

LA TEORÍA DE PRINCIPAL-AGENTE EN LOS ESTUDIOS DE CIENCIA Y TECNOLOGÍA¹

Remo Fernández-Carro
Universidad Carlos III de Madrid

ABSTRACT: *This paper introduces principal-agent theory and its recent application in science and technology studies. First, the text describes the theory itself and its developments in economics and political science. The paper examines the way these disciplines have used it to explain research related issues. Then, it presents specific applications of the theory to science and technology studies. The paper suggests that the Principal-Agent Theory is related to classical theories of the discipline, of which it is sometimes merely a formalization. It proposes, finally, that the theory weaknesses are due to its misuse and its limited development rather than to its real potential, judging by its successful application in other social sciences.*

KEY WORDS: *Principal-agent theory, agency theory, science, technology, research, sociology of science, economics of science, science and technology studies, public policy, science policy, reward system of science.*

1. INTRODUCCIÓN

La teoría de principal-agente (PAT, desde aquí) apareció en la economía a mediados de los años 1970 (Spence y Zeckhauser, 1971). Se trata de un desarrollo formal de las teorías de contratos y se relaciona con las teorías de costes de transacción y de información imperfecta. Aunque muchas de sus ideas sean muy anteriores, como sucede tantas veces. Proceden del campo de la economía actuarial y de los seguros, al igual que su rara terminología². Desde este origen restringido, la teoría de principal-agente se ha extendido al resto de la economía y, en los últimos años, a la mayor parte de las ciencias sociales.

Su objeto de estudio es el *acuerdo entre dos actores* por el que uno solicita al otro que realice, a cambio de un pago, una tarea que el primero no puede o no quiere hacer por sí

PRINCIPAL-AGENT THEORY IN SCIENCE AND TECHNOLOGY STUDIES

RESUMEN: Este trabajo presenta la Teoría de Principal-Agente y la forma en que se ha establecido en los últimos años en los estudios de ciencia y tecnología. Para ello describe, primero, la propia teoría y la forma en que la han desarrollado la economía y la ciencia política. El trabajo revisa el modo en que estas disciplinas la han utilizado para explicar problemas relacionados con ciencia y tecnología o que los tocan lateralmente. Después se ocupa de las aplicaciones concretas en los llamados estudios de ciencia y tecnología. El artículo propone que la teoría se relaciona con las teorías clásicas de la disciplina, de las que a veces es sólo una formalización. Como propone, por fin, que las debilidades de la teoría se deben antes a su mal uso y a su escaso desarrollo que a sus posibilidades reales, a juzgar por su éxito en otras ciencias sociales.

PALABRAS CLAVE: Teoría de principal-agente, teoría de agencia, ciencia, tecnología, sociología de la ciencia, economía de la ciencia, estudios de ciencia y tecnología, políticas públicas, políticas de ciencia, *sistema de recompensas de la ciencia.*

mismo. Es obvio que tal acuerdo no necesita estar mediado por un contrato, que es una institución social algo más compleja. La teoría aplica la metáfora a todos los acuerdos de esta clase, tanto si los protege un contrato como si no. Las dificultades de las partes para llegar a un acuerdo y para garantizar que ambas cumplirán sus obligaciones son lo que le da interés y lo que, al final, orienta los trabajos empíricos. La asimetría de información que aparece entre los actores es, de entre esas dificultades, la que más atención ha merecido. De un lado, quien realiza el trabajo –el agente– obtiene más información sobre su labor que quien lo ha encargado –el principal–; quien, sin embargo, necesita esa información para decidir sobre el propio acuerdo. De otro, sólo el agente tiene información completa sobre su verdadera capacidad para ocuparse de la tarea. La complejidad añade interés a la relación de principal-agente: que se repita o se prolongue en el tiempo, que los actores

sean colectivos en lugar de individuales, que los intereses de cada parte sean muy divergentes o estén enfrentados, que los actores difieran también en su grado de aversión al riesgo, que el resultado de la tarea sea incierto o difícil de juzgar o que el comportamiento del agente sea difícil de evaluar (Eisenhardt, 1985, p. 71).

Estos rasgos hacían casi inevitable que la PAT llegara a los estudios de ciencia y tecnología: es difícil imaginar una relación en que la asimetría de información sea mayor que en la investigación científica. Tampoco es fácil encontrar una actividad cuyo resultado a corto plazo sea más incierto. Por otra parte, la noción de la ciencia pública como un contrato social es antigua y se suele atribuir al muy citado informe de Vannevar Bush de 1945, *Science, the Endless Frontier*³. La metáfora se ha presentado de distintas formas, bien como un contrato entre la sociedad y la comunidad investigadora mediado por el Estado o las empresas; bien como un contrato entre el Estado y los científicos o entre el Estado y las empresas; o bien como una relación entre el Estado y alguna de las agencias que se ocupan de hacer investigación o de administrarla.

En este trabajo me ocuparé de los usos que se ha dado a esta fértil metáfora en los estudios de ciencia y tecnología. Utilizando la literatura que ya la ha desarrollado, trataré de mostrar tanto las ventajas como los límites de este punto de vista. Subrayaré también su flexibilidad y su capacidad para formalizar las ideas más establecidas sobre ciencia, tecnología y sociedad, desde la del *sistema de recompensas de la ciencia* mertoniano o la de *estratificación en la ciencia* de los Cole, hasta las ideas de la Triple Hélice y del Modo 2 de producción científica, pasando por las teorías de los intereses de Bloor o Barnes, o la versión de Latour del Actor-Red.

Para ello, en la **sección primera** presentaré una explicación general del modelo de principal-agente y, a continuación, me ocuparé brevemente del uso de la PAT en economía, ciencia política y sociología, por la influencia que ha tenido en los estudios de ciencia. En la **sección segunda**, resumiré las aplicaciones de la PAT a los estudios de ciencia y tecnología. En primer lugar revisaré sus antecedentes, tanto las teorías que la idea de principal-agente vendría a formalizar –procedentes en su mayoría de la sociología–, como los usos que se le ha dado en ciencia política y sociología y en teoría de las organizaciones. En segundo lugar exploraré la fructífera apli-

cación de la idea en la economía de la ciencia y la tecnología; y pasaré luego a las aplicaciones recientes de la idea en los estudios sociales de la investigación científica y tecnológica. La **sección tercera** reunirá las críticas que recibe la teoría, en parte por relación con su marco teórico, las teorías de elección racional; me detendré en algunos errores comunes en que han caído quienes la utilizan. En su segundo apartado propondré algunos aspectos de la teoría que aún no han sido desarrollados en estos estudios. En la **última sección** aventuraré una valoración de la PAT en su aplicación a los estudios de ciencia, tecnología y sociedad. No será una sorpresa que mi valoración sea positiva, aunque crítica, y que trataré de persuadir al lector de la necesidad de incorporar la teoría a la caja de herramientas de los científicos sociales, se ocupen o no de los científicos y tecnólogos.

2. LA TEORÍA DE PRINCIPAL-AGENTE EN LAS CIENCIAS SOCIALES

La Teoría de Principal-Agente, que he nombrado PAT por sus siglas inglesas, es una más de las teorías de elección racional. Como sucede con éstas, es difícil darle el nombre de "teoría" salvo en la forma en que lo aplicamos a una teoría matemática. Aunque describe una relación social y asume un comportamiento de los actores –la tendencia a maximizar su utilidad– y aunque supone un tipo particular de actor –por el individualismo metodológico que comparte con ese conjunto de teorías– se trata de un modelo *vacío* que se puede aplicar a cualquier par de actores entre los que reconozcamos una relación como la que describen Pratt y Zeckhauser (1985b, 2): "*Allí donde un individuo depende de la acción de otro*". Al igual que en otros ejemplos de la teoría de juegos, en las relaciones de principal-agente los actores obtienen el mejor resultado si son capaces de cooperar. El interés consiste entonces en estudiar aquello que en cada caso se lo dificulta o se lo facilita. La posibilidad de la cooperación a pesar de los contratiempos es uno de los rasgos que hacen atractiva la teoría para estudiar toda esta clase de relaciones de manera formal o, al menos, falsable⁴. Para el científico social es aún más atractiva por la evidencia de que a menudo la solución a los dilemas de la relación es un puñado de instituciones sociales, o que ese par de actores puede consistir en un par de grupos, de organizaciones, de empresas o de clases sociales. De este modo, la teoría promete tender

un puente entre las explicaciones micro y macro, entre la teoría de alcance intermedio y la gran teoría, entre las explicaciones estructurales y las de agencia individual, y entre las metodologías cuantitativas y cualitativas. Promesa que, por lo demás, es la que hace todo ese grupo de teorías (Goldthorpe, 2000).

Pero, dejando de lado las promesas, ¿en qué consisten los modelos de principal-agente? Y, ¿cómo los aplicamos al estudio de la investigación científica? Comenzaré por la primera pregunta y dejaré la segunda para la sección siguiente.

