



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE FILOSOFÍA Y LETRAS

POSGRADO EN FILOSOFÍA

PSICOLOGÍA POPULAR: TEORÍA-TEORÍA O TEORÍA DE LA SIMULACIÓN

TESIS

QUE PARA OBTENER EL GRADO DE:

MAESTRO EN FILOSOFÍA

PRESENTA:

CARLOS HERNÁNDEZ MERCADO

DIRECTOR DE TESIS: DR. SILVIO PINTO

CIUDAD UNIVERSITARIA

JUNIO 2007



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mi Nube

Agradecimientos

Esta tesis de maestría fue realizada gracias al apoyo de profesores y compañeros de la Facultad de Filosofía y Letras y del Instituto de Investigaciones Filosóficas, ambos pertenecientes a la Universidad Nacional Autónoma de México. Quiero reconocer, y agradecer, la ayuda enorme que el Consejo Nacional de Ciencia y Tecnología (CONACYT) me brindó al haberme otorgado una beca por dos años pues me permitió dedicarme de tiempo completo a la realización de mis estudios (becario número 189491).

Mi más amplio agradecimiento para el Dr. Silvio Pinto, coordinador del Posgrado en Historia y Filosofía de la Ciencia de la UAM-I, quien mostró conmigo una eterna paciencia y una generosidad enorme al permitirme escuchar sus invaluable consejos y agudas observaciones. Su guía fue indispensable para la realización de esta tesis. No sólo me liga a él una profunda admiración sino también un enorme aprecio. Gracias por contagiarme de ese infinito amor por la filosofía.

Quiero hacer patente mi gratitud total a la Dra. Salma Saab, pues mi tema de investigación surgió gracias a la asignatura que cursé con ella. Su exposición hizo que sintiera por primera vez una emoción inefable ante un tópico filosófico. He aprendido mucho de su incansable y rigurosa manera de argumentar. Gracias a ella asistí por primera vez a un seminario de investigadores y quedé profundamente impactado por las críticas implacables que ahí se esgrimieron. Por fin, me daba cuenta cómo se practicaba la filosofía. No quisiera seguir sin antes expresar el gran cariño y respeto que le tengo.

Un agradecimiento especial a mis sinodales Dra. Ángeles Eraña, Dr. Gustavo Ortiz-Millán, y Dr. Miguel Ángel Fernández. Sus consejos y observaciones fueron absolutamente benéficos para mi investigación. Siempre recordaré sus agudas pero fraternales críticas. Mi reconocimiento y aprecio absoluto es para ellos.

Quisiera agradecer a la Dra. Maite Ezcurdia cuyas llamadas de atención fueron absolutamente oportunas. Mi admiración total a ella.

Finalmente, quisiera agradecer a las personas que estuvieron conmigo en la realización de este proyecto: Nube Estrada, que con su infinito cariño siempre me motivó a seguir adelante, sin ella no habría escrito ni una sola línea de esta tesis; Israel Grande-García, quien me ha brindado su amistad incondicional y con quien he tenido conversaciones agradablemente infinitas e infinitamente agradables; Kania Flores, por su amistad sincera; Marcos Rivera, cuya generosidad

nunca acabaré de agradecer; mis padres, Carlos Hernández y Antonia Mercado y mis hermanas Vianney, Elisa y Yessica.

A todos mi mayor reconocimiento y gratitud.

Carlos Hernández Mercado

Índice	i
Introducción	1
Capítulo 1. La Teoría-Teoría	
Introducción: ¿En qué sentido la PP puede ser una teoría?	6
1. La PP como una Teoría Científica (primera parte)	7
2. La PP como un Módulo	17
3. La PP como una Teoría Científica (segunda parte)	31
4. La PP como una Teoría Funcionalista	39
Capítulo 2. La Teoría de la Simulación	
Introducción: Una Propuesta Alternativa	43
1. Teoría de la Simulación: Gordon (1986)	44
2. Teoría de la Simulación: Goldman (1989)	55
3. Retos a la Teoría de la Simulación	62
Capítulo 3. Hacia una Teoría Híbrida	
Introducción	69
1. Argumentos	69
2. Posible Estrategia de Integración	85
3. Conclusión	90
Referencias	92

INTRODUCCIÓN
Psicología Popular

Los psicólogos Heider y Simmel (1944) hicieron un estudio con tres grupos de estudiantes. Se les mostró una película en la que tres figuras geométricas (dos triángulos y un círculo. Fig. 1) realizaban distintos movimientos en torno a, y dentro de, un rectángulo.¹

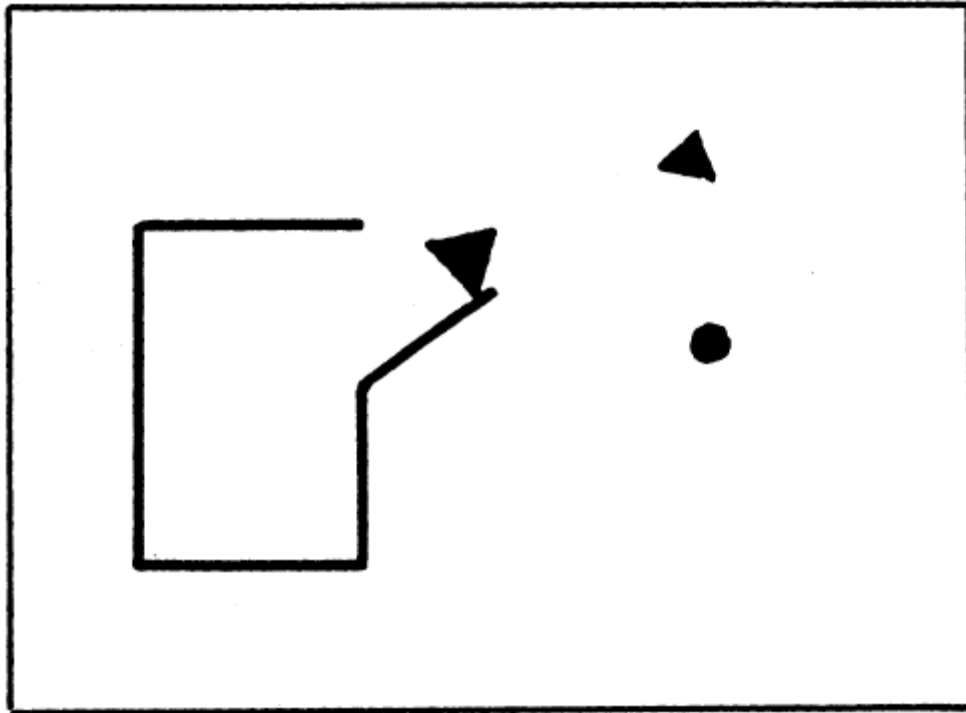


Fig. 1. Figuras del experimento de Heider y Simmel (1944).

Las figuras se acercaban y alejaban entre sí, en ocasiones de manera rápida y en ocasiones de manera lenta. A veces tenían contacto repentino y a veces contacto pausado. A las estudiantes se les solicitó que dieran una explicación de lo que ocurría. Todos los reportes del primer grupo (34 mujeres), excepto uno, contenían interpretaciones mentalistas de los movimientos de los polígonos. Los sujetos de estudio explicaban libremente lo sucedido en el cortometraje con frases como “El primer hombre le dice al segundo que se marche; el segundo le dice lo mismo al primero y mueve su

¹ Heider y Simmel (1944), p. 244.

cabeza. Entonces se inicia una pelea entre los dos hombres... La muchacha se empieza a preocupar y corre de una esquina a otra en el fondo de la habitación... (Los triángulos son vistos como sujetos masculinos y el círculo como femenino) El primer hombre retrocede y trata de abrir su puerta, pero está cegado por la rabia y la frustración que le provoca no poder abrirla...".² Al segundo grupo (36 mujeres) se le hicieron preguntas no neutrales como: ¿qué tipo de persona es el primer triángulo?³ Las respuestas no fueron menos asombrosas el 97% de las estudiantes respondió con adjetivos como agresivo, belicoso, beligerante, pugnaz, peleón, problemático, malo, enfadado, de mal temperamento, temperamental, irritable, etc. Cuando se les pregunto sobre el triángulo pequeño respondieron: heroico, valiente, valeroso, desafiante, no se deja someter, etc.⁴ Estos resultados son intrigantes. ¿Cómo podemos explicar que los sujetos del estudio identifiquen sistemáticamente el movimiento como conducta intencional? ¿Por qué los autistas no pueden hacer lo mismo?

Esta investigación tiene como propósito presentar algunas respuestas que se han formulado ante estas y muchas preguntas más que surgen cuando investigamos la psicología del hombre ordinario.

Para explicar y predecir conductas las personas atribuimos estados mentales a nosotros mismos y a los otros. Cuando un agente ejecuta una acción buscamos qué creencias, deseos, esperanzas, intenciones, sensaciones, emociones etc. pudieron haberla causado. Frecuentemente creemos que si un agente tiene tales o cuales estados mentales realizará tal o cual acción. Frases del siguiente tipo son comunes: el conductor quiere llegar rápido, buscó en la alacena pues cree que ahí hay chocolates, los inversionistas son pesimistas ya que no creen que los mercados repunten, los políticos sólo desean poder, el editor rechazará la columna si se percata de que tus fuentes mienten, etc.

Esta manera de hablar sobre la conducta de la gente implica un cierto tipo de psicología pues supone varias cosas: que los individuos son sistemas intencionales, que

² Heider y Simmel (1944), pp. 246 y 247.

³ Digo no neutrales porque ya implicaban una interpretación de la figura como agente intencional al referirse a las figuras como "personas".

⁴ Heider y Simmel (1944), pp. 248 y 249.

tienen mentes, que poseen un conocimiento implícito de relaciones entre percepciones, estados mentales, y acciones, y que se conducen de acuerdo con él. Explicar una conducta es hacer explícitas las relaciones que tiene con ciertos estados mentales e insumos. De manera similar, predecir una conducta es determinar qué salida conductual será producida a partir de otros tantos estados mentales y estímulos externos. Pero todo esto tiene sentido sólo si consideramos que los demás tienen mentes como las nuestras y que su conducta es intencional (como la nuestra). La posibilidad de predecir y explicar las acciones nos hace ser capaces de “leer las mentes de los otros”. El comportamiento de un agente nos habla de sus estados psicológicos internos, y al conocer estos últimos podemos predecirlos.

Los seres humanos en contextos sociales cotidianos asumimos que todos se conducen según lo que creen y lo que quieren; somos realistas de los estados mentales. Tal suposición está en el núcleo de este tipo de psicología. El empleo de tal suposición, y por ende, de este núcleo mentalista, es universal. Efectivamente, todas las personas independientemente de su origen, pero con un estímulo apropiado, desarrollan esta capacidad y se relacionan con los otros gracias a ella. Como bien señala Fodor (1987), esta psicología nos sirve para interactuar con conocidos y con extraños, ya sea con personas de culturas semejantes o de culturas totalmente diferentes. Todos los agentes exhiben aproximadamente el mismo nivel de dominio y parecen haberlo alcanzado en las mismas etapas de desarrollo. Lo más interesante de esto es que no parece haber existido instrucción explícita sobre cómo emplear esta estrategia y aún así es ubicuamente usada. Esta herramienta psicológica de sentido común ha sido llamada Psicología Popular (*Folk Psychology*).⁵ El nombre surge en gran parte porque esta estrategia no posee el refinamiento teórico-conceptual de una teoría psicológica (y científica) acabada, y además, porque su uso es intuitivo y universal. Del mismo modo existen: una biología popular que consiste en las ideas de sentido común que se tienen acerca de los seres vivos, (que la naturaleza funciona de tal o cual modo, que las seres vivos crecen o se modifican con el tiempo, etc.), y una física popular que consiste en el conjunto de saberes

⁵ En adelante PP.

pre-teóricos que la gente en la vida cotidiana supone sobre los objetos, (como que las cosas caen hacia abajo o que el movimiento puede ser transmitido, etc.).

La universalidad en el uso de la PP, la coincidencia en las etapas de desarrollo, y su aprendizaje no explícito han apoyado la idea de que la PP constituye una competencia psicológica propia del ser humano y no una mera estrategia explicativa/predictiva accidental.

Si lo anterior es el caso, es decir, si la PP es una competencia constitutiva del ser humano, entonces tiene que existir un mecanismo que posibilite su aprendizaje o desarrollo. Su aprendizaje si es que se adquiere completamente a partir de insumos ambientales, o su desarrollo si resulta ser innata pero necesita de determinados estímulos para activarse y cumplir con su ontogenia típica.

Premack y Woodruff (1978) pensaban que la capacidad de atribuir estados mentales a un agente era equivalente a poseer una teoría de la mente. Había dos razones para considerar a esta capacidad como una teoría: a) la postulación de inobservables (estados mentales), y b) la capacidad de predecir el comportamiento de los demás.⁶ Las teorías científicas normalmente postulan inobservables y permiten hacer predicciones, por tanto, la inferencia analógica no era difícil de extraer. Si todos los psicólogos y filósofos hubieran quedado convencidos con este criterio doble, probablemente nunca habría nacido una fructífera discusión como la actual. Afortunadamente, el escenario fue otro pues surgieron desacuerdos sobre si, efectivamente, era suficiente postular inobservables y permitir hacer predicciones para tener una teoría.

En caso de ser una teoría debería tener algo como leyes o principios generales. Sin embargo, los usuarios no parecen ser conscientes de ellos. La solicitud de hacer explícitas las leyes de esta "teoría" se convirtió gradualmente en un argumento en su contra por la dificultad encontrada para satisfacer esta petición. Al momento de buscar enunciar los principios rectores de esta habilidad, las personas que cotidianamente emplean términos mentalistas encuentran complicaciones para formularlos a pesar de que se supone rigen

⁶ Premack, D. y Woodruff, G. (1978), p. 515.

su conducta y sus explicaciones de ella; más aún, cuando se les presiona para que lo hagan sólo ofrecen generalizaciones que muchos consideran triviales o accidentales. Así, tenemos dos hechos claros: por un lado la PP parece ser poseída por todos, pero, por otro lado, nadie parece ser capaz de explicitarla.

