y menor abstracción del lenguaje computacional en comparación al matemático. Además, los programas computacionales son modulares, lo que permite al científico social realizar cambios en algunas partes del código sin necesidad de reescribir todo el programa. Pero una de las mayores ventajas del modelamiento computacional es la posibilidad de construir modelos con altos grados de heterogeneidad. Esto posibilita, por ejemplo, diferenciar entre niveles de complejidad (individual, organizacional, social) o trabajar con agentes que poseen recursos y capacidades diversas, algo muy difícil de hacer mediante el modelamiento matemático. Por lo anterior, las simulaciones computacionales, especialmente aquellas basadas en agentes, ofrecen un excelente entorno para construir explicaciones causales basadas en mecanismos generativos, esto es, establecer las microespecificaciones que producen los fenómenos sociales emergentes que nos interesa comprender. Por cierto, la simulación computacional también presenta algunas limitaciones, entre ellas están la dificultad para explorar su sensibilidad a los cambios en los valores de los parámetros, la dificultad aún presente para comunicar y replicar las simulaciones computacionales

—con lo que se pone en riesgo una parte importante de la práctica científica (repetir los experimentos)— y la tentación que enfrentan los ‘simuladores sociales’ para sobre-especificar sus modelos con un gran número de parámetros, los que aumentan la complejidad del modelo hasta hacerlo, en no pocas ocasiones, incomprensible.

Con todo, es seguro que en poco tiempo la simulación computacional se consolidará como una técnica más de modelamiento dentro del abanico de opciones que los científicos sociales tienen actualmente. De hecho, la división social del trabajo en el campo de las ciencias sociales, y de la sociología en particular, ya se está estructurando en torno a las técnicas de modelamiento que hemos revisado. Con ello, la teoría de la sociedad no desaparece, ni menos se supedita a una posición de dependencia respecto de otros dominios técnicos. Por el contrario, aun creyendo que el carácter científico de la sociología se encuentra estrictamente asociado al desarrollo de modelos explicativos que formalicen mecanismos causales, esto no se traduce en un imperialismo de las técnicas de modelamiento, sean de tipo estadístico, matemático o computacional, por sobre la teoría de la sociedad. Si el corpus teórico que tenemos es deficiente, este no será mejorado por modelos, por muy extravagantes o imaginativos que sean. Ningún modelo volverá rica una teoría pobre. Así las cosas, lo que afirmó hace ya casi cinco décadas James Coleman sobre el modelamiento matemático —aplicable también para el modelamiento estadístico y computacional— sigue siendo hoy cierto:

Muy a menudo la sociología matemática es discutida como un fin en sí misma o para sí misma, antes que en el contexto de los problemas con los que la sociología ha estado típicamente relacionada. La matemática es una herramienta, y si ésta es usada como una herramienta útil por la sociología debe serlo en completo reconocimiento del estado en el que la sociología se encuentra a sí misma hoy día. (Coleman 1964:4)

Recibido julio 2009  
Aceptado noviembre 2009

## Referencias bibliográficas

- Åberg, Yvonne, 2000. "Individual Social Action and Macro Level Dynamics: A Formal Theoretical Model". *Acta Sociologica* 43(3), 193-205.
- Agresti, Allan, Barbara Finlay, 2008. *Statistical Methods for the Social Sciences*. 4th ed. Upper Saddle River, NJ: Pearson Education.
- Ahrweiler, P., Nigel Gilbert, 2005. "Caffè Nero: the Evaluation of Social Simulation". *Journal of Artificial Societies and Social Simulation* 8(4). Available at: <http://jasss.soc.surrey.ac.uk/8/4/14.html> [agosto 2009].