En su expresión más sencilla, como hemos visto, un actor acuerda con otro que haga para él una tarea a cambio de una compensación⁵. Al primer actor lo llamamos principal y al segundo, agente. Asumimos –es fácil– que los actores prefieren llegar a un acuerdo porque es beneficioso para ambos; si no fuera así, no habría contrato. La dificultad comienza con la incertidumbre del resultado y con la diferencia entre lo que saben los actores sobre la relación, en particular lo que sabe el principal sobre el trabajo que realiza el agente. En primer lugar, el resultado de una tarea es siempre incierto: la relación entre una acción y lo que produce es aleatoria, al menos en parte. La habilidad del agente puede reducir esa incertidumbre, pero no eliminarla. Un cirujano no puede anticipar todas las complicaciones de una intervención, por experto que sea. Por eso la *regla de pagos* acordada entre los actores establece cómo se van a repartir entre ellos el riesgo producido por la incertidumbre y cómo esto afecta al pago.

En segundo lugar, el principal tiene menos información que su agente sobre el trabajo que han acordado. Por un lado, no conoce las circunstancias de la tarea concreta, la información local, ni sabe si su agente trabaja de la mejor manera posible. Esto permite a un agente oportunista manipular la regla de pagos y hacer un menor esfuerzo por la misma retribución. A esto se lo conoce como *riesgo moral del principal* o el problema de *acción oculta*. Por otro lado, el principal desconoce las habilidades del agente, su *tipo*. Corre el riesgo, entonces, de escoger un agente inapropiado, que no sabrá desarrollar la tarea o no lo hará de la mejor manera. A este problema se lo conoce como de *selección adversa* o de *información oculta*. La peor posibilidad es que un sistema de incentivos incorrecto o una regla de pagos inadecuada excluyan a los mejores agentes.

Se puede reducir estas inevitables *asimetrías de información* y paliar la incertidumbre, pero siempre a un coste. No todos los *costes de agencia*, empero, son para el principal: si los pagos son *a posteriori* el agente corre el riesgo de que el principal no le pague como acordaron, o no le pague en absoluto. Se trata del *riesgo moral del agente*: una vez hecho el trabajo, el principal tiene incentivos para no cumplir su parte. A este problema general se lo conoce como de *inconsistencia temporal*. Los actores anticipan ese resultado y la relación no se produce. Si los pagos se producen *a priori*, todo el riesgo moral es para el principal y el problema es el mismo: la relación no se produce.

Pero, a pesar de las dificultades, esta clase de relaciones se dan continuamente. Por un lado, los acuerdos no suceden en el vacío y los actores encuentran en la sociedad una serie de herramientas que les permiten superar los problemas derivados de la inconsistencia temporal. Arrow cita a John Stuart Mill para sugerir que en muchos casos "*los términos [de la relación] tienden a ser regulados por la costumbre*" (1991, 48); en otros "*la responsabilidad profesional es claramente reforzada en buena medida por sistemas de valores [ethics], internalizados durante el proceso educativo y reforzados hasta cierto punto por castigos formales y, de forma más amplia, por reputaciones*" (id., 50). Otros autores han insistido en ese papel de la reputación, entre otras instituciones sociales básicas, como forma de limitar el riesgo moral del principal o del agente en algunas situaciones y, en especial, cuando la relación se repite (Wolfson, 1991, 101–126; Milinski, Semmann y Krambeck, 2002; Arrow, 1991, 50). Para la mayoría de las relaciones el contrato, garantizado por el sistema judicial, es la más importante de las instituciones sociales que limitan los costes de agencia (Coleman, 1990, 147–48). Y Kiewiet y McCubbins encuentran la solución a los problemas de la delegación en otros arreglos organizativos (1991, 23) y en los contrapoderes institucionales (ibíd., 33–34).

La PAT ha ido perfeccionando en las ciencias sociales las herramientas que utilizará en los estudios de ciencia y tecnología. En la economía ha tenido un fuerte auge debido a su utilidad para los estudios sobre problemas de información imperfecta y costes de transacción. Tiene un interés evidente para el estudio de las organizaciones, y en este campo y en el de relaciones laborales ha conocido su mayor desarrollo (Eisenhardt, 1989; Pratt y Zeckhauser, 1984a). El interés de la ciencia política por las teorías de

agencia, por otro lado, ha crecido en los últimos años hasta el punto de que algunos de sus primeros trabajos son ya referencias obligadas de la disciplina (Moe, 1984; Miller, 1999; Tirole, 1994; Ferejohn, 1999, 1986; un resumen en Miller, 2005). La metáfora general se adapta bien al estudio de la delegación democrática, en la que el gobierno es un agente del pueblo o del parlamento. Se ha empleado también en el estudio del papel de la separación de poderes o en el de las elecciones como forma de control *ex post*. Se ha mostrado muy adecuada para el estudio de las burocracias, como agentes de la administración, y del papel de sus reglas y procedimientos en el control de los empleados públicos (Miller y Moe, 1983; Bendor, 1988; Bendor, Taylor y Van Gaalen, 1987; Williamson, 1999; Wood, 1988). De la misma forma se ha mostrado útil para el estudio de la delegación de poderes de los gobiernos a agencias independientes, como los Bancos Centrales. O los consejos de investigación, como veremos. La teoría política se ha interesado por la interacción entre múltiples principales –los ministerios de un gobierno, por ejemplo– y por la competición entre muchos agentes. En sociología, la PAT ha mostrado también su afinidad con las teorías weberianas de la burocracia (Kiser, 1999, 1994; Kiser y Kane, 2001; Kiser y Schneider, 1994): Miller (2005) observa la afinidad entre las asimetrías de información y la idea weberiana de saber de servicio o *Dienstwissen*.

Hasta aquí me he ocupado de la relación de principal-agente y de algunos de los usos que ha tenido en las ciencias sociales. Pero, ¿cuáles son los problemas específicos de la relación entre un investigador y su patrón? ¿Cómo utilizan los estudios de ciencia y tecnología los conceptos y herramientas que he descrito?

3. LA TEORÍA DE PRINCIPAL-AGENTE EN EL ESTUDIO DE LA INVESTIGACIÓN CIENTÍFICA Y TÉCNICA

3.1. La forma particular de la PAT en el estudio de la investigación científica y técnica la hace parecida a algunas de las teorías clásicas

Anoté arriba que los problemas de principal-agente en la investigación científica y técnica son más agudos⁶. Por un lado, por las *asimetrías de información*: a los investigadores científicos o técnicos se les demanda conocimiento

original que, por ser original, es difícil de comprender y de valorar por quienes lo costean. Los resultados de una nueva investigación son difíciles de comparar con los de investigaciones anteriores. Tampoco los servicios que presta son comparables con algo ya conocido. No hay una escala de medida para aquello que es *original*. Esto implica mayores problemas de riesgo moral o de selección adversa –que los científicos no hagan *su trabajo* o que no se escoja a los *adecuados*– pero también implica dificultades para decidir el contenido de “su trabajo” o el significado de “adecuado”. Los problemas son más agudos, por otro lado, por la *incertidumbre* de los resultados científicos a corto plazo, incertidumbre que se traslada a la relación: el investigador no puede comprometerse a producir una cantidad o una calidad determinada.

Como explica también la teoría económica de los costes de transacción, la incertidumbre y las asimetrías de información agravan la inconsistencia temporal y amenazan con hacer imposible el acuerdo. ¿Cómo salva el sistema científico estas dificultades?

Van der Meulen explica cómo una de las soluciones es el control profesional de los agentes mediante la evaluación por los colegas o *peer review*: “*En la ciencia los agentes tienen incentivos para vigilarse mutuamente en tanto que son a la vez consumidores y productores de afirmaciones científicas y de conocimiento*” (1998, 400; también Guston, 1996, 236). La evaluación colegiada se encarna, por ejemplo, en los tribunales académicos o en la figura del *referee* de una publicación. Al evaluar la calidad de un artículo de investigación o la carrera de un científico estas instituciones alinean los intereses de la comunidad científica con los de su principal y le facilitan información que le permite monitorizarlos (ibid., 405). De este modo, el *peer review* disminuye los costes de agencia para el principal –le da información para monitorizarlos– y para los agentes –reduce la incertidumbre sobre su capacidad para valorarlos–.