La discusión ha dividido a filósofos, científicos cognitivos, y psicólogos, en dos bandos: aquellos que favorecen la teoría de que subyace una teoría a la PP y aquellos que postulan un mecanismo de simulación “fuera de línea” como componente esencial de nuestra comprensión, explicación y predicción de los demás y que niegan cualquier apoyo teórico. Entre estos dos extremos han surgido versiones menos radicales de las teorías originales que han incorporado algún elemento de las concepciones en principio antagónicas.

El texto, dividido en tres capítulos, tiene como objetivo esclarecer cuál es la base de la Psicología de creencias y deseos. En la primera parte exploro cuatro sentidos según los cuales se puede entender que la PP es, o tiene que ver con, una teoría. La idea final de la sección es que las diferentes concepciones vistas de la PP no son suficientes para dar cuenta de ella. En la segunda parte abordo la alternativa conocida como Teoría de la Simulación. Analizo dos versiones de la misma y concluyo que tampoco son suficientes para explicar la PP. Por último, argumento que sería conveniente pensar en una concepción híbrida frente a ciertos problemas que presentan casos específicos. Añado como posible componente de esta teoría la noción de “Marco” de Minsky (1974) y propongo una estrategia para empezar a integrar las dos alternativas originalmente antagónicas. Concluyo favoreciendo la idea de teoría híbrida.

La respuesta correcta sobre el mecanismo que subyace al uso de la PP debe dar cuenta de cuestiones como coincidencia de desarrollo, universalidad de posesión, fallas de adquisición, y otras que veremos con detalle en el transcurso de la investigación.

Capítulo I

La Teoría-Teoría

Introducción

¿En qué sentido la PP puede ser una teoría?

Hay diversas concepciones sobre la PP que pueden agruparse bajo el nombre “Teoría-Teoría”. Este hecho se explica por que los partidarios de cada variante sostienen que una teoría es un componente esencial de la PP o que ella misma es una teoría. A pesar de esta convergencia, el sentido en que se toma la palabra “Teoría” es diferente en cada caso.

Ravenscroft (2004) explica que hay dos sentidos en los cuales podemos entender a la PP: a) como una teoría implícita acerca de nuestra manera de hablar sobre estados mentales que introduce funcionalmente términos teóricos como “creencia” y “deseo” (visión externa), o bien, b) como una teoría de la psicología humana que está representada en la mente-cerebro y que sostiene o apoya nuestra capacidad de hacer inteligibles las conductas (visión interna). En otras palabras, la psicología popular puede verse, o bien, como una teoría adquirida empíricamente que le da significado a los términos que la constituyen, o bien, como una facultad interna de lo mental que nos hace posible interactuar con los demás seres humanos.⁷ En ambos casos, la noción de “teoría” que se maneja tiene que ver con un conjunto de principios coherentemente estructurados que permiten trazar las relaciones entre los componentes de la misma.

Botterill y Carruthers (1999) favorecen la idea de que la PP es una teoría interna y argumentan que tanto las personas normales la poseen como individuos con ciertas deficiencias mentales (por ejemplo, niños con síndrome de Down o con síndrome de Williams, un tipo especial de autismo, hacen atribuciones de estados mentales y explican la conducta intencional en esos términos). La TT básicamente presupone que el organismo en cuestión posee un cuerpo de conocimientos, organizados y estructurados, al cual

⁷ Ravenscroft, I. (2004) “Folk Psychology as a Theory”, *The Stanford Encyclopedia of Philosophy* (Winter 2004 Edition, Feb.-23-2004), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2003/entries/folkpsych-theory/>.

recurre para realizar explicaciones y predicciones. Éstas últimas son producto de las inferencias realizadas a partir de una base de datos, para la mayoría de los autores tácita, constituida por principios generales sobre entidades psicológicas.

Las diferentes versiones de la TT normalmente son compatibles entre sí y es frecuente que sus partidarios estén a favor de emplear simultáneamente más de un sentido en sus explicaciones. En este capítulo expondré 4 diferentes sentidos en los que se puede entender a la PP como una teoría.

1. La PP como una Teoría Científica (Primera Parte)

Algunos defensores de la TT sostienen que la PP es una teoría precisamente porque tiene las características que constituyen a las teorías científicas. Como veremos a continuación, existen varias similitudes estructurales que favorecen una concepción teórica de la PP. Sin embargo, también existen ciertas diferencias importantes que pueden inclinarnos a rechazar esta analogía como son: que las teorías científicas son aprendidas y que sus principios están formulados explícitamente en contraste con la PP. Los objetivos en esta sección son: exponer esta versión de la TT y ver si resulta satisfactoria como explicación de la PP.

Las teorías científicas son sistemas de entidades abstractas (gen, átomo, fuerza, etc.) relacionadas mediante leyes normalmente enunciadas en un vocabulario basado en no-observables. Las leyes de la teoría establecen relaciones causales entre los fenómenos observables y los no-observables. Las teorías permiten predecir nuevos sucesos así como explicarlos. Los conceptos de las teorías no siempre son definidos de modo explícito. En ocasiones pueden ser definidos al establecer sus interacciones con otros conceptos de la misma teoría (podríamos decir, por ejemplo, que un gen es lo que determina el genotipo y fenotipo de los seres vivos). A su vez, las teorías buscan una integración coherente y sistemática de sus leyes.

George Botterill (1996) proporciona una lista de características que normalmente poseen las teorías científicas:

1. Suministran explicaciones y predicciones a través de leyes;
2. Sus leyes tienen proyección contrafáctica;
3. Postulan inobservables causalmente eficaces;
4. Definen implícitamente a sus conceptos;
5. Contienen principios que proporcionan una integración sistemática del conocimiento.⁸

Si la PP tuviera estas cinco características entonces podría ser considerada como una teoría científica.⁹ Explicaré en lo siguiente en qué sentido la PP podría poseerlas. Es importante resaltar que aisladamente estas condiciones no son suficientes para que la PP sea una teoría. El estatus teórico se obtiene a partir de la posesión de todas. Mi tarea en esta parte es exponer cada una de ellas y concluir si la PP es una teoría en este sentido.

1. Las teorías suministran explicaciones y predicciones a través de leyes. Dos de los requisitos más importantes que tiene que satisfacer una teoría son la explicación y la predicción. La formulación de teorías obedece, entre otras cosas, a la búsqueda de inteligibilidad en un fenómeno. Si se quiere comprender lo que vemos, buscamos una explicación. Cuando el acaecimiento de un evento suscita la duda de por qué ocurrió, o por qué lo hizo de determinado modo, entonces se recurre a una teoría que muestre las relaciones causales que lo detonaron. Las predicciones y explicaciones se dan en virtud de ciertas leyes. El científico apela a ellas cuando enfrenta un evento que pretende explicar. Si quiero explicar porqué el volumen de un gas contenido en un cilindro se incrementa cuando aumento su temperatura, recurro a la ley que relaciona el volumen con la temperatura. La ley que utilizamos nos muestra las relaciones que se presentan entre los elementos involucrados en el fenómeno. Algo similar ocurre cuando formulamos la

⁸ Botterill (1996), pp. 106, 107.

⁹ Lo formulo contrafácticamente porque aún no está claro que las posea.

pregunta: ¿Por qué ese sujeto abrió su paraguas? La PP nos da explicaciones para este caso. La respuesta acostumbrada es: “por que ese sujeto no quiere mojarse”. Se recurre a una especie de principio nomológico que relaciona deseos y creencias con acciones.¹⁰

Si una persona no está satisfecha con la explicación que se le brinda cuando pregunta por qué alguien abrió su paraguas es posible añadir que el sujeto abrió su paraguas porque no deseaba mojarse, creía que el agua moja y que abriendo el paraguas y cubriéndose con él lograría satisfacer su deseo y no tenía otras creencias o deseos que se opusieran a la realización de su acción. Lo anterior es normalmente suficiente para hacer inteligible la acción del sujeto.

Lo que se hace al dar respuesta a una pregunta psicológica es mostrar la relación entre estados psicológicos e insumos y estados mentales y conducta. Postulo la ejemplificación de interacciones causales entre estados mentales y estímulos que llevan a los sujetos a actuar de cierto modo. Hasta este punto la PP se desempeña adecuadamente en el ámbito de la explicación. Otorga inteligibilidad a los fenómenos que ocurren mostrando qué sucesos los provocaron. A su vez brinda explicaciones apelando a principios de tipo nomológico como “para todo S, Q, y R, si S desea que P y cree que Q producirá que P, entonces S hará Q”. Así como las teorías científicas nos dan explicaciones de los fenómenos empleando principios generales, la PP también lo hace.¹¹ Por tanto, dado que ambas comparten este rasgo, la intuición de que son semejantes procede.

Otro factor crucial en una teoría es la predicción.¹² Si efectivamente hay principios que caracterizan correctamente ciertas interacciones entre sucesos entonces nuevas

¹⁰ El punto a resaltar es que ciertos principios generales como “si S quiere $\neg P$ y cree que $Q \rightarrow \neg P$, entonces S hará Q” son empleados para dar explicaciones de las acciones. Un objetor podría rehusarse a ver este principio como ley científica. Considero que tendría razón, sin embargo, sigo sosteniendo, con Boterill (aunque no estoy de acuerdo con su hipótesis general totalmente), que aún así es un principio general que empleamos para dar una explicación de una acción. Para fines prácticos concedamos que hay alguna semejanza explicativa entre principios generales de la PP y leyes científicas.

¹¹ La diferencia, como veremos en la crítica a esta propuesta, es que toda ley es un principio general pero no todo principio general es una ley.

¹² Cabe señalar que la predicción de conductas es diferente de su anticipación. Normalmente la divergencia radica en los recursos lingüísticos empleados. Si S predice algo, lo hace valiéndose en parte de su facultad lingüística. Pero si S anticipa una conducta no es necesario este componente verbal. En ciencia la predicción siempre se da en cierto lenguaje y a partir de esta predicción se pueden anticipar los fenómenos. En cambio, podemos decir que un mamífero anticipa un golpe, y lo evita en consecuencia, pero difícilmente diremos

acciones deben ser subsumidas por aquellos principios. La consecuencia inmediata de tener estas generalizaciones es que pueden servir para trazar un mapa del curso que tomarán los eventos.

En el caso de la psicología popular esto quiere decir que las personas somos capaces de anticipar y predecir varias cosas que pueden ocurrir a los demás o a nosotros mismos. Podemos prever en qué estado psicológico se encuentra alguien o qué acción realizará.¹³ Como bien sabemos, en el ámbito de la psicología popular se formulan predicciones con una alta tasa de éxito, situación que parece indicar que ciertas conjunciones de estados mentales e insumos están relacionadas efectivamente con la aparición de otras conjunciones de estados mentales y de conducta. De hecho, la interacción humana presupone que estas relaciones se presentan con cierta uniformidad en la conducta intencional. Cuando preguntamos a alguien la hora, por ejemplo, asumimos que ese sujeto tiene estados mentales con contenido intencional pues esperamos que al escuchar la frase “¿qué hora es?” se presentará determinada conducta (nos informará qué hora es). Si sabemos que un sujeto siente frío y sabemos que no le agrada esa situación entonces podemos predecir que buscará algún modo de suministrarse calor. Si sabemos que un individuo cree que el café está en la cafetera y sabemos que desea tomar café entonces podremos predecir su curso de acción. En cierto sentido, podemos predecir el futuro leyendo las mentes de los otros a partir de la observación de su conducta. La mayoría de las veces nuestras expectativas se verán cumplidas y conoceremos con anticipación como se desarrollarán los eventos. Como hemos podido observar, la psicología popular cumple también con la condición de habilitarnos para hacer predicciones. Por lo tanto, ya que la PP permite dar explicaciones y hacer predicciones al igual que lo hace una teoría científica, puede verse del mismo modo.¹⁴

que lo ha predicho.

¹³ Una teoría que no ofrezca la posibilidad de hacer predicciones reduce sus explicaciones a un nivel anecdótico pues no muestra ninguna relación efectiva entre los sucesos. Esto por el mismo hecho de que si se dan estas relaciones efectivas entonces servirán para anticipar la ocurrencia de los fenómenos. Este no es el caso de la PP.

¹⁴ El objetor de esta perspectiva fácilmente descartará la analogía mostrando que los principios que se

2. Las leyes de las teorías científicas tienen proyección contrafáctica. Una de las características que distingue a los principios nómicos genuinos de las simples generalizaciones es la posibilidad de soportar una formulación contrafáctica. En la teoría cinética se puede ver como una ley que relaciona la temperatura con el volumen de los gases soporta esta formulación. Si la temperatura no se cambiara entonces el volumen tampoco lo haría, (concediendo que la presión se mantenga constante). Botterill sostiene que la PP soporta formulaciones contrafácticas en casos como: *“If Maxi had not believed the chocolate was there, he wouldn’t have opened the wrong cupboard”*.¹⁵ En casos como “si x no hubiera creído que p entonces x no habría hecho q” se formula el principio “si x cree que p entonces x hará q” contrafácticamente mostrando que se sostiene. Un ejemplo de una generalización que no soporta una formulación contrafáctica sería el enunciado “todas las casas que están en Coyoacán valen más de 100 mil pesos” pues al ser formulado de manera contrafáctica (si la casa no estuviera en Coyoacán entonces no valdría más de 100 mil pesos) deja de ser verdadero. Por tanto, si los principios de la PP soportan la proyección contrafáctica entonces no sólo son generalizaciones empíricas accidentales sino que son principios nómicos genuinos.¹⁶

3. Las teorías científicas postulan inobservables. Botterill escribe al respecto: *“This was what the kinetic theory of gases did and why it operates at a deeper level than a macroscopic generalization such as Boyle’s law”*.¹⁷ En el caso de la psicología, para explicar un comportamiento humano (observable) se postula generalmente una cadena causal de estados mentales (inobservables). Bajo esta perspectiva, la PP tendría una estructura profunda como la de casi todas las teorías científicas. Al igual que ellas, las generalizaciones causales que le subyacen se establecen sobre no observables (estados

emplean en una y otra teoría para explicar y predecir eventos no son del mismo tipo. Reconozco esa limitación y hablaré posteriormente de ella.

