

LA VERDAD Y EL ÉXITO DE LA CIENCIA*

(A propósito de un artículo de P. Kyle Stanford)

MANUEL COMESAÑA

¿En qué consiste lo que se suele llamar el éxito de la ciencia? Básicamente, en que algunas teorías científicas –esto es, algunos sistemas de hipótesis teóricas– han tenido un notable éxito observacional: a partir de ellas se ha obtenido un número considerable de consecuencias observacionales verdaderas e interesantes. Cuando este éxito es predictivo –además o en vez de explicativo– sirve de base a las aplicaciones tecnológicas del conocimiento científico, que son la manifestación más visible del éxito de la ciencia. Este éxito da lugar al argumento más importante (y tal vez el único) a favor del realismo científico, a saber, que sería inexplicable si no se admitiera que las teorías exitosas son verdaderas –parcial o aproximadamente verdaderas, ya que la verdad total es improbable en grado sumo–. Este argumento ha recibido varios nombres, el más pintoresco de los cuales es el de “argumento del milagro”, debido a la siguiente frase de Putnam: “el realismo es la única filosofía que no hace del éxito de la ciencia un milagro”¹.

Una fuente de dificultades para el argumento del milagro es la tesis de que todas las teorías tienen, o al menos pueden tener, rivales empíricamente equivalentes; peor aún: se sabe desde hace mucho que cualquier conjunto de datos puede ser explicado por infinitas teorías distintas, de las cuales es verdadera una sola (leer esto en un texto de Popper hizo que Chomsky se volviera innatista). Un caso especial de equivalencia empírica es la que se da entre una teoría más o menos verdadera y otra totalmente falsa, ejemplificado por la relación entre la teoría de Copérnico y una versión mejorada del sistema tolemaico. Si una teoría totalmente falsa puede tener las mismas consecuencias observacionales verdaderas que una teoría rival parcial o aproximadamente verdadera, parecería *prima facie* que no se necesita verdad teórica para explicar el éxito de la ciencia. Nótese que este argumento intenta prescindir de la verdad teórica pero no de la observacional, es decir, no pretende dar lugar a una concepción de la ciencia que prescinda por completo de la verdad, cosa que algunos han intentado, como por ejemplo Laudan en el libro *El progreso y sus problemas*².

Una manera de defender el argumento del milagro, y por lo tanto al realismo, frente a la tesis de la equivalencia empírica –muy probablemente la única manera– consiste en elaborar la siguiente intuición: las teorías

* Agradezco los útiles comentarios de Juan Comesaña.

¹ Putnam, *Mathematics, Matter and Method*, Cambridge, Cambridge University Press, 1975, p. 73.

² Laudan, *Progress and its Problems*, Berkeley, University of California Press, 1977.

falsas que tienen consecuencias verdaderas, o bien son parcialmente verdaderas, o bien contienen hipótesis *ad hoc*; en este último caso su promesa de éxito observacional futuro es inferior a la de una teoría verdadera. En efecto, sólo por casualidad una teoría falsa que salva los fenómenos gracias a hipótesis *ad hoc* dará lugar a predicciones exitosas nuevas, mientras que una teoría parcialmente verdadera lo hará —o, al menos, podrá hacerlo— de manera más o menos sistemática. Esto que acabo de decir depende de cuál sea el concepto de novedad predictiva que se esté manejando; si “El Sol va a salir mañana” se considera una predicción nueva —lo que seguramente estaría mal— el realista está perdido: Tolomeo formula esa predicción tan bien como Copérnico.