La sociología clásica de la ciencia, que desarrolla el *sistema de recompensas* mertoniano, apunta la segunda solución: una carrera larga con recompensas diferidas y basadas en la reputación [Merton, 1968, 1957; Whitley, 2003, 2000 (1984); Nowakowska, 1975]. Whitley lo resume de esta forma: “*La investigación en la ciencia pública está orientada fundamentalmente por la búsqueda competitiva de*

reputación académica [collegiate] obtenida a través de la contribución –con publicaciones– a fines intelectuales colectivos y los investigadores son recompensados de acuerdo, en lo esencial, con esas reputaciones" (2003, 1016). Es cierto que Merton había insistido en que la recompensa era la propia reputación y que los científicos se guiaban por una serie de normas ideales –el *ethos* que había observado Mill–, como es cierto que esta idea ha sido ya discutida por la literatura posterior (Barnes y Dolby, 1970; Mulkey, 1969, 1991). Más tarde Barnes (1985) recogió sugerencia mertoniana de que la reputación era una suerte de "moneda" de la comunidad científica, el vector de un mecanismo de recompensa diferida⁷. Así, el investigador intercambia reputación por recompensas, principalmente materiales. Latour y Woolgar (1986) la llaman crédito o credibilidad siguiendo a Bourdieu (1975), quien sigue a su vez a Merton (1968). Para Bourdieu, la autoridad científica es capital simbólico y puede ser transformado: "La autoridad científica es así una clase particular de capital que puede ser acumulada, transmitida e, incluso, reconvertida en otras formas de capital bajo ciertas circunstancias" (ibid., 265). En otro lugar añade que esta "clase particular de capital social" da poder sobre los mecanismos constitutivos del campo científico (ibid., 262).

En *The manufacture of knowledge*, Knorr-Cetina (1981, 70-71) subraya esta idea bourdieuana de un mercado gobernado por el crédito y el capital simbólico, "en un universo esencialmente antagonístico" (ibid., 71) en que los colegas son competidores que se resisten a dar ese crédito. Esta autora anota que la idea no habría sido desarrollada en su totalidad: "Tales modelos económicos de agentes científicos no han sido aún empujados a sus límites" (id.).

3.2. La PAT en la economía de la ciencia y la tecnología

A pesar de que la cita que cierra el párrafo anterior no emplea el término "agentes" de la forma que lo hace la PAT, incluye al menos uno de sus rasgos característicos: la elaboración de modelos (Hedström y Swedberg, 1996). Los primeros modelos del acuerdo entre un investigador científico o técnico y su patrón, sea privado o público, aparecen en la economía.

Por un lado, algunos de los trabajos clásicos toman como ejemplo un profesional técnico –un científico o un ingenie-

ro– cuando quieren ilustrar las dificultades que producen las asimetrías de información. Arrow toma el de la relación médico-paciente (1985, 49). Kiewiet y McCubbins detallan el riguroso control que instituciones como la universidad ejercen sobre sus empleados (1991, 34). Eisenhardt pone la innovación técnica entre las causas del riesgo y la incertidumbre de una relación de principal-agente (1989, 61, 65) y subraya los problemas que encuentran en esta relación las empresas de base tecnológica, en las que encuentra todas las dificultades que puede presentar la relación: un fuerte conflicto de metas y el posible oportunismo del agente, una incertidumbre sobre los resultados suficiente para atizar los problemas de riesgo que implica la teoría, y unos puestos de trabajo en que, porque la labor no se puede programar o porque están orientados al trabajo en equipo, la evaluación individual de los comportamientos se hace difícil (ibid., 71). Todos estos autores subrayan el papel de la reputación, clave del mundo científico, en casos severos de asimetría de la información sobre los agentes (también en Pratt y Zeckhauser, 1985a; Moe, 1984; Miller, 1992; y Tirole, 1994). Por último, la discusión que hace Tirole sobre la recogida de información para la toma de decisiones resulta, implícitamente, un modelo a escala del mundo de la investigación (1994, 25).

Por otro lado, los economistas han hecho ya un buen número de aportaciones explícitas. Sus trabajos se centran en las dificultades de la gestión de la investigación –con dinero privado o público– en la empresa (por ejemplo, Agihon y Tirole, 1994). Así, Dasgupta y David (1987), utilizan la PAT para distinguir ciencia y tecnología de acuerdo a la facilidad para apropiarse sus resultados. Diferente *apropiabilidad* induce diferencias en las comunidades científica y tecnológica, en sus instituciones, sus *ethos*, actitudes y acciones. En la ciencia básica se anima la divulgación de la información mediante la publicación de los resultados. En la ciencia aplicada y el desarrollo tecnológico se anima la búsqueda de lucro mediante la explotación del secreto industrial o de las patentes. La diferente naturaleza de estos conocimientos explica la necesidad mutua y la necesidad de la protección pública para la ciencia, como *bien público* que es. Los científicos, en tanto que científicos, se ven empujados a publicar todo el conocimiento nuevo de acuerdo con las reglas de su comunidad. Reglas que incluyen –y aquí Dasgupta y David siguen de cerca a Merton– procedimientos para presentar y verificar las teorías y la obligación de someter el conocimiento nuevo a la crítica de los colegas (también Stephan, 1996, 1201-1203).

Dasgupta y Maskin (1987), encuentran los problemas de *riesgo moral* y *selección adversa* al valorar la utilidad de los "research portfolios" y recurren también al *sistema de recompensas* mertoniano para argumentarlo. Su modelo se basa en el supuesto, quizá demasiado restrictivo, de que en un contexto competitivo sólo el primer descubridor/inventor obtiene todo el beneficio del trabajo científico. Siguiendo esta idea, Olsen (1993) imputa a la asimetría de información el exceso de gasto en los programas de I+D.

David (1988) continúa su trabajo en historia económica con una explicación del surgimiento de las instituciones de la "ciencia abierta" durante el Renacimiento. El autor encuentra sus orígenes en el sistema de patronazgo aristocrático de las primeras cortes de la era moderna y en la creciente dificultad para comprender la nueva "filosofía mecánica" y valorar sus resultados. Ante este inesperado problema, que no se daba en otras artes liberales, los aristócratas recurrieron a la reputación que los agentes obtenían de sus colegas (ibíd., 18). El resultado fue la emergencia de procedimientos de construcción de reputaciones por referencia a los "colegas expertos" (ibíd., 17). David incluye entre los remedios la creación de las redes de correspondencia que terminarían por producir las primeras revistas científicas o los premios a la solución de rompecabezas particulares, "(...) *una respuesta funcional a los mayores problemas de simetría de información* (...)" (ibíd., 19; también David, 2007).

Otro ejemplo interesante es el artículo de Abraham y Prosch (2000) sobre la gestión de Carl-Zeiss, una empresa de base tecnológica que a finales del siglo XIX desarrolló una política de pensiones que equivalía a un incipiente estado del bienestar. Este sistema de incentivos diferidos estaba destinado a limitar el constante abandono de los mejores técnicos –formados en la compañía– por otras empresas o por la práctica profesional particular, alguno de los problemas sugeridos por Eisenhardt (1989). El sistema de incentivos, no obstante, no podría sostenerse debido un problema de inconsistencia temporal: ¿qué garantizaba a esos técnicos que el patrón pagaría las pensiones al cabo de los años? Abraham y Prosch explican cómo Carl-Zeiss estableció su credibilidad y evitó los problemas derivados del riesgo moral del principal mediante su transformación en una fundación privada, propietaria a su vez de la empresa, lo que *ataba* sus manos.

3.3. La llegada de la PAT a los estudios de ciencia y tecnología

En los apartados anteriores he resumido los problemas específicos de la relación de agencia en el mundo de la investigación y he anotado algunos de los usos que los economistas han hecho de la PAT. Sus aportaciones tienden a ignorar las hechas desde los "estudios de la ciencia", en particular desde la ciencia política o la sociología, salvo quizás a Merton. Pero este desconocimiento es recíproco: es difícil encontrar, en los artículos de un área, referencias a trabajos de otra, con el consiguiente empobrecimiento de un campo aún poco extenso. En lo que resta de la sección presentaré el núcleo de estos estudios, más interesados por la versión sociológica de la teoría tal como la desarrolla Coleman (1990) y más preocupados también por la ciencia pública –o pagada por el Estado– que por la investigación privada.

Si dibujáramos la red de relaciones definida por las citas a los artículos, el centro del grafo estaría ocupado por tres trabajos, los de Braun (1993), Guston (1996) y van der Meulen (1998). La PAT llegó a los estudios de la ciencia, así, con algo de retraso sobre el resto de ciencias sociales. Los tres artículos se ocupan de un actor institucional central en la ciencia *moderna*, el consejo de investigación⁸. Si el primero describe la interacción entre el Estado como principal y el consejo de investigación como su agente –con los científicos como "terceras partes" de la relación–, los trabajos de Guston y van der Meulen estudian a los investigadores como agentes del consejo, aunque sin dejar de notar que éste es a su vez agente del Estado. Esta variedad en el enfoque ilustra bien la flexibilidad de la PAT. Braun toma los consejos de investigación como caso particular de las agencias autónomas o *mission agencies*.