¹⁵ Botterill (1996), p. 107. “Si Maxi no hubiera creído que el chocolate estaba ahí, no habría abierto la alacena equivocada”. Más adelante veremos que este rasgo también presenta ciertos problemas para la PP. Cabe señalar que no comparto esta idea de Botterill. (Todas las traducciones son mías).

¹⁶ La verdad del antecedente de esta implicación no es indisputable.

¹⁷ Botterill (1996), p. 108. “Esto fue lo que hizo la teoría cinética de los gases y es la razón de que opere a un nivel más profundo que una generalización macroscópica como la ley de Boyle”.

mentales) y esas generalizaciones conducen a determinadas predicciones.¹⁸ La postulación de inobservables en el sentido anterior le otorga profundidad a la teoría pues explica los fenómenos observables a partir de interacciones con entidades no observables. Las explicaciones que suministra la PP están formuladas en estos términos. Si deseo beber café y creo que la taza contiene café generalmente bebo el café de la taza. La acción de beber el café de la taza se explica en este caso en términos de creencias y deseos. No obstante estos estados intencionales no están visibles propiamente en la conducta; son accesibles en virtud de la relación que guardan con ésta. El comportamiento intencional de un sujeto es causado por la convergencia de insumos y combinaciones de estados psicológicos no observables.

4. Las teorías científicas definen implícitamente sus conceptos. En la ciencia no siempre se hacen definiciones tradicionales en términos de género y diferencia específica. De hecho, no ocurre normalmente que se den listas de términos con su significado. Lo común es enunciar los principios de la teoría y de ahí extraer las definiciones implícitas de sus términos. El empleo de los términos de la PP presupone, si se hace un uso correcto de ella, la posesión de una teoría que los define implícitamente por medio de sus interacciones causales. Es el papel causal en la teoría lo que le da en este caso el significado al concepto. Por ejemplo, el concepto de creencia puede definirse como aquello que en conjunción con percepciones y deseos puede derivar o resultar en cierto tipo de conducta.

5. Las teorías científicas contienen principios que proporcionan una integración sistemática del conocimiento. Botterill señala lo siguiente en referencia a esta condición: "Theories produce cognitive economy through integration of information in a small number of general principles".¹⁹ Una teoría en este sentido debería poseer un cuerpo integrado y limitado de leyes que por su generalidad sean aplicables a una cantidad

¹⁸ Fodor (1987), p. 7.

¹⁹ Botterill (1996), p. 109. "Las teorías producen economía cognitiva por medio de la integración de información en un número pequeño de principios generales".

considerable de casos. Esta idea tiene la ventaja de permitir agrupar una cantidad de fenómenos bajo ciertos principios básicos identificando sus similitudes. En el caso de la psicología popular sucede algo análogo, aunque no idéntico, a lo que sucede con las leyes de las teorías científicas. Algunos principios como “si x quiere que p y x cree que haciendo q alcanzará p, entonces si no hay otros estados mentales que se opongan a q, x hará q” parecen englobar una cantidad considerable de fenómenos y permiten dar explicaciones y predicciones adecuadas en parte por su generalidad. Podemos insertar cualquier proposición en lugar de “p” y “q” y el nombre de cualquier sujeto en lugar de “x” para crear un ejemplar de este principio.

Si la PP permite dar explicaciones y hacer predicciones, sus principios tienen proyección contrafáctica, postula relaciones entre inobservables, define implícitamente a sus conceptos, y contiene principios que proporcionan una integración sistemática del conocimiento, entonces, la PP tiene las mismas características que la mayoría de las teorías científicas. Por lo tanto, su estatus está garantizado.

Hasta aquí la exposición de esta primera versión de la teoría-teoría. Para Botterill (1996) el estatus teórico de la PP está garantizado en virtud de que tiene las 5 características que presentan las teorías científicas.

Ahora realizaré un breve comentario de la propuesta recién expuesta objetando dos de sus puntos principales. Según los párrafos anteriores, si la PP tuviera los rasgos mencionados sería estructuralmente semejante a una teoría y de este modo, se favorecería la concepción de los partidarios de la TT. Quiero señalar que varias de las características que pone Botterill son cumplidas por la Psicología de creencias y deseos de un modo bastante laxo. La analogía de la teoría no es del todo precisa si observamos con cuidado. Considero que entender a la PP como una teoría, según los criterios esbozados, expone a la TT a muchos problemas y objeciones innecesarias. Entre ellas, las objeciones de los materialistas eliminativos que consisten, brevemente, en ver a la PP como una teoría falsa, estática e incompleta. Por ello creo que es más conveniente explorar otras

alternativas que sean menos sensibles a ese tipo de ataques. Por el momento, enfoquémonos en algunos detalles interesantes de la analogía de Botterill.

En primer lugar, tanto las teorías científicas como la PP ofrecen predicciones y explicaciones. Pero en el caso de las teorías científicas se emplean leyes que subsumen a los casos singulares. Si queremos predecir un fenómeno entonces recurrimos a una ley que nos hable de sus relaciones causales. Si queremos explicarlo buscamos el principio nomológico que describa la causa de su aparición. ¿Acaso podemos decir que los principios de la PP son iguales a las leyes científicas? Considero que la respuesta es negativa en varios sentidos. Las relaciones entre insumos, estados mentales y acciones no son completamente directas como en el caso de las relaciones causales de los fenómenos explicados por otras ciencias como la física o la química. Las leyes científicas y psicológicas se dan acompañadas de cierto número de condiciones contextuales, sin embargo, en el caso de las segundas el resultado no siempre se obtiene aún cuando gran parte de los elementos importantes están presentes. Comparemos la ley de la caída libre que señala que en la tierra los cuerpos caen a 9.8 metros por segundo cuadrado con el principio psicológico que señala que “si x desea que q y $(p \rightarrow \neg q)$ entonces x evitará p”. Para que las predicciones que se obtienen de su aplicación sean exactas varios requisitos importantes tienen que ser cumplidos. Sin embargo, parece que siempre es posible encontrar situaciones en las cuales el principio psicológico no se cumple (y siempre es posible agregar cláusulas *ceteris paribus* para salvar la generalización). En este punto son bastante diferentes las leyes científicas y los principios generales de la PP. Más aún si tomamos en cuenta lo que dice Davidson (1970) sobre la anomalía de lo mental (a saber, que no hay leyes psicofísicas estrictas y que no hay leyes estrictas con base a las cuales podamos predecir y explicar los fenómenos mentales), entonces la analogía de Botterill pierde fuerza.²⁰

El punto de Botterill fue que tanto las teorías científicas como la PP permiten hacer predicciones y dar explicaciones. Esto es cierto, sin embargo, los recursos empleados por cada una son diferentes entre sí. Por un lado, tenemos leyes que parecen ser estrictas y,

²⁰ Davidson, (1970) p. 117

por otro lado, tenemos principios que admiten muchas excepciones. A mi juicio, la analogía no puede ser correcta a menos que sólo estemos concentrando la comparación en las ciencias especiales, pues precisamente un rasgo esencial de estas es el de admitir excepciones. Teniendo en mente la concepción de Botterill parece apropiado señalar que toda ley es un principio general, pero no todo principio general es una ley. Los principios generales se aplican a muchos casos semejantes pero no se ve claramente que haya causalidad, y menos necesidad en ellos. El tema del carácter de las leyes es muy controvertido como para ocuparme de él en este texto, pero no sería incoherente afirmar que las leyes hablan de tipos de relaciones causales y en el caso de la PP no parece respetarse esta formulación en términos generales. Tal vez los principios generales puedan algún día descubrirse como leyes, no es necesario que no suceda. Pero por el momento, hay que aclarar que no son lo mismo.

El segundo punto de la lista de Botterill es la formulación contrafáctica. En filosofía de la ciencia ha sido frecuente someter a los enunciados que se quiere examinar si son legaliformes a una formulación de este estilo. Se supone que las simples generalizaciones no se sostendrán ante esta prueba pero que las leyes auténticas sí lo harán. Botterill propone que los principios de la PP se sostienen contrafácticamente. Su ejemplo involucra al famoso Maxi de Wimmer y Perner (1983) y dice: “Si Maxi no hubiera creído que el chocolate estaba ahí, no habría abierto la alacena equivocada”. Sin embargo, es posible imaginar un caso donde Maxi quiere despistar a su abusivo hermano. Esto representa una situación competitiva para Maxi pues él desea el chocolate al igual que su antagonista. La estrategia para conservar la golosina es, precisamente, llevar a su hermano al lugar equivocado. En este caso es cierto que Maxi no cree que el chocolate está en esa alacena, pero la abre para desviar a su hermano. Por tanto, el contrafáctico no se sostiene para todos los casos y resulta que el principio de Botterill no es legaliforme.

Puede mostrarse que muchos principios de la PP sufren precisamente este mismo problema. Tan sólo pensemos en el enunciado (a) “Si x tiene hambre entonces x buscará comida” y su formulación contrafáctica (b) “si x no tuviera hambre entonces x no buscaría

comida". Es razonable considerar al enunciado (a) como un principio de la PP. Es claro que la gente que tiene hambre se procura algún tipo de alimento y que gracias a ese conocimiento podemos predecir la conducta de cualquiera que nos confiesa que tiene hambre. Sin embargo, mucha gente come a pesar de no tener hambre y mucha gente que tiene hambre no come. Nuevamente esto produce un fallo en el contrafáctico general pues se rompe la dependencia causal entre tipos de estados. Lewis (1970) señala que existe dependencia causal si para dos sucesos diferentes e y c expresados por las oraciones " Oe " y " Oc " donde " O " quiere decir "ocurre que", " $Oc \rightarrow Oe$ " y " $\sim Oc \rightarrow \sim Oe$ " son el caso.²¹ El aspecto relevante en lo anterior es que, nuevamente, la analogía de Botterill parece estar lejos de ser exacta. Los principios de las teorías científicas soportan contrafácticos pero esto no ocurre sistemáticamente con los supuestos principios de la PP. Por consiguiente, no podemos estar justificados en afirmar que la PP es una teoría en el mismo sentido que lo es una teoría científica.

La primera perspectiva que hemos visto afirmaba que la PP era una teoría pues tenía todos los rasgos que presentan las teorías científicas. Sin embargo, hay por lo menos, dos rasgos que presentan problemas al trazar la analogía: los tipos de leyes que se emplean en una y otra teoría son diferentes y los principios de la PP no soportan contrafácticos sobre tipos de estados.

En la siguiente sección analizo algunos aspectos de la concepción modular de la TT. La idea principal de esta postura es que la PP es sostenida, o está constituida, por un módulo donde se concentra la teoría de la mente.

²¹ Lewis (1970), p. 562

2. La PP como un Módulo

Muchos partidarios de la TT creen que la psicología popular es una teoría tácita. Esto significa que las predicciones y explicaciones de la conducta son gobernadas por una teoría interna de la cual no se tiene conciencia todo el tiempo. Normalmente se sostiene que el grado de acceso a sus principios indica si la teoría es fuerte o débilmente tácita. Si la teoría es tácita en un sentido fuerte entonces sus principios no pueden ser explicitados por el usuario, ya que no tiene acceso a ellos. Por otro lado, si la teoría es débilmente tácita entonces algunos de sus principios pueden estar disponibles conscientemente para hacerse explícitos. Ha habido varias maneras de concebir este rasgo. Una de ellas ha sido emplear la idea de módulo para explicar su inaccesibilidad. Aunque lo tácito no implica necesariamente lo modular vale la pena explorar este terreno. Eraña (2006) explica que un módulo es clásicamente visto como un mecanismo o sistema de conocimiento de dominio específico, encapsulado, rápido, obligatorio, etc.²² Fodor (2000), por su parte, señala que un módulo es un sistema computacional con una base de datos específica.²³ Ahora sólo me concentraré en dos aspectos de los módulos: la especificidad de dominio y el encapsulamiento informacional. Un módulo tiene especificidad de dominio porque la información que procesa sólo es de cierto tipo y si hay información disponible que no posee las propiedades clave (o que se registran como insumos adecuados), no afecta o influye en su desempeño. Por ejemplo, la información adquirida a través del olfato no influye en las tareas de codificación de color realizadas por la visión. Oler una rosa no nos hará verla más o menos roja o más o menos cercana. El aroma de un objeto no está relacionado con la luz que es reflejada desde el objeto hacia los ojos. La visión no trabaja con información olfativa y el olfato no opera con información visual. En el caso del encapsulamiento informacional podemos recordar la ilusión Müller-Lyer. El contenido de la experiencia perceptual al ver las líneas no es afectado por el estado cognitivo que puede producirse cuando las medimos y vemos que son iguales. Las creencias en este caso

²² Eraña (2006), p. 3.

²³ Fodor, (2000), p. 63.

no alteran la experiencia perceptual. Los intercambios de datos con otros sistemas están restringidos. Fodor señala:

...there are some *prima facie* reasons for doubting that the computations that input systems perform could have anything like unlimited access to high -level expectations or beliefs. These considerations suggest that even if there are some perceptual mechanisms whose operations are extensively subject to feedback, there must be others that compute the structure of a percept largely, perhaps solely, in isolation from background information. (1983, p. 66)²⁴

Son fáciles de encontrar las razones para dudar del acceso informacional irrestricto proveniente de estados cognitivos conceptuales cuando tomamos en cuenta ilusiones como la que mencione arriba o casos como los de visión ciega.²⁵ En ejemplos del último tipo, la información recogida por el sistema visual del sujeto no alcanza los niveles de la evidencia consciente pero su conducta pone en evidencia la asimilación de la información visual. A pesar de que el sujeto cree no haber visto, se comporta como si lo hubiera hecho con una alta tasa de éxitos en experimentos realizados. Parece que cierta información es tratada localmente sin influencia claramente conceptual.