Elaborar satisfactoriamente la intuición mencionada es una tarea difícil, a la que Jarrett Leplin ha dedicado un libro³ muy interesante de doscientas páginas en el que desarrolla un análisis del concepto de predicción “novedosa” de una complejidad sin precedentes, según sus propias palabras (conviene decir “novedosa” en vez de “nueva” porque él, siguiendo a Lakatos, distingue respectivamente entre “*novel*” y “*new*”). Para que un resultado predicho por una teoría sea “novedoso” en el sentido de Leplin, se deben satisfacer dos condiciones, que —formuladas de manera vaga al solo efecto de ofrecer una primera aproximación— son las siguientes: a) *independencia*: el resultado no desempeña ningún papel esencial en la procedencia de la teoría, y b) *unicidad*: ninguna otra teoría puede predecirlo.

Esta estrategia realista supone admitir que una predicción exitosa tiene más fuerza confirmatoria que una explicación adecuada, esto es, supone tomar partido por la teoría histórica de la confirmación contra la teoría lógica, y con ello exige rechazar la idea de que el contexto de descubrimiento no puede influir en el contexto de justificación. En efecto, que una consecuencia de una teoría sea algo que la teoría explica o predice depende de la construcción o desarrollo de la teoría. Por ejemplo, la teoría de Torricelli explicó por qué es limitada la eficacia de las bombas y predijo que la presión atmosférica sostendría una columna de mercurio de aproximadamente setenta centímetros. Seguramente este último resultado no sería “novedoso” en el sentido de Leplin, pero sirve para ilustrar de manera sencilla la diferencia intuitiva entre explicación y predicción en lo concerniente a fuerza probatoria. La explicación tuvo cierto valor justificatorio —sobre todo porque ninguna otra teoría explicaba la limitación mencionada—, pero el éxito de la predicción lo tuvo en mayor grado. Como dice Popper, sin darse cuenta de que atenta contra la autonomía del contexto de justificación —que él defiende, aunque sin usar esa terminología—, el éxito predictivo garantiza que las premisas explicativas no son *ad hoc*.

³ Leplin, *A Novel Defense of Scientific Realism*, Nueva York, Oxford University Press, 1997.

Suelen negar la autonomía del contexto de justificación los partidarios de concepciones irracionalistas de la ciencia, pero no es eso lo que ocurre en este caso. Más bien es al revés: contra lo que sostienen algunos, el realismo científico parece estar fuertemente asociado a las concepciones racionalistas de la ciencia; no parece posible ser racionalista en la concepción de la ciencia sin ser también realista. A eso se agrega que, frente a la tesis de la equivalencia empírica, la única manera de defender el realismo consiste en asignar más peso justificatorio a las predicciones exitosas que a las explicaciones aparentemente adecuadas. De modo que casi parecería, contra lo que muchos han pensado, que una concepción racionalista de la ciencia está obligada a negar la autonomía de la justificación, y a asignarle un papel a la historia, como pedía Kuhn, aunque no el que él pedía. Leplin no admitiría esto último: él pretende reconstruir el *razonamiento* que condujo a una teoría sin adjudicar ningún papel a factores históricos contingentes. No voy a examinar aquí la cuestión de si esto es posible, aunque tengo la sospecha de que no hay razonamientos que conduzcan a teorías. Cito el caso de Leplin porque constituye un esfuerzo reciente –y, hasta donde yo sé, el más exhaustivo– por explotar el concepto de predicción novedosa en beneficio del realismo científico.

La relación de equivalencia empírica entre una teoría más o menos verdadera y otra totalmente falsa le sirve de base a P. Kyle Stanford para proponer “Una explicación antirrealista del éxito de la ciencia”⁴, según lo anuncia el título de su artículo. Stanford usa una terminología distinta de la que hemos venido empleando; en vez de hablar de equivalencia empírica, habla de “similitud predictiva” (en algún momento aclara que usa la palabra “predicción” en un sentido extremadamente amplio, que permite aplicarla también a las retrodicciones y las explicaciones). Según él, el éxito de una teoría falsa se explica por su similitud predictiva con la teoría verdadera en el dominio de que se trate. La teoría tolemaica mejorada tiene éxito porque es predictivamente similar a la teoría de Copérnico. Y no hay que seguir preguntando: esta explicación en términos de similitud predictiva pone fin a la cadena de pedidos legítimos de explicación. Dice, en efecto, Stanford: “*el éxito de una teoría falsa dada es explicado por el hecho de que sus predicciones son (suficientemente) cercanas a las hechas por la explicación teórica verdadera del dominio pertinente*. Y lo que es quizá más importante, es apropiado terminar ahí nuestra demanda de explicación: como vimos antes en el caso de explicar el éxito del sistema tolemaico, es inapropiado preguntar cuál otra característica *de la teoría* da cuenta de o explica su similitud predictiva con la verdad”⁵. No hay, según él, ninguna respuesta general a la pregunta acerca de cómo es que una teoría falsa