La asimetría de información debida a la especificidad de sus conocimientos permite a los científicos tratar a su patrón de igual a igual, como enseñó Weber. Tanto Braun como van der Meulen explican que la asimetría de información les da poder para "capturar" el consejo de investigación y ponerlo a su servicio. La necesidad de la revisión colegiada –e incluso de la selección colegiada– se explica por la fuerte asimetría de información, como vimos antes. Van der Meulen (1998, 400) enfatiza la necesidad de una confianza recíproca entre el agente y el principal en las relaciones a largo plazo. Tal confianza, junto con las

instituciones sociales del mundo de la ciencia y el hecho de que el contrato se extienda en el tiempo, contribuiría a afianzar una relación "inherentemente inestable". Ya en este primer trabajo, Braun se detiene en otro rasgo específico, la contradicción entre la independencia relativa que los autores clásicos describían en la república de la ciencia y la "racionalidad administrativa" de la gestión pública corriente (1993, 155): el Estado jerárquico no funciona bien en un mundo autogobernado en que la reputación resuelve el problema de "*falta de criterios de evaluación cuantificables*" (ibid., 155).

Cerca del centro de ese imaginario mapa bibliográfico encontraríamos otro trabajo de Guston (2000), *Between politics and science*, que poco a poco se ha convertido en la primera referencia del campo. Este trabajo lo relaciona con la teoría también emergente de las *instituciones de frontera*. No muy lejos, veríamos un trabajo de Caswill (1998) y uno más de Braun (1998). Tanto estos trabajos como la mayor parte de los posteriores citan las *Foundations of social theory* de Coleman (1990) tanto como la *Laboratory life* de Latour y Woolgar (1986): la PAT se muestra otra vez versátil, bien relacionada con la sociología matemática tanto como con el constructivismo de los *Social Studies of Science* y, a través de sus autores, con la teoría del Actor-Red⁹. Otros trabajos citados a menudo incluyen los ya nombrados de Bourdieu (1975) y Polanyi (1962), el libro de Ben-David *The scientist's role in society* (1971) y el de Gibbons *et alii* (1994), y dos clásicos de la PAT, *The new economics of organization* de Moe (1984) y el *Congressional oversight overlooked* de McCubbins y Schwartz (1984). Un artículo anterior de Turner (1990) que desarrolla un modelo de principal-agente sin nombrarlo y otro posterior de Morris (2000) cierran este núcleo central. En lo que resta de este apartado detallaré alguno de estos trabajos.

Guston apenas amplía en su libro (2000) la teoría de su anterior artículo, *Stabilizing the boundary between US politics and science* (1999), y se repite en su siguiente trabajo, *Boundary organizations in environmental policy and science* (2001). Los tres son estudios de caso sobre organizaciones mediadoras de la ciencia estadounidense, que se extienden también en la idea de organizaciones de frontera de Gieryn (1995). La mayor extensión permite al autor, no obstante, demorarse en la descripción de los rasgos de la relación de principal-agente en la ciencia pública

de las democracias y sus problemas. Es por eso una buena introducción al tema: sencilla, además, porque no hace un modelo formal y su trabajo empírico es cualitativo. El primero de los rasgos es la asimetría de información entre ejecutores y patrones, aunque admite que "*no es exclusiva de la política científica*" (2000, 17). Para Guston, lo que es singular de esta política es, primero, su posición en las discusiones sobre el estatus epistemológico particular de la ciencia; en segundo lugar, la impredecibilidad de su desarrollo; en el tercero, las dificultades para anticipar sus consecuencias; y, por último, que la distribución de sus productos se dé entre el público cualificado o en un mercado externo en lugar de entregarse al principal que lo financia (2000, 17). El autor se interesa por cómo éste controla a los científicos mientras tiene que delegar poder organizativo en una agencia intermedia, cómo controla su integridad y lealtad además de su productividad. Los científicos por su parte también se preocupan por mostrar esa integridad y esa productividad por su propio interés y a pesar de la oscuridad inherente a la información científica. Los consejos de investigación devienen herramientas para mantener la separación –y la comunicación– entre los ámbitos político y científico: hacen *boundary work*.

Cuando escribió su artículo, Caswill (1998) era director de investigación del Consejo Británico de Investigación Económica y Social (ESRC). Su ejercicio, por ello, no es tanto el de un académico como el de un funcionario con experiencia. La política de investigación, para él, no ha sido suficientemente investigada por las ciencias sociales. Apenas se han ocupado del efecto de la intervención del gobierno, de la interacción con los usuarios, la interdisciplinariedad y la cooperación internacional, que son los principales motivos de interés de los administradores. La teoría con mayor capacidad para aclarar estos temas sería para este autor la de principal-agente (ibid., 293). El modelo de Caswill sigue al de Braun, con un consejo de investigación como agente del gobierno (aunque principal de los científicos). El consejo resultaría la mejor forma de organización para un gobierno "distante" y que carece de la suficiente información de encarar el problema de la selección adversa. Como el de Braun y el de Guston, su modelo es formal; su material empírico, bastante débil, es cualitativo.

El siguiente trabajo de Braun (1998), también teórico, desarrolla con el entramado de la PAT varios de los elementos

que apuntaron Bourdieu (1975) y Latour y Woolgar (1986), en particular el desarrollo e intercambio de distintas formas de capital. Elabora así, desde otro punto de vista, algunos de los temas que apunta Guston acerca de los consejos de investigación¹⁰. Braun observa, en primer lugar, que la literatura ha enfatizado los premios sociales de los científicos por encima de los económicos. Con Bourdieu, sugiere que el "capital social" expresado por el reconocimiento de los colegas es sólo uno de los capitales del científico y diferente de su "capital económico" (salarios, recursos organizativos y de infraestructura) y su "capital cultural". La estructura de la ciencia transforma los intereses "externos" en desarrollo científico. Braun utiliza los conceptos de Merton pero cambia su punto de vista. En segundo lugar, subraya que las "restricciones cognitivas" de los actores políticos –vale decir su asimetría de información– explican la necesidad que tienen los consejos de cooptar científicos para al menos dos de sus tareas, la de selección y la de control (ibíd., 812). Distingue así tres arenas y termina el artículo con un detallado análisis de las normas e intereses que operan en ellas: la arena de selección, la de control y la de políticas o "*policy arena*" (ibíd., 814 y figura 2). Una tercera observación interesante es que la reputación de los directores de los consejos –y no sólo la de los científicos– estabiliza el sistema: formaría parte de la división de poderes que implica la distinción entre "arenas".

El artículo de Turner (1990) es citado a menudo a pesar de que no utiliza la terminología de la PAT ni sus formalismos y de que está escrito en un lenguaje deliberadamente arcaico y oscuro. Se ocupa de la persistencia del Antiguo Régimen en las prácticas de patronazgo científico, lo que lo relaciona con los trabajos de Davis. Su contenido es, quizás, demasiado teórico para ignorar como ignora la teoría anterior –en particular sociológica– y para presentar tan descuidadamente los ejemplos que utiliza¹¹. Una razón para su notoriedad es que ha sido muy citado por Guston (2000, 1999 y 1996).

Turner arguye que las relaciones de patronazgo entre el gobierno y los científicos se basan en la construcción de la confianza sobre las promesas que hacen estos. La confianza y las promesas resuelven a su vez los problemas de no-verificabilidad –problemas de asimetría de información– y los problemas que producen la diferente distribución del conocimiento –asimetría de nuevo– y el diferente

reparto de poder discrecional. Esa "evaluación prospectiva de las promesas" es un rasgo común en los modelos de los autores comentados hasta aquí: los consejos de investigación distribuyen subvenciones o becas de acuerdo con la confianza en un proyecto de investigación. Van de Meulen había enfatizado también el papel de la confianza. Entre las instituciones que, señala Turner, afianzan la relación produciendo confianza, están los "testimonios" (*attestations*) de otros científicos como las cartas de recomendación o la evaluación de manuscritos enviados a una revista (ibíd., 191). Otras instituciones, que la PAT llamaría "de señalamiento" o "*signaling*", son las que él llama *metonimias*: "(...) partes que señalan algo de la totalidad [y que] en la ciencia son las señales visibles de competencia, validez y fiabilidad que hacen creíble una promesa" (idem). La confianza que cuenta para Turner sigue siendo la que tiene el patrón en el empleado al que va a delegar parte de sus recursos. Si hay crisis de confianza, termina el autor, el patrón la resuelve por la demanda de "productos tangibles", de productos terminados, en lugar de sólo la promesa de lo que se va a hacer.

Otro mecanismo que Turner presenta como generador de confianza es la división de poderes –de "autoridad", en sus palabras–. Es semejante a la que en cualquier negocio separa la parte financiera de la teneduría contable¹². El autor distingue dos dimensiones: delegación y división. La primera consiste en la transferencia de poder de decisión en los niveles más bajos de la cadena de mando. La segunda, la división, consiste en la distribución de diferentes aspectos de la toma de decisiones entre diferentes instancias o personas. En esto se asemeja a las "arenas" de Braun o al papel que describe en los directores de consejo de investigación, que Turner llamaría "puntos nodales"¹³.