Si concedemos que el uso de la PP es posible gracias a una facultad psicológica propia del ser humano (i.e. la capacidad de desplegar conceptos como percepción, creencia, deseo, ubicados dentro de una cadena de generalizaciones causales para explicar y predecir conductas) entonces podríamos pensar que es modular si tuviera dominio específico y estuviera encapsulada informacionalmente según su formulación clásica. Esto se cumpliría si la facultad psicológica operara sólo con elementos de cierta

²⁴ "Hay razones *prima facie* para dudar de que las computaciones que realiza el sistema de insumos pudieran tener algo como acceso ilimitado a expectativas o creencias de nivel superior. Estas consideraciones sugieren que incluso si hay algunos mecanismos perceptuales cuyas operaciones están sujetas extensivamente a la retroalimentación debe haber otras que computen la estructura de un percepto en su totalidad, tal vez sólo, en aislamiento de la información de fondo".

²⁵ Tye (2000).

clase y no recibiera influencia de otros procesos cognitivos. Esta es en parte la concepción que Fodor (1983) y (2000) sostiene.

Bajo esta lectura de la modularidad, estos dos requisitos no se cumplen. El requisito de dominio específico falla pues hay muchos estados mentales diferentes que participan en el despliegue de la psicología popular (deseos, creencias, dolores, emociones, etc.). Si hubiera un módulo mental entonces su dominio sería muy amplio. No sólo eso, el tipo de insumos con los cuales haría sus transformaciones además de incorporar todo tipo de estados mentales tendría que lidiar con su formulación o contraparte lingüística. Esto quiere decir que si el módulo opera con actitudes proposicionales entonces tiene que incorporar un estado de actitud y un elemento lingüístico al cual está dirigida la actitud con el resultado de quebrantar la especificidad de dominio.²⁶ Por otro lado, el requisito del encapsulamiento informacional parece no cumplirse pues los resultados de otros procesos cognitivos, como la percepción en todas sus modalidades sensoriales, influyen de manera importante en las interacciones entre los estados mentales. Incluso, en cierto grado, las salidas del sistema parecen modularse cuando se tiene conocimiento de una situación. Un ejemplo de esto puede encontrarse en la reacción de los super-espartanos que Putnam (1967) concibe en su experimento de pensamiento. Según este célebre texto, el sujeto reacciona de tal manera que no sólo modula, sino que corta la salida conductual que se produce típicamente en los episodios de dolor. Al ser la manifestación pública de dolor algo ignominioso se inhibe la conducta observable que es natural para todos los no-super-espartanos. Según lo anterior, lo que sabemos parece afectar la manera de relacionar los estados mentales con inputs y outputs. Por tanto, parece que sí hay acceso informacional o penetrabilidad cognitiva.²⁷

²⁶ Esta objeción tiene éxito siempre que consideremos que los estados mentales son de tipos diferentes. Pero si seguimos a Tye (1995a), (1995b), y (2000) y concebimos a todos los estados mentales como representacionales entonces no habría en principio problema para que la especificidad de dominio se cumpliera. Pues todos tendrían algo en común, a saber, ser intencionales, tener contenido representacional. Sin embargo, este punto sigue en pugna.

²⁷ El argumento de la penetrabilidad cognitiva será tratado en el siguiente capítulo.

La consecuencia de estos dos últimos párrafos es que la idea de que hay un módulo mental de dominio específico e informacionalmente encapsulado pierde fuerza.²⁸

A pesar de que la conclusión de los párrafos anteriores es negativa en cuanto a una noción de modularidad de lo mental, hay otras razones que favorecen no sólo la idea de un módulo, sino también de una teoría innata sobre lo mental. A continuación expondré tres argumentos favorables hacía las dos características anteriores de un sistema cognitivo sobre lo mental.

A) Los niños coinciden temporalmente en la capacidad de mentalizar. Normalmente se considera que existen tres etapas para llegar a un punto de madurez en el desarrollo de la PP. Ellas están comprendidas desde el nacimiento hasta los 4 años y cada una tiene sus rasgos particulares en los cuales se perfeccionan ciertos aspectos.

La primera etapa (comprendida entre el nacimiento y los 18 meses de edad) es conocida como “psicología de metas”. Consiste básicamente en interpretar las acciones como orientadas a alcanzar metas. El motor principal para la realización de una acción es uno o varios deseos. A partir de atribuir deseos y considerar el entorno se generan predicciones.²⁹

La segunda etapa (comprendida entre los 2 y 3 años) se llama etapa de “psicología de deseos y percepción”. La diferencia con la etapa anterior radica en que ahora el usuario puede hacer ciertos ajustes en los deseos según las variaciones en los estímulos perceptuales relevantes. El niño no ha trazado una división entre lo que él percibe y lo que los demás perciben. Generalmente considera que lo que él percibe es lo que los demás perciben y lo que no le es asequible tampoco lo es para otros.³⁰

²⁸ Las concepciones de módulo varían en la literatura cognitivista y excedería los propósitos de este trabajo abordarlas todas. Ciertamente si se cambia la noción de modularidad por una más permisiva sería posible que la mente fuera modular en ese sentido.

²⁹ Botterill y Carruthers (1999), p. 79.

³⁰ Botterill y Carruthers (1999), p.80.

La tercera etapa (normalmente alcanzada en el cuarto año de edad) es conocida como “psicología de creencias y deseos”. El niño en esta etapa es capaz de hacer diferencias claras entre sus creencias y las creencias de los demás de modo que le es posible considerar la idea de tener creencias falsas. Se crea una escisión entre apariencia y realidad. El niño está conciente de que sus representaciones del mundo podrían ser erróneas. Se constituye el ámbito subjetivo al poder concebir la posibilidad de que lo que se percibe no sea lo que es.³¹ En esta etapa se alcanza una teoría representacional de la mente.

Una explicación plausible de este desarrollo por etapas sería la postulación de una teoría básica subyacente contenida en un módulo o representada internamente de algún modo.

Otro argumento para favorecer la idea de un módulo es el siguiente:

B) Las diferencias culturales no afectan de manera importante los patrones de desarrollo.

Botterill y Carruthers (1999) apoyándose en el argumento anterior sugieren que una manera de apoyar la idea de que tenemos un módulo dedicado a la teoría de la mente es apelando a lo que ellos llaman “rigidez del desarrollo”. Ésta idea consiste en que la adquisición y despliegue de la habilidad de “leer la mente” presenta una invariabilidad a través de los distintos contextos sociales. Si poseer una teoría de la mente fuera un hecho dependiente enteramente del estímulo social, entonces no se presentaría tal constancia en el núcleo de principios y conceptos que la constituyen. Sin embargo, existe cierta uniformidad notable en los elementos que conforman la PP. Por lo tanto, tenemos buenas razones para pensar que la capacidad de mentalizar no depende completamente de los estímulos sociales.³² El punto a notar es que hay una habilidad para “leer las mentes”

³¹ Botterill y Carruthers (1999), p. 80.

³² Botterill y Carruthers (1999), pp. 78, 79.

presente en casi todos los seres humanos normales y sus etapas de desarrollo son invariables a pesar de las diferencias en los entornos socio-culturales.

Un tercer argumento para apoyar la idea de un módulo es:

C) Niños con deficiencias mentales específicas, como el síndrome de Williams, son incapaces de adquirir ciertas habilidades teóricas (como capacidades matemáticas) pero explican y predicen las conductas de los otros, además de que superan pruebas complejas como las de creencia falsa.³³ En pruebas donde se evalúa el razonamiento matemático obtienen resultados muy pobres comparados con niños que padecen de otras deficiencias mentales. En cambio, en estudios donde se precisa emplear su competencia lingüística y mental igualan en sus resultados a los niños normales. Sus habilidades lingüística y psicológica se encuentran casi intactas. En cambio, niños con autismo son hábiles en la adquisición y desarrollo de teorías matemáticas, por ejemplo, pero no desarrollan la capacidad de mentalizar adecuadamente.³⁴

Estas tres razones sugieren que hay un módulo, o por lo menos un mecanismo determinado de manera innata, que está a la base del desarrollo de la teoría psicológica popular. Dejaré hasta este punto la reflexión sobre esta manera de explicar el carácter tácito de la PP. Exploraré a continuación la idea de un acceso parcial a los principios que integran a la psicología de creencias y deseos.

Ahora bien, si la información no está completamente encapsulada entonces es posible que los principios que constituyen a la PP sean débilmente tácitos. Si lo anterior es el caso, entonces sus reglas pueden ser accesibles a la mente consciente. Botterill escribe sobre la accesibilidad: "On the widely held view that explanation involves subsuming particular cases under general principles, the prominence and precociousness of explanatory practice must be a strong argument in favor of the accessibility of folk

³³ Segal (1996).

³⁴ Baron-Cohen et. al. (1985).

psychological principles”.³⁵ Esta situación tiene un aspecto importante a notar. Si se tiene acceso por lo menos a algunos aspectos de la TT, en principio no habría razón para no intentar hacerlos explícitos. Un primer intento de cumplir con esta tarea lo ha realizado Botterill, como veremos enseguida. Otra razón que favorece la idea de que la TT es débilmente tácita es la significatividad de la explicación psicológica. En la práctica psicológica cotidiana, al observar conductas atribuimos normalmente estados psicológicos al sujeto que realiza la acción. Pretendemos que las interacciones de los estados psicológicos junto con los estímulos ambientales producen esa conducta. Esta última nos resulta inteligible cuando reconocemos su conexión con los estados antecedentes adecuados. La razón para que esto sea satisfactoriamente explicativo podría consistir en que tenemos cierto acceso a las leyes con las que opera la PP. De este modo, si la explicación que recibimos es un caso o cae bajo alguno de nuestros principios nomológicos entonces nos resulta significativa.

En la siguiente sección expondré un intento de hacer explícitos los principios que subyacen a la habilidad de mentalizar. La idea es que si la teoría que subyace a la PP es parcialmente tácita entonces sus principios pueden ser accesibles. George Botterill piensa que así es. Aquí recopiló lo que para él podría ser el núcleo de la PP:

1. Principio de Acción. Un agente actuará de tal modo para satisfacer, o al menos para incrementar la probabilidad de la satisfacción, de su más fuerte deseo actual bajo la luz de sus creencias.
2. Principio de Percepción. Cuando un agente A presencia una situación S en determinado modo, y p es un factor perceptualmente saliente de S, entonces A adquiere la creencia de que p

³⁵ Botterill (1996), p. 114. “Sobre la visión generalizadamente sostenida de que la explicación involucra subsumir casos particulares bajo principios generales, la relevancia y precocidad de la práctica explicativa debe ser un argumento fuerte a favor de la accesibilidad de los principios de la psicología popular”.

3. Principio de Inferencia. Cuando un agente A adquiere la creencia de que p y un pensador racional debería inferir q de la conjunción de p con otras creencias que A tiene, A llega a creer que q .³⁶

El Principio de Acción relaciona creencias y deseos con conductas. No especifica qué tipo de deseos deben ser instanciados y tampoco lo hace para las creencias. Lo único que señala es que dado determinado deseo y dado un conjunto de creencias, el agente realizará una acción que lo acerque a la satisfacción, según su perspectiva, del deseo en cuestión. Podría ser que la realización de esa conducta lo alejara de la satisfacción de su deseo. No obstante, el principio sólo se limita a mencionar lo que el agente en específico cree que le acercará a su meta y no a lo que en realidad le pudiera llevar a la conclusión exitosa de su acción. Este principio es tanto explicativo como predictivo. Permite explicar una conducta a la luz de la satisfacción de determinado deseo y de un conjunto de creencias. A su vez, permite predecir una acción si conozco qué deseo se quiere satisfacer y qué creencias se tienen al respecto. Así que aunque parece demasiado general funciona como un esquema central en la PP. Botterill comenta: “If the *Action Principle* seems numbingly platitudinous, so much the better. That is exactly the way we might recognize a weakly tacit basic principle”.³⁷ Muchos autores han denunciado la vaguedad e imprecisión de los principios de la PP. Sin embargo, la declaración de Botterill nos alerta sobre un punto muy importante, a saber, un principio tácito no tendría por qué ser absolutamente preciso y fino.

Los principios de la PP pueden llegar a ser empleados en la práctica psicológica cotidiana de una manera eficaz una vez que son acompañados con información adicional referente a la situación en cuestión y a sus manifestaciones típicas. Si en un caso el principio general parece no cumplirse, entonces se procede a verificar si alguna información relevante está ausente. Es decir, si resulta que al tratar de explicar la

³⁶ Botterill (1996), pp. 115, 116.

³⁷ Botterill (1996), p. 115. “Si el principio de acción parece insensiblemente trivial, mucho mejor. Exactamente esa es la manera en que podríamos reconocer un principio básico débilmente tácito”.

conducta de un agente no se puede hacerla inteligible, se tiene que añadir hipótesis adicionales con el propósito de lograrlo. Botterill sugiere: "... an agent must have had *some reason* for what she did, so if prior attributions of intentional states to the agent fail to rationalise conduct, those attributions are to be corrected or supplemented".³⁸ Sin embargo, bajo esta perspectiva, el principio central permanece intacto; lo que ocurre cuando no se logra racionalizar la conducta del agente es que se formulan hipótesis alternativas que ayuden a hacerlo.

El Principio de la Percepción relaciona estímulos ambientales con la formación de estados mentales. Una de las fuentes cruciales para la formación de creencias es la recepción de información a partir de las distintas modalidades sensoriales. De este modo si un agente percibe algo que se expresa como *p*, la tendencia natural es creer que *p* cuando el agente tiene los conceptos adecuados.

El Principio de Inferencia es normativo al establecer qué curso debe seguir un razonamiento para preservar la verdad. La manera en la que solemos predecir la conducta es suponiendo que el agente (blanco de nuestra predicción) hace inferencias como lo establece este principio. Suponemos que a partir de ciertas creencias llegará a otras creencias siguiendo un proceso determinado.