- Axelrod, Robert, 1997a. "Advancing the art of simulation in the social sciences". *Complexity* 3(2), 16-22.
- , 1997b. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. New Jersey: Princeton University Press.
- Axtell, Robert, Robert Axelrod, Joshua M. Epstein, Michael D. Cohen, 1996. "Aligning simulation models: A case study and results". *Computational & Mathematical Organization Theory* 1(2), 123-141.
- Aydinonat, N. Emrah, 2007. "Models, conjectures and exploration: an analysis of Schelling's checkerboard model of residential segregation". *Journal of Economic Methodology* 14(4), 429-454.
- Bar, Aníbal, 2001. "La explicación como producto lógico o como producto de la praxis". Cinta de Moebio 11. Disponible en <http://rehue.csociales.uchile.cl/publicaciones/moebio/11/bar.htm> [agosto 2009].
- Bianchi, Carlo, Pasquale Cirillo, Mauro Gallegati, Pietro Vagliasindi, 2007. "Validating and calibrating agent-based models: A case study". *Computational Economics* 30(3), 245-264.
- Brown, Courtney, 2007. *Differential Equations: A Modeling Approach: 150 (Quantitative Applications in the Social Sciences): A Modeling Approach*. California: Sage Publications, Inc.
- Burt, Ronald S., 1987. "Social contagion and innovation: Cohesion versus structural equivalence". *The American Journal of Sociology* 92(6), 1287-1335.
- Carpenter, Malinda, Katherine Nagell, Michael Tomasello, George Butterworth, Chris Moore, 1998. "Social cognition, joint attention, and communicative competence from 9 to 15 months of age". *Monographs of the Society for Research in Child Development* 63(4), i-174.
- Castellani, Brian, Frederic William Hafferty, 2009. *Sociology and Complexity Science*. Berlin/Heidelberg: Springer.
- Cederman, Lars-Erik, 2005. "Computational models of social forms: Advancing generative process theory". *American Journal of Sociology* 110(4), 864-893.
- Cho, Adrian, 2009. "Ourselves and our interactions: The ultimate physics problem?". *Science* 325(5939), 406-408.
- Church, Alonzo, 1936. "An unsolvable problem of elementary number theory". *American Journal of Mathematics* 58(2), 345-363.
- Coleman, James, 1964. *Introduction to Mathematical Sociology*. New York: Free Press of Glencoe.
- Coleman, James, Elihu Katz, Herbert Menzel, 1957. "The diffusion of an innovation among physicians". *Sociometry* 20(4), 253-270.
- Derrida, Jacques, 1985. *La voz y el fenómeno: Introducción al problema del signo en la fenomenología de Husserl*. Valencia: Pre-textos.
- Dilthey, Wilhelm, Rudolf A. Makkreel, Frithjof Rodi, 1991. *Wilhelm Dilthey. Selected Works: Introduction to the human sciences*. Princeton: Princeton University Press.
- Edmonds, Bruce, David Hales, 2005. "Computational simulation as theoretical experiment". *Journal of Mathematical Sociology* 29, 209-232.

- Edmonds, Bruce, Cesareo Hernandez, Klaus G. Troitzsch, 2007. *Social Simulation: Technologies, Advances and New Discoveries*. Information Science Reference.
- Edmonds, Bruce, Scott Moss, S., 2005. "From KISS to KIDS – An 'Anti-simplistic' Modelling Approach". En *Multi-Agent and Multi-Agent-Based Simulation*, 130-144. Disponible en <http://www.springerlink.com/content/9yclq08wfxq4f3a7> [marzo 2009].
- Epstein, Joshua M., 1999. "Agent-based computational models and generative social science". *Complexity* 4(5), 41-60.
- , 2007. *Generative Social Science: Studies in Agent-Based Computational Modeling*. New Jersey: Princeton University Press.
- Epstein, Joshua M., Robert Axtell, 1995. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington D.C.: Brookings Institution, US/MIT Press.
- Feyerabend, Paul, 1984. *Against Method*. London: Verso.
- Fiorina, Morris P., 1975. "Formal models in Political Science". *American Journal of Political Science* 19(1), 133-159.
- Foucault, Michel, 1997a. *La arqueología del saber*. Buenos Aires: Siglo XXI.
- , 1997b. *Las palabras y las cosas: una arqueología de las ciencias humanas*. Buenos Aires: Siglo Veintiuno.
- Gardner, Martin, 1970. "Mathematical games: The fantastic combinations of John Conway's new solitaire game 'Life'". *Scientific American* 223(4), 120-123.
- Giere, Ronald N., 2004. "How models are used to represent reality". *Philosophy of Science* 71(5), 742-752.
- , 2006. *Scientific Perspectivism*. Annotated edition. Chicago: Chicago University Press.
- Gilbert, Nigel, 2007. *Agent-Based Models*. California: Sage Publications, Inc.
- Gilbert, Nigel, Andrew Abbott, 2005. "Introduction". *American Journal of Sociology* 110(4), 859-863.
- Gilbert, Nigel, Rosario Conte, 1995. *Artificial societies: computer simulation of social life*. 1st. ed. London: Routledge.
- Gilbert, Nigel, Jim Doran, 1994. *Simulating Societies: Computer Simulation of Social Phenomena*. London: Routledge.
- Gilbert, Nigel, Pietro Terna, 2000. "How to build and use agent-based models in social science". *Mind & Society* 1(1), 57-72.
- Gilbert, Nigel, Klaus G. Troitzsch, 2005. *Simulation for the Social Scientist*. Glasgow: Open University Press.
- Gill, Jeff, 2006. *Essential Mathematics for Political and Social Research*. Cambridge: Cambridge University Press.
- Gladwell, Malcolm, 2007. *El punto clave: cómo los pequeños cambios pueden provocar grandes efectos = The Tipping Point*. Madrid: Taurus.
- Godfrey-Smith, Peter, 2006. "The strategy of model-based science". *Biology and Philosophy* 21(5), 725-740.
- Gould, Stephen Jay, 1999. *The Mismeasure of Man*. New York: Econo-Clad Books, Div. of American Cos., Inc.