El último de los artículos citados hasta aquí, el de Morris (2000), es uno de los primeros trabajos empíricos extensos que hacen uso de la PAT. Se trata de un buen ejemplo de uno de los principales defectos de esta literatura, que presenta un entramado teórico que después no desarrolla y al que, por eso, apenas saca partido. En su investigación, cualitativa en este caso, Morris distingue y clasifica los *principales* que reconocen los propios investigadores de departamentos universitarios británicos en ciencias biomédicas. Su interés es estudiar el impacto del cambio en las políticas de investigación, la demanda por el gobierno de mayor responsabilidad, la introducción de gestión priva-

da o casi privada en los organismos públicos y la búsqueda general de productividad.

No mucho después, en octubre de 2003, la revista *Science and Public Policy* dedicó a la PAT un número monográfico. El volumen contiene seis artículos, de Braun (2003), van der Meulen (2003), Caswill (2003), Guston (2003), Morris (2003) y Shove (2003), y una introducción de Dietmar Braun y David H. Guston (2003), los editores. Lo que debía haber sido la puesta de largo de la PAT en los estudios de la ciencia resulta, sin embargo, un poco decepcionante. Los artículos no sacan partido a las posibilidades de formalización de la teoría –pese al intento de Braun– ni remedian la separación que denunciaba Goldthorpe entre teoría e investigación empírica. La literatura acumula así descripciones de consejos de investigación que no producen entusiasmo: desde 2003 parece estancada tanto en el número de sus seguidores como en el de sus publicaciones anuales o en el de citas a sus artículos.

El volumen podría parecer falta de ambición, quizá por la complacencia de sus autores con la aparente consistencia de la idea central. Sin embargo, muestra propuestas nuevas y algunas críticas. Así, la de Shove quien, a pesar de cierta debilidad en su argumento, presenta un primer reparo a la teoría a la vez que sugiere el estudio de los planes de investigación –*funding programs* en sus términos– como objeto de estudio añadido. Van der Meulen sugiere estudiar los públicos y usuarios de la investigación como “cuartas partes” de la relación de agencia. Guston, por su parte, estudia la relación de los gobiernos con los asesores científicos o técnicos. Caswill y Morris apuntan otras instituciones o “rasgos contextuales” que contribuyen a estabilizar la relación, y varios de los autores señalan la riqueza de actores diferentes que toman parte en ella. Caswill además hace notar el peso que tiene para el agente la evaluación *a posteriori*, o *post hoc*, de haber ganado un contrato con un consejo de investigación (idea que había adelantado Rip (1994): no es el mismo el dinero que procede de uno u otro fondo). Braun hace un intento de formalización para describir distintos períodos en la ciencia pública del siglo XX (como había hecho van der Meulen, 1998), intento meritorio pese a la lasitud de su argumentación y a lo opinable de su descripción de cada período.

La *introducción* al volumen, de Braun y Guston (2003), es una presentación alternativa de la literatura. Los au-

tores subrayan varios temas, además de los comentados: el equilibrio entre los dos actores, su interdependencia y mutua vigilancia –una “calle de doble sentido”– y la observación de que el principal también puede defraudar, observación que repite Guston en su artículo pero que no llega a desarrollar¹⁴.

4. CRÍTICAS, CARENCIAS Y POSIBILIDADES

4.1. Críticas a las teorías de elección racional

La teoría de principal-agente apenas ha sido criticada en los estudios de ciencia y tecnología, salvo por la crítica de Shove (2003) que mencioné antes. Para esta autora la teoría olvida la posibilidad de colusión entre agentes o su capacidad para eludir el control cambiando de patrón de acuerdo a su conveniencia (cambiando de programa de investigación, por ejemplo, o solicitando fondos a instancias distintas). La crítica es débil porque, por un lado, la teoría ha estudiado estos extremos al tratar el problema de múltiples principales, aunque Shove lo desconozca; por otro, la relación que presenta no es de principal-agente porque los programas y los consejos de investigación no pagan a los investigadores, quienes tienen su contrato con otra organización. El primer punto ilustra uno de los problemas de la teoría, el uso descuidado de muchos de sus autores; sobre el segundo volveré en un poco más adelante.

Las críticas a la PAT se subsumen en las críticas a las teorías de la elección racional (TER). La discusión es ya vieja y ocupa muchos volúmenes de la literatura. De hecho, las críticas más profundas han sido incorporadas ya al acervo de la teoría. Como tal discusión, sin embargo, quizá tenga menos interés en el campo de los estudios de la ciencia y la tecnología. Como propuse al principio de este trabajo, más que una teoría sobre la investigación científica y técnica la de principal-agente es una metodología general aplicable a cualquiera de sus aspectos, y el mismo argumento se puede aplicar a la teoría de la elección racional. Por eso, en lo que sigue sólo voy a esbozar las críticas generales a la corriente sólo en aquellos aspectos que puedan interesarnos. Pasaré luego a proponer alguna crítica concreta sobre el modo en que se ha utilizado la PAT y los errores que, en mi opinión, ha producido en la comprensión de la actividad investigadora.

Las críticas habituales comienzan por los reproches, desde otras ciencias sociales, al imperialismo de la economía o al dogmatismo de algunos de sus practicantes, a menudo cierto. Lo primero aparece incluso en libros de texto (Baert, 2001, cap. 7)¹⁵. Pero la primera crítica seria a las TER es a su *reduccionismo*. Estas teorías hacen modelos simplificados de la realidad basados a menudo en supuestos más o menos verosímiles y apoyados en una concepción muy restringida del concepto de racionalidad. La respuesta más corriente es que todas las teorías hacen modelos limitados que dividen el mundo discrecionalmente en porciones de acuerdo a convenciones no siempre explícitas. Los modelos de elección racional, en cambio, tratan de hacer explícitos sus supuestos, correctos o no, y la discusión se organiza en torno a la pertinencia o al provecho de incorporarlos al modelo. Los supuestos de un buen modelo, por lo demás, proceden de la teoría previa y de las observaciones de cada objeto de estudio. Como en cualquier buena teoría, las explicaciones *ad hoc* implican un mal uso, no un defecto de la propia matriz teórica. Frente a una de estas explicaciones defectuosas el investigador puede proponer un modelo distinto que someter a una siguiente falsación.

La discusión sobre el propio concepto de racionalidad, en cambio, no está cerrada aunque esto afecta poco al uso que le damos. En su versión general, la teoría describe a un actor que entre varias opciones elige la que cree que, dadas sus preferencias, le conviene más; decimos racional a este comportamiento. Esta sencilla operacionalización del concepto, parsimoniosa y práctica, es aceptada por la mayor parte de los autores. La dificultad comienza por el origen de las preferencias individuales y de su ordenación. Ese origen podría estar en la sociedad o la cultura, que limitarían así el juego de elecciones de cada persona haciendo del albedrío individual un ejercicio de cómputo. Diríamos que las preferencias personales serían *exógenas* a la teoría. Pero incluso los autores menos "creyentes" conceden que es posible –y útil– estudiar ese "cómputo" y los agregados de sus consecuencias. Goldthorpe (2000, 125) avanza un paso más para presentar la TER como una teoría especial, en lugar de general "(...) *alrededor de la que se puede organizar mejor el esfuerzo de lograr una mayor generalidad*". Esto recuerda el consejo mertoniano de buscar teorías de alcance intermedio y las TER han tenido hasta ahora un relativo éxito en este propósito de "organizar el esfuerzo".

Dejamos para otro lugar el problema de que conceder incluso ese libre albedrío –racional– en quien toma una elección implica una posición antropológica fuerte, no sólo metodológica.

Dos problemas que vienen de la mano de los anteriores han orientado la discusión teórica y han sido, por fin, incorporados a la teoría¹⁶. El primero tiene que ver con la calidad de la información con que cuenta el "elector racional". Además de atender a sus preferencias, el actor toma sus decisiones de acuerdo con sus *creencias* sobre el estado del mundo. Las creencias –sobre las metas y los mejores medios para alcanzarlas– pueden ser falsas, sin embargo. El actor puede no tener suficiente información para decidir (o puede no saber si la tiene). Elster se preguntó, por ello, en qué momento es racional dejar de buscar información para decidir (1986). El segundo problema se refiere a la capacidad de los actores para hacer ese cálculo racional. En ocasiones se pueden equivocar al anticipar las consecuencias de su elección, por ejemplo, o quizá no alcancen a "computarlas"; se pueden comportar de forma irracional en algunos momentos o para algunos temas y ello puede influir en su propio cálculo racional en las otras ocasiones. Desde las críticas clásicas de Simon (1957) y los trabajos de Kahneman y Tversky, entre otros, la llamada racionalidad acotada o *bounded rationality* se ha incorporado a la caja de herramientas de la teoría.