Los principios que Botterill supone subyacen a la PP son poco específicos pero su validez generalmente está protegida por medio de cláusulas *ceteris paribus*. Esto significa que hay restricciones que imponen cierto tipo de normalidad referente al contexto en el que se emplea el principio. La idea es demarcar adecuadamente cuándo se aplica un principio teórico en una situación práctica sin que una falla en su aplicación lo falsifique. Una manera de ejemplificar esto es la siguiente: Si un sujeto ve una pared roja entonces adquiere una creencia acerca del color de esa pared, a saber, cree que la pared es roja. La aplicación de la cláusula *ceteris paribus* nos indica que lo anterior ocurrirá si existe un

³⁸ Botterill (1996), p. 117. "Un agente debe haber tenido alguna razón para lo que hizo, de tal modo, si las atribuciones previas de estados intencionales del agente fallan en racionalizar la conducta, aquellas atribuciones serán corregidas o complementadas".

contexto de normalidad cuando se presenta el evento. En caso de que el sujeto esté en una habitación donde la luz es roja puede resultarle difícil en principio tener una creencia sobre el color “real” de la pared pues las condiciones de iluminación habituales no se presentan y por lo mismo no se cumple la cláusula *ceteris paribus*. De lo anterior resulta que el principio “las personas creen lo que ven” no resultaría falsado en caso de que nuestro sujeto no llegue a formarse la creencia de que la pared es roja. La razón es simple, no se presentaron las circunstancias habituales en las que ese principio se aplica. Si los elementos contextuales no forman parte del suceso en cuestión (como tener una iluminación determinada) entonces el fallo en la aplicación del principio general (las personas creen lo que ven), radica precisamente en una implementación defectuosa y no en un error del principio teórico. Esto quiere decir que las cláusulas *ceteris paribus* protegen la validez de la ley cuando se trata de instanciar en un caso que no cumple con el requisito de normalidad o de uniformidad.

Ahora bien, la PP, por un lado, nos permite explicar qué estados mentales puede tener el sujeto cuyas acciones queremos hacer inteligibles y por otro lado, nos posibilita hacer predicciones acerca de su conducta. Para explicar qué estados mentales tiene el sujeto en cuestión consideramos las acciones que está efectuando junto con los estímulos ambientales a los que está siendo expuesto y recurrimos a nuestros principios teóricos para “llenar los huecos” con estados mentales que hagan inteligible ese comportamiento. A partir de la consideración de factores perceptibles podemos acceder a los posibles estados psicológicos del sujeto en cuestión. La PP no sólo nos facilita esta tarea sino que gracias a ella tenemos todo un conjunto de conocimientos sobre como relacionar entre sí a estados psicológicos con estímulos ambientales y conducta. Uno de los rasgos más importantes de la PP es que nos permite hacer predicciones acerca de la conducta de los sujetos. Si tomamos en cuenta los insumos externos a los que está siendo expuesto un sujeto más los estados psicológicos que pudiera tener, y si recurrimos a nuestros

principios generales, entonces podemos ser capaces de anticipar que acción llevará a cabo el agente.³⁹

El trío de principios postulado por Botterill parece intuitivamente correcto. Sin embargo, da la impresión de ser un tanto esquemático e inexacto. No especifica cómo reconoce un sujeto a un agente como tal para instanciar los principios. Tampoco especifica qué criterio hay que emplear para atribuir deseos al otro una vez que se identificó como agente y menos habla de rivalidad entre deseos mutuamente excluyentes. Por otro lado, el principio de Inferencia que postula parece ser violado en muchas ocasiones por sujetos que consideramos que actúan racionalmente. Además no es claro a primera vista que el principio se respete por animales no humanos y bebés. El principio de Percepción, por su parte, suele violarse en casos de ilusión como cuando percibimos la ilusión Müller Lyer o el cubo de Necker. También se debilita frente a casos cuando no tenemos los conceptos adecuados para identificar al objeto saliente como Dretske (1968, 1981) lo menciona. Los principios expresados por Botterill parecen hablar más de un paradigma de racionalidad que de un mecanismo básico para mentalizar. Considero que aunque es interesante el conjunto de principios que Botterill ofrece, no es adecuado. Sobre todo si consideramos que niños con autismo que pueden perfectamente hacer inferencias según el *modus ponens* no dominan la PP.

Antes de terminar con esta sección quiero aclarar en qué sentido la PP puede entenderse como una teoría tácita. Esta pequeña digresión aclarará por qué no ser capaz de hacer explícitos los principios que regulan nuestras acciones no implica que no haya tales principios.

Otra manera de entender el carácter tácito de la psicología popular lo ofrece Fodor (1968). Al realizar ciertas actividades muchas personas se dan instrucciones a sí mismas.

³⁹ Botterill y Carruthers (1999).

Por ejemplo, cuando los niños aprenden a atarse las agujetas repiten en voz baja las instrucciones para hacerlo. Algo análogo sucede cuando para escribir una palabra en un idioma desconocido, se procede a deletrearlo y simultáneamente a escribirlo. Sin embargo, gran parte de las actividades que realiza el ser humano en la vida cotidiana no se consideran primero y luego se ejecutan. No es usual pensar en los movimientos que se necesitarán para tomar un vaso con agua. Normalmente pasa desapercibido para el agente el proceso que lleva a realizar una conducta. Fodor (1968), defendiendo una visión intelectualista de las competencias mentales, señala:

...not just that there are causal interactions in which the organism is unconsciously involved, but also that there are unconscious processes of learning, storing, and applying rules which in some sense “go on” within the organism and contribute to the etiology of its behavior. (Fodor 1968:632)⁴⁰

Al hablar, pocas personas son conscientes de las reglas que están empleando para construir enunciados. Esto se realiza de manera automática. Se construyen enunciados gramaticalmente bien formados sin reparar en ello. En conversaciones la demanda de respuesta es continua. Exige un dominio enorme de reglas involucradas en la articulación de frases. Este intercambio conversacional no se podría ejecutar tan rápido si se tuviera que decir en voz baja, primero, qué principio se va a emplear, y luego, la frase que se va a construir. Incluso buenos conversadores no son capaces de hacer explícitas las reglas que emplean, ni pueden, en muchas ocasiones, enseñar cómo lo hacen. Para ellos es algo que sale naturalmente. Fodor considera que para saber cómo hacer X no es necesario ser capaz de explicar cómo se hace X. Alguien puede explicar perfectamente bien cómo hacer una acrobacia en el aire al saltar del trampolín pero no ser capaz de hacerla. Por otro lado, alguien puede ser capaz de cantar absolutamente bien entonado pero no ser capaz de explicar cómo lo hizo. La capacidad de hacer X no está ligada necesariamente a la

⁴⁰ “... no sólo que hay interrelaciones causales en las que el organismo está involucrado inconscientemente, sino además que hay procesos inconscientes de aprendizaje, almacenamiento, y aplicación de reglas que en algún sentido “funcionan” dentro del organismo y contribuyen a la etiología de su conducta”.

capacidad de explicar como se hace X. Pueden coincidir. Tal vez eso debería ser lo ideal para cualquier actividad. Sin embargo, no ocurre de ese modo. En todo momento las personas hacen cosas que no pueden explicar como las hacen.

El argumento de Fodor es el siguiente: Si X es algo que un organismo sabe cómo hacer pero es incapaz de explicar cómo hacerlo, y si S es alguna sucesión de operaciones, cuya especificación constituiría una respuesta a la pregunta “¿cómo haces X?”, y si una repetición óptima de la conducta X del organismo se realiza corriendo la sucesión de operaciones especificadas por S, entonces el organismo sabe tácitamente la respuesta a la pregunta “¿Cómo haces X?”, y S es una formulación del conocimiento tácito de ese organismo.⁴¹

Para realizar una acción un sujeto emplea ciertas reglas (como en el caso de la formación de enunciados gramaticalmente correctas) pero es incapaz de hacerlas explícitas. No puede articularlas verbalmente pero las emplea. Si lo anterior es el caso, entonces las posee tácitamente como integradoras de su conducta. Para decirlo de otro modo, si hay una secuencia de operaciones que se realizan para producir una conducta, y son efectuadas sin que el mismo individuo pueda reportarlo, entonces tal secuencia es conocida tácitamente por él.

Tal vez el hombre ordinario no pueda encontrar las leyes de una u otra teoría. No obstante, para el teórico la posibilidad de articular los principios de estas teorías está latente. No cualquier persona puede hacer explícitos los principios que integran su conducta pero eso no implica que no los posea siempre y cuando se conciben como tácitos.

En esta sección he expuesto una caracterización más de la PP. Si entendemos a los módulos como cuerpos de conocimiento con reglas internas de manipulación de información conectados con el exterior gracias a los insumos que reciben y los resultados que producen y con una ontogenia específica e invariable ante la pluralidad de contextos (altamente canalizada) entonces podría ser que la PP esté contenida en un módulo. Según

⁴¹ Fodor (1968), p. 638.

dos rasgos de la caracterización fodoriana de módulo, especificidad de dominio y encapsulamiento informacional, la PP no podría estar en un módulo pues incumple ambos rasgos. Después expuse tres argumentos, que a pesar del pesimismo de Fodor, podrían hablar a favor de un sistema genéticamente determinado que apoya la habilidad de mentalizar. Más adelante, enuncié algunos principios que Botterill considera que constituyen el núcleo de la PP. Finalmente expuse un argumento de Fodor a favor del carácter tácito de ciertos tipos de conocimientos mediante la separación entre saber hacer X y explicar como se hace X.

¿Cuál es el resultado, o resultados, de esta sección? Primero, es posible entender a la PP como una teoría sin tener que comprometernos con la analogía de la teoría científica expuesta en la sección anterior. Segundo, bajo la concepción fodoriana de lo que es un módulo la PP parece resistirse a ser caracterizada como tal. Tercero, la habilidad de atribuir estados mentales y de percibir a los objetos como agentes intencionales se presenta en etapas claras y su ausencia parece estar involucrada con aspectos genéticos (Autismo). Estos hechos sugieren que el desarrollo de la PP podría estar relacionado con un sistema innato. Cuarto, aunque hay intentos de formular los principios de la PP la tarea no es tan sencilla. Quinto, es distinto saber hacer X de saber explicar cómo se hace X.

En la siguiente sección voy a exponer una manera alternativa de entender a la PP como una teoría científica. Esta postura está relacionada con la concepción de Botterill (1996) pero parte de otros rasgos para hacer la analogía.

3. La PP como una Teoría Científica (Segunda Parte)

Las psicólogas del desarrollo Alison Gopnik (1996), Susan Carey y Elizabeth Spelke (1996), llevaron la analogía entre la actividad científica y el desarrollo psicológico del infante a otro nivel precisando las relaciones entre los dos campos. Su propuesta tiene muchos puntos de contacto con la postura sostenida por Botterill (1996) pero también se distingue en aspectos importantes. En esta sección caracterizaré esta versión de la TT con el propósito de averiguar si es adecuada para dar cuenta de la PP. La postura básica que sostienen es que tanto la capacidad de hacer ciencia como la capacidad de mentalizar tienen en su base un elemento común constituido por un conjunto de reglas y representaciones.⁴²

Expondré esta versión de la TT tomando a Gopnik (1996) como paradigma. En “The Scientist as Child”, la autora sostiene que hay semejanzas estructurales relevantes entre el desarrollo cognitivo del infante y el cambio científico. Al desarrollar nuevo conocimiento, los científicos utilizan los mismos recursos cognitivos que subyacen a la maduración psicológica en los niños. En el caso de los científicos, el resultado de sus investigaciones se integra en una teoría o incluso constituye una nueva teoría. Si la analogía ha de sostenerse es preciso que el infante también posea algo equivalente al cuerpo teórico que resulta del trabajo científico. La concepción de Gopnik no sólo implica esto, sino que además afirma que los científicos son como niños en el sentido de que sus prácticas se asemejan a los primeros pasos en el desarrollo mental que el infante recorre para hacerse de una visión coherente del mundo que le permita explicar los fenómenos y anticiparlos y que para ambas actividades se emplean los mismos recursos cognitivos.

Las semejanzas entre la ciencia y los procesos cognitivos son muchas. Veremos algunas a continuación. La ciencia y la cognición buscan llegar a concepciones verídicas del mundo. Los seres humanos tenemos ciertos mecanismos producto de la evolución que nos permiten conocer de qué modo son las cosas. Básicamente, los mecanismos son

⁴² Entiendo “mentalizar” como la capacidad de atribuir estados mentales a otros seres y a uno mismo, de modo no necesariamente explícito.

concebidos en términos de reglas y representaciones. Estas estructuras son susceptibles de transformación y se entienden según un modelo computacional. Del mismo modo, podemos pensar en la ciencia como un sistema de reglas y representaciones. Este sistema es la herramienta que nos permite acceder a la verdadera naturaleza de los fenómenos con el deseable resultado de un cambio conceptual.

Para sostener la analogía entre ciencia y cognición, Gopnik (1996) expone 3 tipos de rasgos de las teorías: estructurales, funcionales y dinámicos.⁴³ Entre los rasgos estructurales de las teorías se destacan: que son sistemas de entidades abstractas relacionadas por leyes que soportan contrafácticos no enunciadas en un vocabulario observacional. Las entidades postuladas están estrechamente relacionadas una con otra en virtud de una estructura causal profunda que las hace responsables de la evidencia empírica. Las teorías también implican compromisos ontológicos pues dicen qué cosas son las que hay.

Los rasgos funcionales de las teorías incluyen la posibilidad de dar explicaciones, hacer predicciones e interpretar los fenómenos. En este punto Gopnik converge con Botterill (1996) al hacer notar la relevancia de la capacidad predictiva y explicativa de una teoría. Gopnik también aclara que la interpretación de la evidencia es un rasgo crucial de las teorías pues nos permite concentrarnos en los fenómenos que son importantes y dejar a un lado los que no lo son al formular y reconocer criterios y patrones de relevancia.⁴⁴

Los rasgos dinámicos de las teorías son los procesos involucrados en su formación y transformación. El cambio de los elementos de una ciencia puede ser motivado por evidencia en contra de la teoría que en principio es ignorada o no tomada en cuenta. Más tarde, y si persisten los casos difíciles, es habitual elaborar hipótesis auxiliares que los traten de explicar en el vocabulario de la teoría actual. Cuando las excepciones se presentan continuamente y las hipótesis auxiliares no son suficientes se suele crear un modelo alternativo que de cuenta de los fenómenos. Se inicia después un período de

⁴³ Gopnik (1996), pp. 496 y 499.