- Grunbaum, Daniel, Karen Chan, Elizabeth Tobin, Michael T. Nishizaki, 2008. "Non-linear advection-diffusion equations approximate swarming but not schooling populations". *Mathematical Biosciences* 214(1-2), 38-48.
- Hedström, Peter, 2005. *Dissecting the Social: On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.
- Hedström, Peter, Richard Swedberg, 1996. "Social mechanisms". *Acta Sociologica* 39(3), 281-308.
- Heidegger, Martin, 1997. *Ser y tiempo*. Santiago: Universitaria.
- Hempel, Carl G., Paul Oppenheim, 1948. "Studies in the logic of explanation". *Philosophy of Science* 15(2), 135-175.
- Kiser, Edgar, Michael Hechter, 1991. "The role of General Theory in Comparative-Historical Sociology". *The American Journal of Sociology* 97(1), 1-30.
- Knuuttila, Tarja, 2005. "Models, representation, and mediation". *Philosophy of Science* 72(5), 1260-1271.
- , 2009. "Isolating representations versus credible constructions? Economic Modelling in Theory and Practice". *Erkenntnis* 70(1), 59-80.
- Kuhn, Thomas. S., 1962. *The Structure of Scientific Revolutions*. Chicago: Chicago University Press.
- Land, Kenneth C., 1970. "Mathematical formalization of Durkheim's Theory of Division of Labor". *Sociological Methodology* 2, 257-282.
- Lane, David C., 1999. "Social theory and system dynamics practice". *European Journal of Operational Research* 113(3), 501-527.
- Lazarsfeld, Paul F., 1955. "Interpretation of statistical relations as a research operation". En P. F. Lazarsfeld & M. Rosenberg, eds. *The Language of Social Research*. New York: Free Press.
- , 1961. "Notes on the History of Quantification in Sociology-Trends, Sources and Problems". *Isis* 52(2), 277-333.
- Louie, Marcus A., Kathleen M. Carley, 2008. "Balancing the criticisms: Validating multi-agent models of social systems". *Simulation Modelling Practice and Theory* 16(2), 242-256.
- Luhmann, Niklas, 1990. *Essays on Self-Reference*. New York: Columbia University Press.
- , 1998. *Sistemas sociales*. Barcelona: Anthropos.
- Macy, Michael W., Robert Willer, 2002. "From factors to actors: Computational Sociology and Agent-Based Modeling". *Annual Review of Sociology* 28(1), 143-166.
- Marchione, Elio, Mauricio Salgado & Nigel Gilbert, 2009. "What did you say? Emergent Communication in a Multi-Agent Spatial Configuration". Proceedings of The European Social Simulation Conference 2009. Guildford, United Kingdom.
- Marks, Robert, 2007. "Validating simulation models: A general framework and four applied examples". *Computational Economics* 30(3), 265-290.
- Martínez Coll, Juan Carlos, 1986. "A bioeconomic model of Hobbes' 'state of nature'". *Social Science Information* 25(2), 493-505.
- Mascareño, Aldo, 2008. "Acción, estructura y emergencia en la teoría sociológica". *Revista de Sociología* 22, 217-256.