⁴ Stanford, “An Antirealist Explanation of the Success of Science”, *Philosophy of Science*, 67 (Junio de 2000), pp. 266-284.

⁵ *Ibid.*, p. 275; todas las cursivas son de Stanford.

puede generar, en un momento dado, las mismas "predicciones" que una verdadera; tal pregunta sólo se puede responder haciendo referencia a los mecanismos específicos postulados por cada teoría particular.

Esta propuesta de Stanford da lugar a algunas dudas. Por lo pronto, él dice reiteradamente que acepta la exigencia de explicar el éxito de las teorías exitosas por alguna propiedad *intrínseca* de dichas teorías —y cuando lo dice subraya la palabra "intrínseca"—, pero propone un *explanans* notoriamente relacional, como la similitud predictiva. Pregunta ingenua: ¿por qué no propone directamente el éxito predictivo como *explanans*? Respuesta: porque eso es lo que tiene que explicar, el *explanandum*. Lo que él sostiene es entonces lo siguiente: algunas teorías totalmente falsas tienen éxito predictivo porque hacen más o menos las mismas predicciones que las respectivas teorías verdaderas. Explicar el éxito de una teoría de esta manera es análogo a dar la siguiente explicación de por qué determinado cisne es negro: "Este cisne es negro porque es cromáticamente similar a aquel cuervo". Conviene decirlo de manera más sencilla, para no correr el riesgo de que suene interesante: "Este cisne es negro porque es del mismo color que aquel cuervo". Estas no son explicaciones satisfactorias y últimas; reformulan o trasladan el problema, en vez de resolverlo. El problema expresado por la pregunta "¿Por qué algunas teorías falsas tienen éxito?" es, en efecto, prácticamente el mismo que el expresado por la pregunta "¿Cómo puede una teoría falsa ser empíricamente equivalente, o predictivamente similar, a una verdadera?".

Prima facie, hay una respuesta general a esa pregunta. El aspecto lógico del asunto consiste en que, como es sabido, la deducción no conserva la falsedad, y ésta es la explicación genérica de cómo es que teorías falsas pueden tener consecuencias verdaderas. Pero hay distintos casos de no-conservación de la falsedad.

Caso 1. De la premisa falsa (pero parcialmente verdadera) "Todos los presidentes argentinos del siglo XX murieron en 1974" se deduce la conclusión verdadera "Perón murió en 1974". En un caso como éste no hay mucho misterio: la conclusión verdadera está "contenida" en la premisa falsa; la premisa equivale a una conjunción de enunciados singulares, uno de los cuales —la conclusión— es verdadero. Esto vale también, por supuesto, para teorías enteras; algunas teorías falsas, como por ejemplo la de Copérnico, tienen consecuencias verdaderas porque son *parcialmente* verdaderas (*grosso modo*, es cierto que los planetas giran alrededor del Sol y es falso que sus órbitas son circulares). (No hace falta considerar por separado el caso de las teorías *aproximadamente* verdaderas, como lo es la de Newton para ciertas velocidades, porque en sentido estricto esas teorías son totalmente falsas y no tienen consecuencias verdaderas —aunque sí consecuencias que son excelentes aproximaciones a la verdad, a tal punto que su diferencia con las verdades respectivas es tecnológicamente indetectable—.)