⁴⁴ Gopnik (1996), p. 497.

experimentación intensa que si es acompañado por falsaciones explícitas de la teoría desemboca en un cambio.

La noción de Gopnik sobre la TT concibe a las teorías como sistemas que asignan representaciones a insumos. Los insumos de la teoría están constituidos por información sensorial que es traducida a conocimiento teórico de nivel superior. Las representaciones son procesadas según ciertas reglas que producen nuevas representaciones. De este modo, podemos entender cómo ocurre que la conducta observada, entendida como un insumo transformado a otro vocabulario superior, se convierte en una representación que se integra a una cadena de procesamiento que deriva en el reconocimiento de la instancia como un ejemplar de cierto tipo para el sistema.

Ahora bien, ¿cómo se conecta esto con la analogía entre ciencia y cognición? Gopnik expone los resultados obtenidos en una serie de estudios realizados con niños menores de diez años (especialmente de cuatro y cinco años) en los cuales el infante ofrece explicaciones y predicciones de la conducta de otros individuos que son sometidos a ciertas situaciones que involucran creencias falsas. El niño justifica sus afirmaciones postulando eventos causales subyacentes en términos de entidades mentales que se relacionan con determinado comportamiento. Además interpreta la conducta de las personas en términos intencionales de manera natural. Cuando se le pregunta donde buscará un agente cierto objeto, predice el lugar adecuado y explica su preferencia según las creencias y percepciones que el otro pudo haber tenido. En este sentido la comprensión de la mente exhibida por el infante ostenta los mismos rasgos funcionales de una teoría, a saber, predicción, explicación, e interpretación.

El caso de los rasgos estructurales de la teoría es similar pues el infante postula entidades abstractas que considera existentes, a saber, creencias y deseos que relaciona causalmente con conductas. Sólo falta entonces que el conocimiento mentalista infantil cumpla con los mismos rasgos dinámicos que presentan las teorías científicas.

Aunque los niños menores de 4 años también hacen atribuciones de estados mentales para explicar y predecir conductas, su alcance es bastante limitado. Tienen dificultades en enfrentar casos en los que se involucran creencias falsas. Parece que sólo son efectivos

cuando se trata de predecir percepciones y deseos pero no sucede lo mismo con creencias. Gopnik señala: “They predict accurately, for example, that another person on the other side of a screen will not be able to see what they see themselves”.⁴⁵ También predicen que alguien experimentará una emoción o que sentirá un deseo después de que les sucede algo. Si le quitan un juguete a un niño pueden predecir que se sentirá triste y que deseará que se lo devuelvan y cuando esto ocurra pueden anticipar su alegría.

El dominio de sólo un grupo de estados mentales por los niños menores de 4 años nos habla de que existe algo como una teoría implícita sobre la mente que todavía no está acabada pues, según Gopnik, carece de un componente conceptual crucial para entender a los demás: el concepto de creencia falsa. El niño falla sistemáticamente en la predicción de conductas en las que el agente no tuvo acceso a la evidencia y en la atribución a sí mismo de creencias que resultan falsas. Un caso claro puede verse en los estudios de Baron-Cohen et al. (1985) en el que dos muñecas (Sally y Anne) presencian la colocación de una canica en una canasta. Posteriormente la canica es trasladada a una caja cuando una de ellas sale de la habitación. El resultado es que los niños menores de 4 años y los autistas que presenciaron la escena fallan en responder a la pregunta ¿Dónde buscará Sally la canica? No logran distinguir entre sus estados epistémicos y los de los otros. Más intrigante aún resulta el hecho de que luego de haber presenciado donde efectivamente Sally busca la canica ellos parecen olvidar que habían señalado otro lugar como destino de la búsqueda.⁴⁶ Gopnik sugiere que esto se debe a que los niños pequeños aún no son capaces de hacer distinciones entre conocimiento por percepción y conocimiento por inferencia.⁴⁷ Los niños pueden enfrentar casos donde hay creencias verdaderas pero no donde hay creencias falsas involucradas.

Lo anterior lleva a Gopnik a pensar que los rasgos estructurales de la teoría de la mente del infante son como las de las teorías científicas. Las razones son las siguientes: se inicia con una teoría en la que la percepción es siempre verídica. El infante se enfrenta con

⁴⁵ Gopnik (1996), p. 506. “Por ejemplo, ellos predicen con precisión que otra persona en el otro lado de la pantalla no será capaz de ver lo que ellos ven”.

⁴⁶ Baron-Cohen, et al. (1985).

⁴⁷ Gopnik (1996), p. 507.

casos que invalidan sus predicciones pues involucran el concepto de creencia falsa. Se desencadena un proceso de intensa experimentación en el que el niño concibe la necesidad de una modificación conceptual para salvar los fenómenos. Obviamente el niño no está consciente de esto, la fenomenología de este cambio no es clara. Cuando los errores de predicción y explicación son lo suficientemente salientes el niño cambia el concepto de creencia para permitir que haya casos donde pueda haber creencias falsas. Finalmente el niño termina con una teoría diferente a la anterior que le habilita para concebir estados con contenido erróneo. De este modo, logra entender que los otros y él mismo no siempre tienen creencias verdaderas pues hay diferencias entre los estados epistémicos de las personas.

En resumen, Gopnik concibe a la PP como una teoría en virtud de las notables similitudes que tiene con las teorías científicas en tres tipos de rasgos: estructurales, funcionales, y dinámicos. En primer lugar, hay similitudes estructurales pues ambas teorías son sistemas de entidades abstractas y de leyes que las conectan las cuales soportan formulaciones contrafácticas. Tienen una estructura profunda y hacen compromisos ontológicos con las entidades que postulan. En segundo lugar, hay similitudes funcionales pues ambas permiten hacer predicciones, explicaciones, e interpretar los fenómenos de su dominio. En último lugar, la dinámica de ambas supone etapas claras y distintas de cambio donde gradualmente se abandona la teoría anterior para pasar a otra o completar la actual con algún concepto crucial.

Hasta aquí la exposición de la teoría de Gopnik. Considero que hay ciertos comentarios que hacer al respecto de esta tercera versión de la PP como una teoría. Es probable que tenga que enfrentarse a las mismas objeciones arriba dirigidas a Botterill (1996).

La primera observación a la analogía trazada por Gopnik tiene que ver con los aspectos estructurales que ella menciona. La razón de que nos sorprenda cuando una predicción nuestra no se cumple en el ámbito psicológico, es señal de que estamos

comprometidos con la verdad de la teoría.⁴⁸ Sin embargo, la reacción de sorpresa no es suficiente para indicar que nuestros contrafácticos se sostienen. Gopnik es bastante frugal en este punto pues no argumenta demasiado al respecto. Considero que las críticas hechas hacia Botterill (1996) en la primera parte de esta investigación se aplican de igual modo al texto de la autora. Se supone que con Hume aprendimos que las relaciones causales se dan cuando un objeto es seguido por otro y cuando objetos de la primera clase siempre acompañan a objetos de la segunda clase. El problema es que esto no es respetado por las supuestas leyes de la PP. Creencias del mismo tipo, ya no hablemos del mismo contenido, acompañadas de los mismos deseos no siempre causan las mismas acciones en nosotros mismos y menos en los demás. La confianza que depositamos en una teoría psicológica de sentido común nunca es total. El problema que enfrentan las “leyes” de la PP es que tienen que protegerse con cláusulas *ceteris paribus*, pero esto hace más difícil poder dar sentido a la idea de que soportan contrafácticos. Si recordamos la segunda formulación de la causalidad de Hume, a saber, que si el primer objeto no se hubiera presentado, el segundo no hubiera existido, nos damos cuenta de que la satisfacción de este requisito para las “leyes” de la PP se complica mucho más. Tenemos que conformarnos con relaciones causales singulares entre estados mentales, insumos, y conductas en las que no conocemos la ley general.⁴⁹ Por tanto, considero que uno de los elementos que Gopnik supone ambas teorías poseen no es del mismo tipo. Esto quiere decir que su analogía se ve debilitada.

La siguiente observación es acerca de las propiedades dinámicas de las teorías según la visión de Gopnik. No arribamos a nuevas teorías científicas siempre del mismo modo y según las mismas etapas. De hecho es difícil trazar claramente las etapas en las que se encuentran nuestras teorías científicas actuales. ¿Estamos a punto de abandonar la mecánica cuántica? Creo que no lo sabemos en este momento. Parece que no, pero es arriesgado comprometerse. Para gente como Churchland (1988) la PP es una teoría

⁴⁸ Gopnik (1996), p. 496.

⁴⁹ Davidson (1963). Hay que aclarar que Davidson habla de sucesos y no de estados.

estática que va a ser remplazada en algún momento por otra teoría. ¿Acaso sabemos esto? ¿Podemos decir en qué etapa está la PP con exactitud? Creo que para este caso la respuesta es negativa nuevamente. El mismo Churchland reconoce que no podemos mudarnos de casa hasta que tengamos una nueva y todavía no la tenemos.⁵⁰ Gente como Fodor (1987) es pesimista sobre esa posible mudanza. En algunos casos sí se puede detectar el malestar de una teoría pero no en todos. La diferencia con la cognición infantil es relevante. Es mucho más sencillo determinar en que etapa de desarrollo psicológico está un infante. Lo hacemos con estudios como los de Baron-Cohen et al. (1985) o como los de Wimmer y Perner (1983). Creo que, en este caso, la analogía de Gopnik tampoco es demasiado precisa. Presenta una visión idealizada de lo que es una teoría científica y como tal no corresponde a la diversidad de maneras en que puede llegarse a una nueva teoría. Se puede pensar que mi argumento es meramente epistemológico y que se apoya en nuestro desconocimiento de ciertos hechos y que cuando conozcamos más de las teorías científicas y de los procesos cognitivos del ser humano será sencillo trazar el puente entre un campo y otro porque las semejanzas estructurales son decisivas. Creo que esto no es el caso. Aunque conozcamos el proceso detallado de cómo pasar de una teoría científica a otra (si es que hay un proceso único) y las etapas típicas (si es que las hay) eso no podrá equipararse con completar una teoría de la mente con un elemento más como lo es el concepto de creencia falsa. En un caso tenemos diferentes caminos y en el segundo caso un solo camino.

La analogía trazada por Gopnik (1996), aunque bastante evocativa, no logra establecer relaciones constitutivas entre las teorías científicas y la teoría de la mente. En primera lugar, por las diferencias en el tipo de relaciones causales que se dan entre sus entidades postuladas, y en segundo lugar, por la arriesgada apuesta por un isomorfismo en las etapas de desarrollo de ambas teorías que no se ve claramente. Esta tercera manera de ver a la PP como una teoría científica, por tanto, no parece ser definitiva ni totalmente satisfactoria.

⁵⁰ Churchland (1988).

En la siguiente sección hablaré de un último sentido según el cual se puede concebir que la PP implique una teoría. La versión que expondré puede verse como una teoría externa en el sentido mencionado por Ravenscroft (2004).

4. La PP como una Teoría Funcionalista

Otra manera de entender a la PP es verla como una teoría que introduce términos mentalistas al vocabulario cotidiano. Esta es la visión externalista de la TT. La PP bajo este enfoque es una teoría funcionalista pues individúa sus conceptos a partir del papel causal que juegan en ella.

Los términos empleados para referirse a estados y procesos mentales obtienen su significado a partir de las conexiones con otros términos de la misma teoría. Es claro que las definiciones de los términos psicológicos no son explícitas. Normalmente la PP revela el carácter de sus términos mostrando sus relaciones causales. Si se muestra el papel causal que juega un término en una teoría entonces se da una definición funcional de él.

En la vida cotidiana este tipo de definiciones es común. Podemos decir que el café es lo que nos ayuda a despertar totalmente en la mañana, que una ventana es lo que permite la entrada de luz a nuestras casas y nos protege de la lluvia, que un calefactor es lo que nos quita el frío, o que el corazón es lo que bombea sangre al cuerpo. No se recurre a una definición tradicional del tipo género y diferencia específica. Lo que se hace es mencionar el papel causal que el referente del término tiene en el conjunto de conocimientos que tenemos sobre determinado dominio.

En el caso de los términos psicológicos la situación es similar. “Creencia” puede entenderse como refiriéndose a cierto estado que junto con ciertos deseos y percepciones motiva una acción. Cuando hacemos este tipo de definiciones funcionales empleamos otros términos psicológicos para individuar al tipo de estado que tenemos en mente junto a sus relaciones causales.

Los elementos básicos de la PP son predicados mentales, predicados acerca de estímulos ambientales, y predicados acerca de conductas. Los predicados mentales incluyen a todos los verbos psicológicos. Algunos atribuyen actitudes proposicionales pero otros no. Un ejemplo de los primeros. Normalmente se emplea ese verbo en oraciones

como “Juan cree que el piso está mojado”. En este caso un sujeto posee una creencia con determinado contenido proposicional (el piso está mojado). Un ejemplo de los segundos (de los verbos psicológicos que no atribuyen actitudes proposicionales) sería “doler”. Habitualmente tal verbo se utiliza en oraciones como “tengo un dolor en el brazo”. En este último caso es posible que el afectado por el dolor pueda creer que tiene el dolor en el brazo. Sin embargo, poder verbalizar tal situación (expresando una oración declarativa) no implica que se tenga necesariamente una actitud hacia cierta proposición. El sujeto en cuestión podría no haber adquirido jamás el lenguaje y sin embargo tener dolor.

Los predicados acerca de estímulos ambientales son los verbos que incluyen a todas las modalidades sensoriales como, ver, oler, saborear, sentir, oír. Atribuirle a alguien un predicado de este tipo se realiza, por ejemplo, cuando digo “Luis vio una luz” En el caso de la primera persona ocurre lo mismo pues emito oraciones acerca de mis percepciones al decir “oí pisadas en el sótano” o “percibí un aroma agradable” y de este modo me atribuyo ciertas experiencias perceptuales.