El mal uso de la teoría –o de sus herramientas– es una situación distinta. Un error corriente es abandonar el individualismo metodológico en que estas teorías se concibieron. Es aún más común en los estudios de ciencia y tecnología: comencé este trabajo refiriéndome a aquellos que explotan la metáfora del contrato social. Así, por ejemplo, la teoría de principal-agente se suele aplicar a los sistemas nacionales de investigación para decir que la comunidad científica es agente de la sociedad (o del Estado). Se trata de una aplicación metafórica, es obvio, porque no hablamos de individuos que tomen decisiones teniendo en cuenta las de otros actores. No se trata de que no podamos trabajar con agregados, sino de que no perdamos de vista a ese individuo que en algún momento tomará su decisión, solo¹⁷.

La metáfora del contrato social puede llevar a un error diferente: confundir el principal. Muchos de los trabajos citados toman a los consejos de investigación como

"principales" de los investigadores. La metáfora sólo sirve en un sentido muy general porque los *consejos* no pagan a los investigadores. El dinero que distribuyen –en forma de subvenciones o ayudas– no forma parte de la *regla de pagos* y sólo de una forma muy indirecta forma parte del *sistema de recompensas*. Son las "organizaciones anfitrionas" de Caswill (2003) las que pagan a los agentes y la *regla de pagos* son esos complejos sistemas de incentivos de carrera basados en el mérito¹⁸. Los investigadores tienen *contratos* menos metafóricos con las organizaciones anfitrionas, y los *consejos* de investigación tienen *contratos* con los gobiernos y no con los científicos. De manera que no pueden ser agentes de éstos, como sugiere Rip (1994, 12-13). Cuando los hay, los *contratos* explícitos son una buena guía para reconocer las relaciones de principal-agente y evitan la confusión en su análisis formal, aunque es evidente que para estudiar la regla de pagos tengamos que considerar el contexto de instituciones sociales de la ciencia.

4.2. Posibles caminos

Es fácil suponer que la PAT se desarrollará en los estudios de ciencia y tecnología de dos formas: por imitación de los caminos que ha tomado en las disciplinas que los alimentan –la economía, la sociología y la ciencia política, al menos–; y por incorporación de aspectos descuidados de la teoría, descuido debido en parte a esa imitación. Si la teoría de principal-agente sigue desarrollándose en las ciencias sociales, con sus pequeños tropiezos, a un ritmo más o menos rápido, terminará por calar en ellos.

¿En qué direcciones se desarrollará la teoría? Dentro de la obvia incertidumbre, es fácil suponer que sus trabajos se diversificarán, alejándose de los estudios de caso. En primer lugar, se incorporarán otros actores tecnocientíficos y relaciones distintas o más complejas: cada contrato en el mundo de la investigación es susceptible de un modelo de principal-agente. Es posible que se estudie mejor el papel de las organizaciones anfitrionas o *host institutions* de Caswill (2003) como los institutos, laboratorios o universidades. O el papel de los clientes, como sugirió van der Meulen (2003); o el de los asesores, como Guston (2003). Es posible que se estudie mejor a los actores básicos: no se ha mencionado la influencia de la ideología de los gobiernos en la forma de la relación, a pesar de la vieja discusión

entre los partidarios con Polanyi de una ciencia autoorganizada y los de Bernal de una investigación planificada por los Estados-. Ni se ha tratado apenas del impacto del régimen político sobre las instituciones de la ciencia o su productividad, a pesar de los trabajos clásicos de Merton (véase Fernández-Carro, 2002, y Fernández-Carro y Lapuente-Giné, 2006). Pocos autores se han interesado por gobiernos que no sean democracias occidentales y por científicos de sociedades pobres.

En segundo lugar, la PAT en los estudios de ciencia y tecnología se irá haciendo más formal. Es el mejor modo de sacarle partido. Los resultados se presentarán en términos más fáciles de falsar lo que, de un lado, animará la discusión teórica y las propuestas de investigaciones empíricas y, de otro, limitará los errores groseros que ha producido su uso "metafórico". El formalismo, incluso matemático, no implica la necesidad de una corroboración cuantitativa, pero desde luego la facilita. El empleo de datos estadísticos permitirá *calibrar* algunos modelos y contribuirá así, por otra vía, al debate teórico. La investigación científica y técnica genera una buena cantidad de datos sociométricos o econométricos que apenas han sido explotados aún desde este punto de vista. Ni siquiera por disciplinas que, como la economía, tienen más hábito de hacerlo.

La teoría de principal-agente es un caso extremo de asimetrías de información, porque la tarea que se demanda a los agentes es precisamente producir esa información que resulta asimétrica. Pero se ha explorado poco las diferencias en la aceptación de riesgo por parte de los actores y la forma en que las instituciones sociales del mundo de la ciencia contribuyen a paliarlas –o no-. Los problemas de la multiplicidad de agentes o de principales han sido discutidos a menudo –por ejemplo, para sugerir que a mayor número de patrones más difícil el control de los agentes (Shove, 2003; Guston, 2003)– pero los pobres resultados reclaman mayor atención para este aspecto. Apenas se ha analizado la regla de pagos de la relación, que en otras disciplinas ocupa buena parte del debate: la mayor parte de los modelos citados suponen que el pago se da *ex ante* y pocos autores sugieren un modelo de pagos *ex post*¹⁹. De la misma forma, apenas se ha desarrollado la idea que presenta Miller (2005) de *riesgo moral del principal*: si el patrón de la relación puede defraudar al agente y no pagarle lo convenido, por ejemplo, cambian los extremos de

la relación (véase Fernández-Carro 2007; Guston, 2003; y especialmente, Bendor 1988). El sistema de instituciones que estabiliza la relación debe ser diferente, como debe serlo el sistema de incentivos que se crea con la regla de pagos. El problema se agudiza cuando los pagos son *ex post*, por la tarea realizada. Sobre todo si el pago es muy diferido como en el mundo científico.

Los autores de la PAT tienden a olvidar una parte de la discusión teórica anterior, como anoté arriba. La teoría se ha ocupado de muchos de sus problemas y presenta en ocasiones modelos informales de una relación de principal-agente. Una tercera dirección en la aplicación de la PAT sería la falsación de las teorías, clásicas o contemporáneas. El descuido general de la teoría o de su verificación empírica revela acaso un problema en la institucionalización de la disciplina: la falta de un consenso sobre el paradigma y, más sencillo, la falta de un *manual* conocido por la mayoría. Las teorías formales, en mi opinión, contribuyen a centrar el debate académico y permiten en ocasiones construir esos consensos paradigmáticos. Ésa sería una de las promesas de la PAR en los estudios de la ciencia: si no la de producir un paradigma, sí al menos la de aclarar cuál es la teoría útil y cuál la que está en discusión.

5. PUNTO Y SEGUIDO: UNA VALORACIÓN DE LA PAT

Teoría o metodología, es posible que la de principal-agente se extienda por las ciencias sociales con toda la familia de la elección racional y el individualismo metodológico. Aunque para ello no basta con su relativo éxito en las ciencias económicas, en mi opinión: tiene que persuadir a los investigadores. Para el ancho campo de los estudios de la

ciencia y la tecnología puede ser una oportunidad de construir un raro puente que permita a los practicantes de distintas disciplinas hablar un idioma semejante y comparar y compartir resultados. Incluso para responder preguntas diferentes y desde diferentes orientaciones teóricas. Ésa es la promesa, al menos, como he escrito. El campo comparte objeto de estudio, como mínimo pero no sólo, cuando se ocupa de la relación del Estado con sus investigadores –públicos o privados– a través de una agencia más o menos independiente que llamamos consejo de investigación y que distribuye fondos competitivos²⁰. Las peculiaridades de este objeto de estudio, y también su interés, tienen relación con aquellos dos problemas centrales: las asimetrías de información y la incertidumbre de los resultados o de su apropiación. Estos dos rasgos, reconocidos bajo diferentes nombres por todas las disciplinas, son centrales también en la teoría de principal-agente. He mostrado con varios ejemplos que los investigadores que se ocupan de ellos tienden a cruzar las barreras disciplinares en busca de respuestas, con mayor o menor acierto. La PAT podría ser un vehículo más seguro para pasar esas fronteras. La unidad de lenguaje –como de objeto de estudio– no contradice la flexibilidad, dado el pluralismo teórico que encontramos bajo el paraguas de la elección racional y la posibilidad de estudiar diferentes actores desde diferentes puntos de vista: he anotado las relaciones con el constructivismo y con el neoinstitucionalismo, con los estudios de historia institucional y de políticas de investigación, con la sociología clásica mertoniana y weberiana y la nueva economía de la ciencia; Guston ha ilustrado las relaciones con la nueva teoría de los *objetos de frontera* (2001, 2000). A lo largo de esta introducción he citado a autores de muy diferentes disciplinas e intereses teóricos cuya posible coherencia se revelaba gracias a este marco analítico. La PAT no ofrece "la" respuesta: ofrece un juego de herramientas para buscar la de cada pregunta.

NOTAS

Recibido: 24 de abril de 2008

Aceptado: 1 de junio de 2008

1 Agradezco a Manuel Hidalgo sus atinados comentarios a la primera versión de este trabajo y a Manuel Fernández Esquinas sus cuidadosas notas.