Caso 2. Las reglas de inferencia llamadas generalización existencial y adición dan lugar a casos de no-conservación de la falsedad más complicados

que el de la verdad parcial. De la afirmación falsa según la cual Perón murió en 1945 se deduce por generalización existencial que alguien murió en 1945 –lo que seguramente es cierto– y por adición, que Perón murió en 1945 o París es la capital de Francia. ¿O acaso es parcialmente verdadera la creencia de que Perón murió en 1945? A primera vista no lo parece; y no estaría bien decir, como si se tratara de una información interesante, que en los razonamientos deductivos la conclusión está contenida en las premisas, y al mismo tiempo explicar “contenido en” diciendo que significa *deducible de*.

Caso 3. Las premisas de algunos razonamientos, y algunas teorías, contienen falsedades que se compensan mutuamente, de modo de dar lugar a conclusiones o consecuencias verdaderas. Es lo que pasa en el siguiente razonamiento: “Todos los catamarqueños son franceses, todos los franceses son argentinos; por lo tanto, todos los catamarqueños son argentinos”. Otro ejemplo, apenas un poco más interesante: la hipótesis de que la Tierra es redonda más la hipótesis auxiliar de que la luz se propaga en línea recta tienen la consecuencia de que el casco de un barco que se aleja desaparece de la vista antes que el mástil, consecuencia que también se deduce de la hipótesis de que la Tierra es plana más la hipótesis auxiliar de que los rayos de luz siguen una trayectoria curva. Esta última hipótesis es *ad hoc* en el siguiente sentido: permite obtener la consecuencia observacional buscada –que de otro modo la hipótesis principal no podría producir– y no se la puede confirmar de manera independiente, ya que en distancias tan cortas los rayos de luz siguen una trayectoria recta (o de curvatura despreciable).

Los casos que nos interesan son el primero y el último, el de las teorías falsas que tienen consecuencias verdaderas porque son parcialmente verdaderas y el de las teorías falsas que tienen consecuencias verdaderas gracias a alguna maniobra *ad hoc*, ejemplificados por la teoría de Copérnico y la de Tolomeo. La posición de Stanford ignora el hecho de que no *cualquier* teoría *totalmente* falsa puede tener todas las consecuencias verdaderas que hagan falta para “salvar los fenómenos”. Para que una teoría totalmente falsa tenga tales consecuencias verdaderas, debe contener falsedades que se compensen mutuamente. Por ejemplo, la teoría de Tolomeo dice que los planetas giran alrededor de la Tierra. Si dijera solamente esto, no podría explicar todas las posiciones observadas de los planetas. Para poder hacerlo, debe agregar que los planetas se mueven en epiciclos. Ambas afirmaciones –que los planetas se mueven alrededor de la Tierra y que lo hacen en epiciclos– son falsas, y ninguna de las dos tiene por sí sola todas las consecuencias verdaderas que debería tener, pero su conjunción sí las tiene. Como la idea fundamental que guía la construcción del sistema tolemaico es la idea de que la Tierra está en el centro del universo, los epiciclos son el prototipo de estrategia *ad hoc* (es cierto que también la teoría de Copérnico necesita epiciclos debido a que postula órbitas circulares, pero ésta es otra historia).

Stanford sostiene que el éxito predictivo novedoso –tanto si se lo entiende de la manera intuitiva habitual como si se lo entiende a la manera de Leplin– puede ser explicado por la “similitud predictiva” tan plausiblemen-