- McElreath, Richard, Robert Boyd, 2007. *Mathematical Models of Social Evolution: A Guide for the Perplexed*. Chicago: Chicago University Press.
- Miller, John H., Scott E. Page, 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Illustrated edition. Princeton University Press.
- Neumann, John Von, 1967. *Theory of Self-reproducing Automata*. Illinois: University of Illinois Press.
- Nowotny, Helga, 2005. "The increase of complexity and its reduction: Emergent interfaces between the Natural Sciences, Humanities and Social Sciences". *Theory, Culture & Society* 22(5), 15-31.
- Oreskes, Naomi, Kristin Shrader-Frechette, Kenneth Belitz, 1994. "Verification, validation, and confirmation of numerical models in the Earth Sciences". *Science* 263(5147), 641-646.
- Pavón, Juan, Millan Arroyo, Samer Hassan, Candelaria Sansores, 2008. "Agent-based modelling and simulation for the analysis of social patterns". *Pattern Recognition Letters* 29(8), 1039-1048.
- Press, William H., Saul A. Teukolsky, William T. Vetterling, Brian Flannery, 2007. *Numerical Recipes: The Art of Scientific Computing*. Cambridge: Cambridge University Press.
- Quine, Willard Van Orman, 1977. *Ontological Relativity and Other Essays*. Columbia: Columbia University Press.
- R Development Core Team, 2008. "R: A language and environment for statistical computing", Vienna, Austria. Disponible en <http://www.R-project.org> [noviembre 2009].
- Rogers, Hartley, 1987. *Theory of Recursive Functions and Effective Computability* New York: MIT Press.
- Sagar, Ambuj D., Adil Najam, 1998. "The human development index: a critical review". *Ecological Economics* 25(3), 249-264.
- Sawyer, R. Keith, 2003. "Artificial societies: Multiagent systems and the micro-macro link in Sociological Theory". *Sociological Methods Research* 31(3), 325-363.
- Schelling, Thomas C., 1971. "Dynamic models of segregation". *Journal of Mathematical Sociology* 1, 143-186.
- , 1978. *Micromotives and Macrobehavior*. New York: W. W. Norton.
- Snow, Charles Percy, 1993. *The Two Cultures*. Cambridge: Cambridge University Press.
- Song, Sen, Kenneth D. Miller, L.F. Abbott, 2000. "Competitive Hebbian learning through spike-timing-dependent synaptic plasticity". *Nat Neurosci* 3(9), 919-926.
- Sorensen, Aage B., 1978. "Mathematical models in Sociology". *Annual Review of Sociology* 4(1), 345-371.
- Steels, Luc, 1996. "Emergent adaptive lexicons". *From Animals to Animats* 4, 562-567.
- Steels, Luc, Frédéric Kaplan, 2002. "Bootstrapping Grounded Word Semantics". En Ted Briscoe, ed. *Linguistic Evolution Through Language Acquisition*. Cambridge: Cambridge University Press, 53-73.
- Taber, Charles S., Richard John Timpone, 1996. *Computational Modeling*. Illustrated edition. London: Sage Publications, Inc.



- Takahashi, Shingo, David Sallach, Juliette Rouchier, 2007. *Advancing Social Simulation: The First World Congress*, Berlin/Heidelberg: Springer.
- Tarling, Roger, 2008. *Statistical Modelling for Social Researchers: Principles and Practice*. London: Routledge.
- Turing, Allan M., 1936. "On computable numbers, with an application to the Entscheidungsproblem". *Proceedings of the London Mathematical Society* 42(2), 230-265.
- United Nations, 2004. *Human Development Report 2004*. Oxford: Oxford University Press.
- Vogt, Paul, Hans Coumans, 2003. "Investigating social interaction strategies for bootstrapping lexicon development". *The Journal of Artificial Societies and Social Simulation* 6(1). Disponible en <http://jasss.soc.surrey.ac.uk/6/1/4.html> [mayo 2008].
- Watzlawick, Paul, 1984. *The Invented Reality: How Do We Know What We Believe We Know?* New York: W. W. Norton & Co.
- White, Harrison C., 2000. "Parameterize!: Notes on mathematical modeling for Sociology". *Sociological Theory* 18(3), 505-509.
- Wilensky, Uri, 2009. "NetLogo. Center for Connected Learning and Computer-Based Modeling". Northwestern University, Evanston, IL. Disponible en <http://ccl.northwestern.edu/netlogo> [noviembre 2009].