2 Véase Crosby (1905), quien la toma de los sistemas legales anglosajones (Coleman, 1990, 146-157).

3 Bush no utiliza el término, aunque se le cite por él (Guston, 2000, 39-48; Smith, 1994).

- 4 Sobre las ventajas de la formalización en las ciencias sociales véase Goldthorpe (2000) o Hedström y Swedberg (1996). También Morrow (1994, 6-8 y 302-305) y Mari-Klose (2000, 35-53).
- 5 Véase Arrow (1991), y Mas-Colell, Whinston y Green (1995); en sociología, Coleman (1990).
- 6 Van der Meulen hace una excelente enumeración de las características del contrato (1998, 400): las diferencias de intereses de las partes, las asimetrías de información, el coste de monitorización en que incurre el principal y la necesidad de confianza mutua entre los actores.
- 7 La idea de la "moneda" es de Merton (1968, 56) aunque él considere las recompensas materiales como secundarias en el sistema de recompensas o instrumentales; sigue en esto a Hagstrom (1965).
- 8 Rip (1994) y Fernández-Carro (2002).
- 9 Una discusión más completa en Guston (2000, 26 y 66); quien, no obstante, olvida la contribución de la sociología clásica (Merton, 1968, 2 y nota 16) y de Bourdieu a la discusión sobre las transformaciones de los distintos tipos de capitales.
- 10 No olvida a Merton ni a Bourdieu, como hace Guston, pero ignora a Barnes y los autores de la Teoría de los Intereses.
- 11 Turner se apoya en un artículo de Shils sobre la delegación en el mundo de la ciencia que adelanta algunas de las ideas de la PAT (1962a; también, 1962b).
- 12 "Finance" y "accounting" (ibid, 186). La idea aparece también en Guston (2000, 105) y la hemos visto en Braun (1998).
- 13 Para los interesados por un desarrollo de esta idea, referido a las consecuencias políticas y la estructura de

los consejos de investigación y a las carreras profesionales de los científicos, véase Fernández-Carro (2007, 2002).

- 14 Sobre esto, véase Fernández-Carro (2007).
- 15 Para una asunción humorística y rigurosa de lo segundo véase Wuffle (1999), quien propone la vía moderada del "elector razonable".
- 16 Para esta explicación sigo a Goldthorpe (2000, capítulo 6), quien discute además otros problemas.
- 17 Es en el estudio de los contratos laborales donde la teoría ha encontrado su mejor aplicación. Véase, por ejemplo, Epstein (1985).
- 18 Las recompensas son personales y, además de las materiales, pueden incluir la "satisfacción emocional" que explicaba Merton.
- 19 Como los citados Caswill (2003) y Morris (2003, 2000). Fernández-Carro (2007, 320-21) propone una forma doble de pago *retrospectivo*: un salario fijo (*insurance pay* en los términos de la literatura) y otro por incentivos (*incentive pay*).
- 20 Aunque muchos trabajos sobre economía de la ciencia no consideren a los actores estatales, gobiernos u organismos públicos, la naturaleza de bien público del conocimiento los trae de nuevo a la discusión.

REFERENCIAS

Abraham, M. y Prosch, B. (2000): "Long-term employment relationships by credible commitments: The Carl Zeiss Foundation", *Rationality and Society*, 12(3), 283-306.

Aghion, P. y Tirole, J. (1994): "The Management of Innovation", *The Quarterly Journal of Economics*, 109(4), pp. 1185-1209.

Alchian, A. A. y Demsetz, H. (1972): "Production, information costs, and economic organization", *American Economic Review*, 62(5), pp. 777-795.

Baert, P. (2001): *La teoría social en el siglo XX*, Madrid, Alianza.

Barnes, B. (1985): *About Science*, Oxford, Basil Blackwell.

Barnes, B. (1977): *Interests and the Growth of Knowledge*, Londres, Routledge and Kegan Paul.

Barnes, B. y Dolby, R. G. A. (1970): "The Scientific Ethos: a Deviant Viewpoint", *Archive of European Sociology*, 11.

Ben-David, J. (1971): *The scientist's role in society*, Englewood Cliffs, NJ, Prentice Hall.

Bendor, J. (1988): "Review Article: Formal Models of Bureaucracy", *British Journal of Political Science*, 18(3), pp. 353-395.

Bendor, J.; Taylor, S. y Van Gaalen, R. (1987): "Politicians, Bureaucrats, and Asymmetric Information", *American Journal of Political Science*, 31(4), pp. 796-828.

Bourdieu, P. (1975): "The Specificity of the Scientific Field and the Social Conditions of the Progress of Reason", *Social Science Information*, 14(6), pp. 19-47.

Braun, D. (2003): "Lasting tensions in research policy-making - A delegation problem", *Science and Public Policy*, 30(5), pp. 399-321.

Braun, D. (1998): "The role of funding agencies in the cognitive development of science", *Research Policy*, 27(8), pp. 807-821.

Braun, D. (1993): "Who Governs Intermediary Agencies? Principal-Agent Relations in Research Policy-Making", *Journal of Public Policy*, 13(2), pp. 135-162.

Braun, D. y Guston, D. H. (2003): "Principal-agent theory and research policy: an introduction", *Science and Public Policy*, 30(5), pp. 302-308.

- Bush, V. (1945): *Science, the Endless Frontier: A Report to the President on a program for postwar scientific research*, Washington, United States Government Print Office.
- Caswill, C. (2003): "Principals, agents and contracts", *Science and Public Policy*, 30(5), pp. 337-346.
- Caswill, C. (1998): "Social science policy: Challenges, interactions, principals and agents", *Science and Public Policy*, 25(5), pp. 286-296.
- Cole, J. y Cole, S. (1973): *Social Stratification in Science*, Chicago, Il, University of Chicago.
- Coleman, J. S. (1990): *Foundations of Social Theory*, Cambridge, Mass., Harvard University.
- Crosby, E. U. (1905): "Fire Prevention", *Annals of the American Academy of Political and Social Science*, 26 (Insurance), pp. 224-238.
- Dasgupta, P. y David, P. A. (1987): "Information Disclosure and the economics of Science and Technology", en G. R. Feiwel (comp.), *Arrow and the Ascent of Modern Economic Theory*, Londres, Macmillan, pp. 519-542.
- Dasgupta, P. y Maskin, E. (1987): "The Simple Economics of Research Portfolios", *Economic Journal*, 97(387), pp. 581-595.
- David, P. A. (2007): "The Historical Origins of 'Open Science' An Essay on Patronage, Reputation and Common Agency Contracting in the Scientific Revolution", *Documento de Trabajo 06-008*, Stanford University Department of Economics.
- David, P. A. (1998): "Common Agency Contracting and the Emergence of 'Open Science' Institutions", *The American Economic Review*, 88(2) (Proceedings of the 110th Annual Meeting of the American Economic Association), pp. 15-21.
- Eisenhardt, K. M. (1989): "Agency Theory: An assessment and review", *Academy of Management Review*, 14(1), pp. 57-74.
- Elster, J. (1986): "The Nature and Scope of Rational-Choice Explanation", en E. Lepore y B. P. McLaughlin, *Actions and Events*, Londres, Basil Blackwell, pp. 60-72.
- Epstein, R. A. (1985): "Agency costs, employment contracts, and labor unions", en J. W. Pratt y R. J. Zeckhauser (comps.), *Principals and Agents: The Structure of Business*, Boston, Harvard Business School, pp. 127-149.
- Ferejohn, J. A. (1999): "Accountability and Authority: Toward a Theory of Political Accountability", en A. Przeworski, S. C. Stokes y B. Manin (comps.), *Democracy, Accountability, and Representation*, Cambridge, Cambridge University, pp. 131-153.
- Ferejohn, J. A. (1986): "Incumbent Performance and electoral control", *Public Choice*, 50, pp. 5-25.
- Fernández-Carro, R. (2007): "A principal-agent model of public research with a retrospective payoff rule", *Science and Public Policy*, 34(5), pp. 317-328.
- Fernández-Carro, R. y Lapuente-Giné, V. (2006): "A Pied-Piper situation: Do bureaucratic researchers produce more science?", *Documento de Trabajo 8/2006*, Departamento de Ciencia Política y Sociología, Universidad Carlos III de Madrid.
- Fernández-Carro, R. (2002): *Regímenes políticos y actividad científica: Las políticas de la ciencia en las dictaduras y las democracias*, Madrid, Instituto Juan March.
- Ferrater Mora, J. (1979): *Diccionario de Filosofía*, Madrid, Alianza.
- Gibbons, M.; Limoges, C.; Nowotny, H.; Schwartzman, S.; Scott, P. y Trow, M. (1994): *The new production of knowledge. The dynamics of science and research in contemporary societies*, Londres, SAGE.
- Gieryn, T. F. (1995): "Boundaries of science", en S. Jasanoff et al. (comps.), *Handbook of science and technology studies*, Thousand Oaks, CA, Sage, 393-443.
- Goldthorpe, J. H. (2000): *On Sociology: Numbers, Narratives and the Integration of Research and Theory*, Oxford, Oxford University.
- Guston, D. H. (2003): "Principal-agent theory and the structure of science policy, revisited: 'science in policy' and the US Report on Carcinogens", *Science and Public Policy*, 30(5), pp. 347-357.
- Guston, D. H. (2001): "Boundary organizations in environmental policy and science: An introduction", *Science, Technology, & Human Values*, 26(4), pp. 399-408.
- Guston, D. H. (2000): *Between Politics and Science: Assuring the Integrity and Productivity of Research*, Cambridge, Cambridge University.
- Guston, D. H. (1999): "Stabilizing the boundary between US politics and science: The rôle of the Office of Technology Transfer as a boundary organization", *Social Studies of Science*, 29(1), pp. 87-111.
- Guston, D. H. (1996): "Principal-Agent Theory and the Structure of Science Policy", *Science and Public Policy*, 23(4), pp. 229-240.
- Hagstrom, W. O. (1965): *The Scientific Community*, New York, Basic Books.
- Hedström, P. y Swedberg, R. (1996): "Rational Choice, Empirical Research, and the Sociological Tradition", *European Sociological Review*, 12(2) (Rational Choice Theory and Large-Scale Data Analysis), 127-46.
- Kiewiet, D. R. y McCubbins, M. D. (1991): Capítulo dos, "Delegation and Agency Problems", en *The Logic of Delegation: Congressional Parties and the Appropriations Process*, Chicago, University of Chicago.