te como por la verdad parcial. Un punto clave en la argumentación con que defiende esta tesis es la idea de que no hay razones para creer que la probabilidad de que una teoría parcialmente verdadera genere éxito predictivo novedoso es mayor que la probabilidad de que dicho éxito sea generado por una teoría predictivamente similar con respecto a todos los fenómenos conocidos hasta ahora en el dominio de la teoría (esto es, no hay razones para creer que $\text{pr}(\text{PN}/\text{VP}) > \text{pr}(\text{PN}/\text{SP})$). Desde luego, no hay manera de cuantificar estas probabilidades, y eso deja margen para una amplia variedad de maniobras argumentativas. Así, Stanford sostiene *a)* que las predicciones derivadas de una teoría parcialmente verdadera pueden proceder de su parte falsa, y *b)* que una teoría totalmente falsa pero que hasta el momento se las ha arreglado para salvar todos los fenómenos conocidos en su dominio de aplicación, hará predicciones sobre la aparición de fenómenos novedosos en ese mismo dominio, en el que la maquinaria de la teoría ya se ha mostrado capaz de salvar los fenómenos, y eso le dará la oportunidad de que dichas predicciones sean exitosas. Esto último es casi textual; lo primero Stanford lo dice así: "cualquiera de las predicciones novedosas de una teoría, o incluso todas ellas, podrían muy fácilmente resultar uno de los (posiblemente muchos) modos en que la teoría falló o se extravió"⁶.

Esta argumentación es tendenciosa. Una teoría parcialmente verdadera tiene, valga la perogrullada, una parte verdadera, y no tiene por qué ser más probable derivar predicciones de su parte falsa que de su parte verdadera. En esto no tiene nada que ver la cuestión de que por cada verdad hay infinitas falsedades⁷; la parte falsa de la teoría contiene determinadas falsedades particulares, no todas las falsedades posibles. Y, sea como fuere, con respecto a esto es peor la situación de una teoría totalmente falsa, que, para seguir con las perogrulladas, no tiene parte verdadera y sólo mediante hipótesis *ad hoc* logra "salvar" fenómenos *ya conocidos*, siendo sumamente improbable que sea capaz de predecir otros que sean de distinto tipo en algún sentido razonablemente fuerte –por ejemplo, "novedosos" en el sentido de Leplin, o tal vez en algún otro sentido–. En suma, no es improbable que la parte verdadera de una teoría parcialmente verdadera dé lugar a predicciones novedosas; en cambio, que eso lo haga una teoría diseñada *ad hoc* para salvar ciertos fenómenos, sería un milagro⁸.

UNIVERSIDAD NACIONAL DE MAR DEL PLATA
mcomesan@mdp.edu.ar

⁶ Stanford, ob. cit., p. 281.

⁷ Que le hizo decir a Simone de Beauvoir, en el prólogo de su librito *El pensamiento de la derecha*, el siguiente disparate: "La verdad es una, el error múltiple; no es casual que la derecha profese el pluralismo".

⁸ Esto también lo dice Putnam en "What Is Realism?", en Leplin (ed.), *Scientific Realism*, Berkeley, University of California Press, 1984, p. 141.

ABSTRACT

P. Kyle Stanford proposes "An Antirealist Explanation of the Success of Science" according to which the success of a wholly false scientific theory (as for example an enhanced version of the Ptolemaic system) can be explained by its "predictive similitude" (= empirical equivalence) to the true theory in the relevant domain (the Copernican theory), and he holds that this explanation: (a) is antirealist, for it doesn't require that the successful theory be true—not even partially or approximately; (b) it ends the chain of legitimate requests for explanation: there is no general answer to the question how a false theory can generate the same predictions as a true one, which can only be answered by appealing to the specific mechanisms posited by each particular theory; (c) is as good as a realist explanation, even regarding the success of the theories in generating "novel" predictions.

In this work I wish to suggest that: (a) to explain the success of a theory in this way is analogous to explaining the blackness of a swan by saying that it is chromatically similar to a crow; (b) there is a plausible explanation of how a totally false theory can have the mentioned capacity, viz., that it contains *ad-hoc* hypothesis; and (c) that, thus, a false theory having also the same capacity as a (partially) true one to generate predictions that are "novel" in some reasonably strong sense would be a miracle.