Por último, hay predicados acerca de conductas. La atribución de estos predicados describe los aspectos relativos a las acciones de las entidades consideradas. Oraciones con predicados de conducta serían “subió rápidamente las escaleras”, “habló toda la noche”, o “estoy escribiendo”. Sin embargo, en la TT normalmente se habla de conducta intencional y no cualquier tipo de conducta. Bajo este concepto no cabría hablar de la conducta de las rocas, o de la conducta de los átomos. Como Davidson bien señala “... los sucesos dejan de ser acciones o conducta sólo cuando no hay manera de describirlos en términos de intención”.⁵¹ Los predicados de conducta que la TT considera relevantes para dar la explicación psicológica son precisamente los que están relacionados con conceptos como intención, creencia, deseo, etc.

Ahora bien, como sugiere Lewis (1972) podemos individuar un término mental si encontramos cuál es el papel causal que ocupa en la teoría. Para esto hay que unir todos

⁵¹ Davidson (1971), p. 289.

los lugares comunes acerca de las relaciones causales, conocidas por todos, de los estados mentales, estímulos sensoriales, y respuestas motoras. Como resultado, se habrá constituido una teoría que le da significado a los términos que emplea mediante su rol causal. A todos los términos mentalistas se les puede llamar términos T. A los términos de estímulos y respuestas términos O. Si sé que ciertos estados mentales ocupan los papeles causales que la teoría psicológica afirma entonces sé qué son. Del mismo modo, si sé que un órgano bombea sangre por todo el cuerpo a través de venas y arterias entonces sé que es el corazón. El rol causal que juega un término en una teoría verdadera nos indica qué cosa se refiere.

La PP, bajo esta perspectiva, es una teoría que introduce términos psicológicos al vocabulario del hombre. Al aprender la PP, aprendo qué son los estados mentales pues adquiero un grupo de generalizaciones que relacionan términos T entre sí y con términos O.

He señalado que la PP puede entenderse como una teoría que introduce sus términos funcionalmente. El significado de un término psicológico como dolor no se da a partir de una definición tradicional. Sucede que al adquirir o desarrollar las leyes o principios de la PP se obtiene a su vez el criterio para individuar los referentes de sus términos. Este criterio consiste en que si las relaciones causales de dos sucesos mentales son iguales entonces se les puede identificar como del mismo tipo. Este enfoque aunque plausible no explica cuestiones como la coincidencia de desarrollo o la invariabilidad transcultural. Por otro lado, tampoco nos da cuenta de los problemas que enfrentan los autistas ante la atribución, por ello, parece ser iluminativo pero no definitivo ni completo.

En este capítulo he explicado 4 sentidos en los que se puede entender a la TT. En conclusión, las distintas versiones de la TT no han dejado satisfechos a todos los filósofos que han reflexionado sobre la PP. Muchos creen que no es la mejor apuesta pues hay maneras de explicar la competencia psicológica sin recurrir a un cuerpo teórico

internamente representado que a la vez permiten dar cuenta de todos los detalles de esta habilidad. En el siguiente capítulo expondré dos puntos de vista de cómo podría explicarse la PP sin recurrir a una teoría. Los dos autores que revisaré tienen en común el rechazo a la postulación de una teoría psicológica para predecir y explicar la conducta intencional de un agente. El grado de fuerza con el que sostienen sus tesis es lo que parcialmente los separa. En el último capítulo de esta investigación mostraré qué aspectos sería conveniente rescatar de estas 4 versiones de la TT.

Capítulo II

La Teoría de la Simulación

Introducción

Una Propuesta Alternativa

Para dar cuenta de la PP, varios filósofos proporcionaron un mecanismo alternativo a la concepción teoricista tradicional basado en la idea intuitiva de “ponerse en los zapatos de otro”. Esta postura es diferente a la visión establecida pues no se apoya en una teoría. Para dispensar ese recurso un individuo tiene que realizar un razonamiento práctico muy particular. Idealmente un razonamiento práctico da como conclusión una acción (o una intención) a partir de creencias y deseos como premisas. En este uso novel, la acción que se produciría es inhibida evitando activar los sistemas motores. Lo que resulta es una intención de acción o una anticipación. Pero no sólo el resultado del razonamiento práctico es diferente del habitual, también los insumos que tiene que recibir el sistema. Naturalmente, cuando uno se pone en los zapatos del otro es recomendable modificar el tipo de estímulos que se están recibiendo. Para realizar esto es conveniente imaginar que estamos en la situación ajena, con lo que esto implica, i. e. tener que imaginarse como recibiendo cierta información ambiental desde otro lugar, como teniendo ciertas creencias diferentes producto de esto, y como teniendo otros deseos como consecuencia de lo mismo. Esto puede parecer complicado, no obstante, lo hacemos en todo momento cuando consideramos posibles cursos de acción. Al realizar planes para acciones futuras efectuamos un razonamiento práctico-hipotético. Pensamos qué tendríamos que hacer si ciertas condiciones estuvieran presentes y concluimos con un comportamiento posible. Las condiciones que tomamos en consideración podrían ser las que de hecho alguien esté enfrentando. De este modo, si queremos predecir la conducta de un agente, pensamos en las condiciones a las que se enfrenta el blanco de nuestra predicción y realizamos la

simulación. El resultado de este ejercicio es lo que atribuimos como conducta futura para el agente en cuestión. El usuario de esta estrategia (que para los simulacionistas somos todos) no emplea una teoría de la mente para coordinarse en sociedad, sino que le basta con los frugales recursos que esta concepción presenta. Por tanto, se presume como más económica al no tener que postular un cuerpo de conocimientos de dominio específico.

Ahora veremos con más detalle las posturas de Gordon y Goldman en ese orden.

1. Teoría de la Simulación: Gordon (1986)

Robert Gordon fue uno de los primeros filósofos en desafiar a la influyente idea de que la PP es una teoría en su artículo “Folk Psychology as Simulation”.⁵² El estatus teórico de la PP había sido fuertemente criticado por Churchland (1988) y este nuevo enfoque le daba frescura al debate sin saber que las nociones precedentes sobre la capacidad de mentalizar cambiarían definitivamente afinándose y detallándose.

Veamos en qué consiste su propuesta. Para la estrategia simulacionista es esencial la noción de razonamiento práctico. El razonamiento práctico habitual se realiza si a partir de ciertas premisas que incluyen creencias y deseos se concluye con una acción. Por ejemplo, si creo que tengo dos pesos en el bolsillo y deseo comprar un boleto del metro, entonces la acción resultante será que compre un boleto del metro. En el caso anterior ocurren varias cosas: creo que tengo dos pesos en el bolsillo. También ocurre que deseo comprar un boleto del metro. Tener estas dos actitudes proposicionales me lleva normalmente a la realización de una acción que en este caso es la de comprar el boleto.

No obstante, puedo pensar: “¿qué pasaría si quisiera comprar un boleto del metro pero no tuviera dinero?” Es decir, que un aspecto del mundo actual fuera diferente. Ante una situación como ésta se suele concluir con una hipótesis sobre lo que se haría en tal caso. Alguien podría responder “tal vez pediría dinero” o “pediría un boleto del metro a un

⁵² Gordon (1986).

usuario". Este tipo de razonamiento es el que Gordon ubica a la base de su estrategia, y por lo mismo, a la base de la PP. La idea de la modalidad simulada tiene que ver con emplear la estrategia de "fingir" estar en otra situación o "pretender" que algún aspecto de lo que se considera verdadero y real se presenta de forma distinta o no se presenta. Se puede pensar que el mecanismo de razonamiento práctico funciona como un sistema que dados ciertos insumos produce ciertos resultados. En una instancia típica de razonamiento práctico los insumos son percepciones, creencias y deseos genuinos y el resultado es una acción. En el caso de que el razonamiento práctico se haga funcionar en su modalidad contrafáctica, por decirlo de algún modo, los datos de entrada son actitudes proposicionales que no posee el simulador y el resultado es una hipótesis de acción.

Dicho esto, ¿qué tipo de relación puede establecerse entre el manejo del razonamiento práctico y las predicciones de la conducta propia y de los otros? Es bastante intuitivo aceptar que empleamos este mecanismo de razonamiento para actuar con un fin. Por ejemplo, si tengo sed, entonces el mecanismo de razonamiento práctico me señala el camino para satisfacer esa necesidad. Para saber que hacer si tengo sed, veo en qué situación estoy. El tipo de acción que vaya a ejecutar cambiará según lo que crea. Si creo que el agua está envenenada, el resultado del razonamiento práctico podría ser diferente que si parto de la creencia que el agua es pura. El fin que persigo puede hacerse público mediante una declaración de lo que pretendo hacer. Puedo decir, "voy a tomar agua" y enseguida hacerlo. Según Gordon, las declaraciones de tienden el puente entre acción el razonamiento práctico y las predicciones si se emplea la estrategia con insumos hipotéticos.⁵³ Gordon señala: "it is easy to see how, by simulating the appropriate practical reasoning, we can extend our capacity for self-prediction in a way that would enable us to predict *our own behaviour in hypothetical situations*".⁵⁴ Esto es, no sólo puedo saber qué hacer si tengo sed sino que puedo concebir qué tendría que hacer si tuviera sed a pesar de que en la situación actual no tenga sed. El razonamiento práctico simulado nos permite

⁵³ Gordon (1986), p. 62.

⁵⁴ Gordon (1986), p. 62. "Es fácil ver cómo, al simular el razonamiento práctico apropiado, podemos extender nuestra capacidad de autopredicción de una manera que nos permitiría predecir nuestra propia conducta en situaciones hipotéticas".

considerar un posible camino de acción ante un posible estado de cosas diferente del actual. La mayoría de las personas responde fácilmente cuando se les hace la pregunta: “¿qué harías si P?” donde “P” es un estado de cosas diferente al actual. En la vida cotidiana evaluamos escenarios alternativos para actuar y predecimos qué haríamos en esas situaciones posibles.

Uno de los rasgos más importantes de la PP es que sirve para coordinarnos socialmente y un elemento clave de esta coordinación es la predicción de las conductas de los otros. La PP nos permite esperar que se den ciertas conductas o ciertas configuraciones mentales en los demás; en virtud de eso actuamos. Para dar el paso de predecir la conducta propia a predecir la conducta de los otros hay que retomar la idea de ‘fingir’ o ‘pretender’ estar en cierto escenario alternativo. Considero ciertas creencias hipotéticas que pudiera poseer el sujeto en cuestión y echo a andar el mecanismo de razonamiento práctico. Ocurre que al recibir insumos hipotéticos se desacopla de sus salidas motoras habituales. La hipótesis producto de este razonamiento práctico simulado se atribuye al sujeto en cuyos zapatos me puse. En otras palabras, el escenario alternativo con el que alimento al mecanismo de razonamiento práctico producirá una conducta hipotética que servirá de predicción. Una ilustración de esto puede hallarse en la siguiente situación: supongamos que vemos a un sujeto que está formado en la fila de la taquilla del metro y tiene en su mano una moneda de dos pesos. Si quiero predecir su conducta tengo que cambiar ciertas cosas, como mis deseos (tal vez yo no quiero un boleto ahora), y mis creencias (imaginarme qué estoy formado en la fila de la taquilla) para que se ajusten a las del blanco de mi predicción. Una vez que logré captar el escenario en el que está el individuo, alimento con esos estados hipotéticos mi mecanismo de razonamiento práctico. El resultado de este proceso lo atribuyo al individuo que estoy considerando. “Va a comprar un boleto”, predigo con seguridad.

La Teoría de la Simulación en la versión de Gordon (1986) supone que el sistema de razonamiento práctico puede ser alimentado con dos tipos de datos. En primer lugar, reconoce estados mentales y escenarios genuinos. Segundo, estados mentales y escenarios hipotéticos o contrafácticos. También produce a su vez dos tipos de resultados: motrices e hipotéticos. Las salidas motrices son las acciones o conductas que se efectúan cuando el razonamiento es alimentado con insumos genuinos. Las salidas hipotéticas, por su parte, son causadas por entradas hipotéticas.

Tenemos entonces un mecanismo de razonamiento práctico a la base de la simulación que se desacopla de sus entradas naturales para aceptar entradas hipotéticas y originar salidas tentativas que se atribuyen al sujeto cuyo escenario construimos en nuestra imaginación. Gordon expone el mecanismo de simulación claramente en el siguiente pasaje:

... Our decision-making or practical reasoning system gets partially disengaged from its 'natural' inputs and fed instead with suppositions and images (or their 'subpersonal' or 'subdoxastic' counterparts). Given these artificial pretend inputs the system then 'makes up its mind' what to do. Since the system is being run off-line, as it were, disengaged also from its natural output systems, its 'decision' isn't actually executed but rather ends up as an anticipation, perhaps just an unconscious motor anticipation, of the others behavior.⁵⁵

Algo importante a notar en lo que expone Gordon es que el resultado del proceso de simulación no proporciona una salida conductual sino una anticipación motriz. La manera de expresarla para atribuirla a los otros no queda especificada, si tomamos en cuenta que en principio es subpersonal habría que mostrar un mecanismo para traerla al plano consciente. El resultado de la simulación *off-line* en este pasaje parece ser una

⁵⁵ Gordon, (1986), p.70. "... Nuestro sistema de toma de decisiones o de razonamiento práctico, se desacopla parcialmente de sus entradas "naturales" y es alimentado en su lugar con suposiciones e imágenes (o sus contrapartes subpersonales o subdoxásticas). Dadas estas entradas artificiales simuladas el sistema decide qué hacer. Dado que el sistema está corriendo "fuera de línea", como si estuviera, desacoplado de sus sistemas de salida natural, su decisión no es ejecutada en realidad, sino que en su lugar termina como una anticipación, tal vez una anticipación motora inconsciente, de la conducta de los otros".

proposición o un estado representacional acerca de un estado de cosas. Los pasos para realizar la simulación son los siguientes:

Previo a la simulación:

1. Detectar a un objeto como agente intencional. Animales no humanos y otros seres también son susceptibles de ser vistos de este modo. (Este punto puede ocasionalmente ir después del número 2).
2. Ver sus movimientos como conducta. A partir de la percepción de movimiento autónomo atribuir intencionalidad. Si el movimiento no es autónomo o es accidental entonces se detiene la simulación y se pueden buscar causas no mentales. (Parece que los autistas fallan desde este punto).
3. Intención de comprender, predecir, o explicar las acciones. Deseo o necesidad de dar sentido al comportamiento. Si esto no es el caso entonces no inicia la simulación.