- Kiser, E. (1999): "Comparing Varieties of Agency Theory in Economics, Political Science, and Sociology: An Illustration from State Policy Implementation", *Sociological Theory*, 17(2), pp. 146-170.
- Kiser, E. (1994): "Markets and Hierarchies in Early Modern Tax Systems: A Principal-Agent Analysis", *Politics & Society*, 22(3), pp. 284-315.
- Kiser, E. y Kane, J. (2001): "Revolution and State Structure: The Bureaucratization of Tax Administration in Early Modern England and France", *American Journal of Sociology*, 107(1), pp. 183-223.
- Kiser, E. y Schneider, J. (1994): "Bureaucracy and Efficiency: An Analysis of Taxation in Early Modern Prussia", *American Sociological Review*, 59(2), pp. 187-204.
- Latour, B. y Woolgar, S. (1986): *Laboratory Life: The Construction of Scientific Facts*, Princeton, NJ, Princeton University, Reimpresión, 1979.
- Mari-Klose, P. (2000): *Elección Racional*, Madrid, Centro de Investigaciones Sociológicas.
- Mas-Colell, A.; Whinston, M. D. y Green, J. R. (1995): *Microeconomic Theory*, Oxford, Oxford University.
- McCubbins, M. y Schwartz, T. (1984): "Congressional Oversight Overlooked: Police Patrols versus Fire Alarms", *American Journal of Political Science*, 28(1), pp. 165-179.
- Merton, R. K. (1968): "The Matthew Effect in science: The reward and communication systems of science are considered", *Science*, 159(3810), pp. 56-63.
- Merton, R. K. (1957): "Priorities in scientific discovery: A Chapter in the sociology of science", *American Sociological Review*, 22(6), pp. 635-659.
- Milinski, M.; Semmann, D. y Krambeck, H.-J. (2002): "Reputation Helps Solve the 'Tragedy of the Commons'", *Nature*, 415(24), pp. 424-426.
- Miller, G. J. (2005): "The political evolution of Principal-Agent models", *Annual Review of Political Science*, 8, pp. 203-225.
- Miller, G. J. (1992): *Managerial Dilemmas: The Political Economy of Hierarchy*, Cambridge, Cambridge University Press.
- Miller, G. J. y Moe, T. M. (1983): "Bureaucrats, Legislators, and the Size of Government", *The American Political Science Review*, 77(2), pp. 297-322.
- Moe, T. M. (1984): "The New Economics of Organization", *American Journal of Political Science*, 28(4), pp. 739-775.
- Morris, N. (2003): "Academic researchers as 'agents' of science policy", *Science and Public Policy*, 30(5), pp. 359-370.
- Morris, N. (2000): "Science policy in action: Policy and the Researcher", *Minerva*, 38(4), pp. 425-451.
- Morrow, J. D. (1994): *Game Theory for Political Scientists*, Princeton, NJ, Princeton University.
- Mulkay, M. (1969): "Some Aspects of Cultural Growth in the Natural Sciences", *Social Research*, 36, pp. 22-52.
- Mulkay, M. (1991): *Sociology of Science: A sociological pilgrimage*, Bloomington, Open University.
- Nowakowska, M. (1975): "Measurable aspects of the concept of scientific career", en *Determinants and Controls of Scientific Development*, K. D. Knorr, H. Strasser y H. G. Zilian (comps.), Dordrecht, D. Reidel Publishing Company, pp. 295-322.
- Olsen, T. E. (1993): "Regulation of Multiagent Research and Development", *The RAND Journal of Economics*, 24(4), pp. 529-541.
- Polanyi, M. (1969): *Knowing and Being: Essays by Michael Polanyi*, Marjorie G. (comp.), Londres, Routledge and Kegan Paul.
- Polanyi, M. (1962): "The republic of science: Its political and economic theory", *Minerva*, 1, pp. 54-73.
- Pratt, J. W. y Zeckhauser, R. J. (comps.) (1985a): *Principals and Agents: The Structure of Business*, Boston, Harvard Business School.
- Pratt, J. W. y Zeckhauser, R. J. (1985b): "Principals and agents: An overview", en *Principals and Agents: The Structure of Business*, Boston, Harvard Business School, pp. 1-35.
- Rip, A. (1994): "The Republic of Science in the 1990s", *Higher Education*, 28(1), pp. 3-23.
- Shils, E. A. (1962a): "Politicians and scientists", *Encounter*, 18 (enero), pp. 103-107.
- Shils, E. A. (1962b): "The autonomy of science", en B. Barber y W. Hirsch (comps.), *The Sociology of Science*, Nueva York, NY, The Free Press of Glencoe, pp. 610-22.
- Shove, E. (2003): "Principals, agents and research programmes", *Science and Public Policy*, 30(5), pp. 371-381.
- Simon, H. A. (1957): *Models of man, social and rational: Mathematical essays on rational human behavior in a social setting*, Nueva York, John Wiley & Sons.
- Smith, B. L. R. (1994): "The United States: The formation and breakdown of the postwar Government-Science compact", en E. Solingen (comp.), *Scientists and the State. Domestic Structures in the International Context*, Ann Arbor, University of Michigan, pp. 33-61.
- Spence, M. y Zeckhauser, R. (1971): "Insurance, information, and individual action", *The American Economic Review*, 61(2), pp. 380-387.
- Stephan, P. E. (1996): "The Economics of Science", *Journal of Economic Literature*, 34(3), pp. 1199-1235.
- Tirole, J. (1994): "The internal organization of government", *Oxford Economic Papers*, 46, pp. 1-29.

- Turner, S. (1990): "Forms of Patronage", en S. E. Cozzens y T. F. Gieryn (comps.), *Theories of Science in Society*, Bloomington, IN, Indiana University.
- Van der Meulen, B. (2003): "New roles and strategies of a research council: inter-mediation of the principal-agent relationship", *Science and Public Policy*, 30(5): pp. 323-336.
- Van der Meulen, B. (1998): "Science Policies as Principal-Agent Games. Institutionalization and Path Dependency in the Relation between Government and Science", *Research Policy*, 27(4), pp. 397-414.
- Whitley, R. D. (2003): "Competition and pluralism in the public sciences: the impact of institutional frameworks on the organisation of academic science", *Research Policy*, 32(6), pp. 1015-1029.
- Whitley, R. D. (2000): *The Intellectual and Social Organization of the Sciences*, Reimpresión, 1984, Oxford, Oxford University.
- Williamson, O. E. (1999): "Public and Private Bureaucracies: A Transaction Cost Economics Perspective", *Journal of Law, Economics and Organization*, 15(1), pp. 306-347.
- Wolfson, M. A. (1985): "Empirical evidence of incentive problems and their mitigation in Oil and Gas Tax Shelter Programs", en J. W. Pratt y R. J. Zeckhauser (comps.), *Principals and Agents: The Structure of Business*, Boston, Harvard Business School, pp. 101-126.
- Wood, B. D. (1988): "Principals, Bureaucrats, and Responsiveness in Clean Air Enforcements", *The American Political Science Review*, 82(1), pp. 213-234.
- Wuffle, A. (1999): "Credo of a 'Reasonable Choice' Modeler", *Journal of Theoretical Politics*, 11(2), 202-206.