Inicio de la simulación:

4. Cambiar imaginativamente perspectivas espacio-temporales con el agente en cuestión.
5. Cambiar papeles institucionales.
6. Racionalizar la conducta del agente buscando y alterando sus posibles creencias y deseos preservando coherencia.
7. Determinar aspectos de origen que pudieran dar sentido a la conducta del agente.
8. Crear hipótesis alternativas sobre su conducta.
9. Eliminar hipótesis descabelladas si están demasiado alejadas de las tendencias naturales del observador.
10. Probar las hipótesis buscando coherencia.
11. En caso de encontrar un par de hipótesis con adecuación empírica, juzgar nuevamente según la propia tendencia y repetir el paso 10.

12. Al hallar una hipótesis correcta, crear un mundo simulado que corresponda al agente y que tenga adecuación empírica.⁵⁶

Fin de la simulación.

Gordon no dice qué pasaría si encontramos al blanco de nuestra predicción en situaciones futuras o cuando vemos a otro individuo con el mismo comportamiento del agente. ¿Acaso tendríamos que repetir la simulación en cada caso? Probablemente no pues eso sería muy costoso cognitivamente. Más adelante, veremos que Goldman (1989) sugiere que los resultados de las simulaciones son almacenados para futuros usos.

Un aspecto relevante en la formulación de Gordon (1986) es que la teoría de la simulación no nos pone en la mente del otro sino sólo en su lugar. Gordon señala que al efectuar la simulación no se hace ningún intento de entrar a la mente del otro. Lo único que se hace es un traslado imaginativo a su situación.⁵⁷ Una pregunta surge ineluctablemente ¿Cómo podría predecir la conducta de los otros a partir de mi razonamiento práctico alimentado con insumos hipotéticos si en principio soy diferente en aspectos relevantes a los demás? Lo natural es pensar que si no tengo una teoría que me proporcione leyes que relacionen tipos de insumos con tipos de estados psicológicos y con tipos de conductas, entonces mis alcances predictivos están reducidos a las personas que son muy semejantes a mí. Gordon reconoce que normalmente la gente es capaz de predecir la conducta a pesar de las diferencias entre ellos. Si la Teoría de la Simulación está a la base de la PP entonces tiene que explicar este rasgo. La idea que propone Gordon es que normalmente cuando hay diferencias entre uno y el blanco de la predicción se procede a hacer ciertos ajustes como: "... trade one set of idiosyncrasies for another" como en el los pasos 6, 7, y 8.⁵⁸ Puedo imaginar situaciones en las cuales muchos de los

⁵⁶ Gordon (1986), pp. 64 y 65. Gordon no hace la formulación explícita pero puede fácilmente seguirse de sus ejemplos.

⁵⁷ Gordon (1986), p. 63.

⁵⁸ Gordon (1986), p. 63. "cambiar un conjunto de idiosincrasias por otro".

estados psicológicos y estímulos ambientales que experimento sean diferentes de lo que son ahora y obtener un posible curso de acción si se cumplen las condiciones mencionadas. Si la situación hipotética imaginada es en términos generales muy similar a la que un sujeto diferente de mí está experimentando ¿por qué no sería capaz de tomar el resultado de mi razonamiento práctico como una predicción de su conducta? A la posibilidad de sortear las diferencias entre conjuntos de creencias subyace, sin embargo, la idea de que se preserva cierta semejanza entre los seres humanos. La Teoría de la Simulación presupone que las otras personas son como uno.⁵⁹

Ahora bien, Gordon (1986) interpreta dos estudios extraídos de la literatura psicológica como argumentos a favor de la teoría de la simulación. El primero es Wimmer y Perner (1983). En este trabajo los autores formulan un tipo de prueba conocida como tarea de “creencia falsa” o experimentos tipo “Maxi”. La idea básica fue exponer a un grupo de infantes de diferentes edades a dos escenas donde se presenciaba la colocación de un objeto (un chocolate) en un sitio x por un personaje (llamado Maxi) y más tarde, en ausencia del personaje el objeto se cambiaba del lugar inicial x, a un lugar y. El objetivo era descubrir qué grado de estabilidad sobre las representaciones de las creencias del otro tenían los participantes. En la primera escena, Maxi entraba a buscar su chocolate y se le cuestionaba a los infantes donde lo haría. En la segunda escena, Maxi nuevamente entraba a buscar su chocolate pero ahora se enfrentaba con una situación competitiva donde su hermano también lo deseaba. La cuestión era determinar cómo respondería Maxi a la pregunta de su hermano, ¿Dónde está el chocolate? Los resultados arrojados fueron que los niños de 3-4 años de edad tuvieron 58% de las respuestas correctas, los de 4-6 tuvieron 86% de respuestas correctas y los de 6-9 acertaron en todas las respuestas. ¿Por qué surge esta diferencia de exactitud según las edades? Parece que los niños incrementan su capacidad de manejar meta-representaciones según se incrementa la

⁵⁹ Gordon no abunda demasiado sobre qué tipo de conocimiento constituyen los conjuntos de idiosincrasias y las normas y valores.

edad. Gradualmente son capaces de considerar estados mentales más complejos. ¿Cómo puede explicarse esto? Gordon (1986) considera que el niño desarrolla la habilidad de hacer afirmaciones dentro del contexto de la simulación práctica. De este modo, si el niño puede ponerse en los zapatos del otro entonces sus respuestas serán las correctas al trascender la visión egocéntrica del mundo. Los infantes de 4 años en adelante pueden resolver las pruebas correctamente pues ya manejan la simulación práctica. Para Gordon la explicación correcta de esto es que han desarrollado su capacidad de alimentar su mecanismo de razonamiento práctico con insumos hipotéticos.

El siguiente argumento es extraído de Baron-Cohen et. al. (1985). El estudio fue realizados con infantes de diferentes tipos psicológicos: normales, con síndrome de down, y con autismo. Se procuró que el nivel de edad mental de los infantes fuera semejante. Se tomo como base a los experimentos realizados por Wimmer y Perner (1983). Dos personajes (Sally y Anne) son observados por los niños de los diferentes grupos. Sally coloca una canica en una canasta en presencia de Anne y abandona la escena. En ausencia de Sally, Anne traslada la canica a una caja. (Fig. 2) Sally regresa a escena y el experimentador cuestiona al público con las siguientes preguntas: ¿Quién es Sally y quién es Anne? (pregunta de identificación) ¿Dónde buscará Sally la canica? (pregunta de creencia) ¿Dónde está la canica realmente? (pregunta de realidad) y ¿Dónde estaba la canica al principio? (pregunta de memoria).⁶⁰

⁶⁰ Baron-Cohen (1985), p. 41.

Figure 1. *Experimental scenario.*

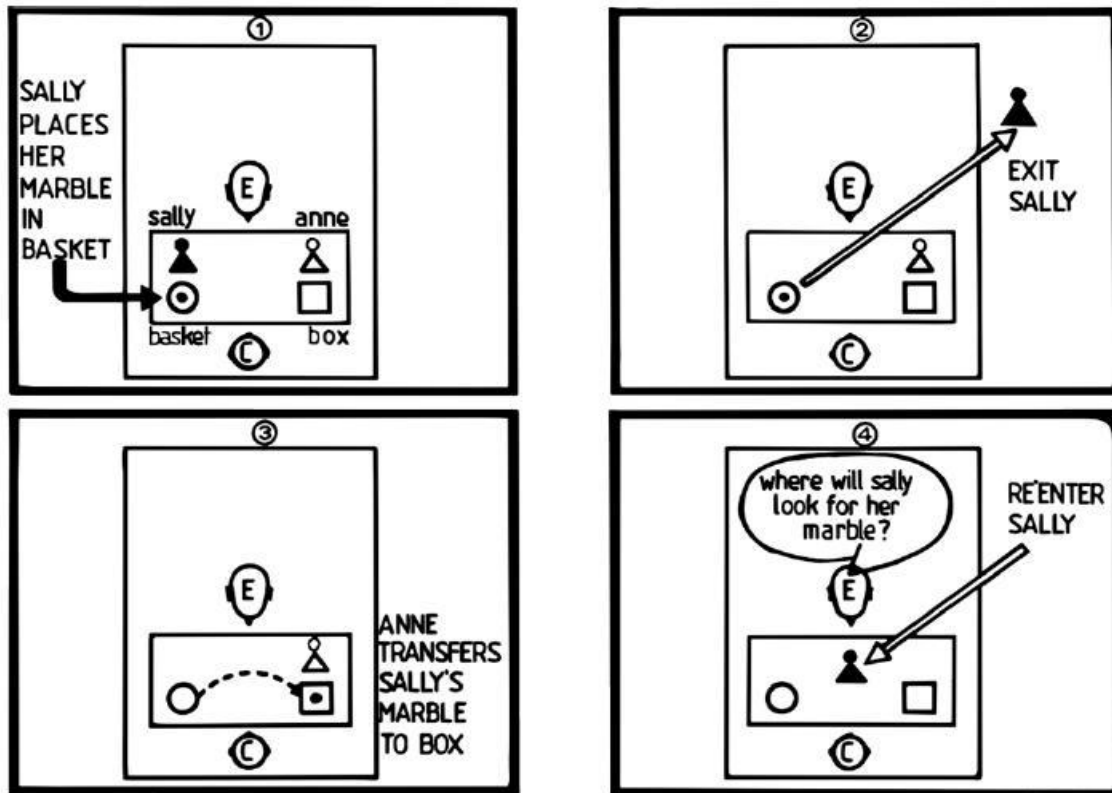


Fig. 2. En el esquema se muestra la estructura del estudio.

Los resultados arrojaron evidencia muy interesante. Los niños normales y los que sufren de síndrome de Down respondieron con una tasa de éxitos muy semejante a todas las preguntas. No obstante, esto no ocurrió con los autistas. De 20 niños con autismo que se analizaron 16 fallaron en la pregunta de creencia en las dos ocasiones que se realizó. Esto sugiere que los niños autistas carecen de la capacidad de mentalizar. No parecen tener la habilidad de atribuir estados mentales a los otros. Fallan en la capacidad de cambiar de perspectiva. Gordon (1986) sugiere que esto puede explicarse gracias a la Teoría de la Simulación. El hecho de que incluso infantes con una deficiencia mental severa como los niños con síndrome de Down hayan tenido éxito en las pruebas tiene que ver con que son capaces de involucrarse en la simulación. Lo que les hace falta a los

autistas es, precisamente, este mecanismo que les permita atribuir creencias a los otros gracias al traslado egocéntrico. Si pudieran ponerse en los zapatos de los otros, entonces habrían podido saber donde buscaría Sally la canica. Cabe señalar que los autistas respondieron adecuadamente a las tres preguntas adicionales lo cual indica una inteligencia adecuada en otros campos.

Hasta aquí los dos argumentos que Gordon ofrece para apoyar a la Teoría de la Simulación. Su conclusión es que la incapacidad de mentalizar es explicada por una falla en el mecanismo de razonamiento práctico cuando es puesto a trabajar *off-line*. Los autistas no pueden desacoplarlo de sus entradas y salidas típicas, por ello no pueden simular, y en consecuencia, no pueden leer mentes. Finalmente, Gordon sugiere que podría haber un módulo mental que se activa cada vez que detecta seres humanos para permitir la interacción social. Su precursor podría ser el módulo de atención compartida (SAM). Sin duda, la alternativa que Gordon (1986) plantea contribuye a enriquecer la imagen que tenemos de los procesos mentales en el ámbito psicológico. En este punto, los resultados empíricos arrojarán importantes datos sobre qué tipo de teoría favorecer.

Antes de exponer la variante simulacionista de Goldman (1989) quisiera plantear un par de dudas hacia la postura de Gordon (1986).

Una primera cuestión que salta a la vista con el simulacionismo de Gordon es que si se apoya en roles institucionales para hacer sus predicciones, como fue evidenciado en el punto 5 de mi reconstrucción de la teoría de Gordon. Los roles institucionales son entendidos como esquemas de acción dirigida a ciertos fines. Si el usuario emplea un cuerpo internalizado de conocimientos que dice cómo relacionar estados mentales con comportamientos y demás, parece que está empleando un tipo de teoría sobre cierta parcela psicológica.⁶¹ En este caso, una simulación sin soporte teórico no es suficiente para obtener resultados satisfactorios. Por tanto, esto constituye un debilitamiento de la

⁶¹ Gordon (1986), pp. 63-65.

formulación original que pretende no echar mano de teoría alguna. Si algunas conductas son predichas según los principios establecidos por un marco institucional entonces no toda interacción social se efectúa bajo simulación. Así el dominio de la Teoría de la Simulación en principio se ve disminuido.

Un segundo punto de interés surge a partir de reflexionar sobre la predicción y explicación de las conductas. Normalmente hay apoyo de conocimiento previamente adquirido sobre el agente en cuestión o por lo menos expectativas. La teoría de la simulación reconoce el hecho de que al hacer predicciones de la conducta de un sujeto S, se toman en cuenta varios aspectos como: (a) Los estímulos que recibe el sujeto en cuestión S, (b) la conducta actual de S, (c) los posibles estados mentales de S en ese momento determinado. Sin embargo, en toda predicción/explicación subyacen ciertos elementos que apuntan al empleo de algún tipo de conocimiento previo. Por ejemplo, cuando se hace una simulación también se toman en cuenta cosas como: (d) reportes pasados de creencias y deseos de S, (e) regularidades en la conducta de S, (f) atribuciones exitosas de estados mentales a S, (g) los contextos sociales en que S está inserto. La consideración de los incisos (a) – (g) sugiere el empleo de un cuerpo previo de conocimientos para predecir y explicar las conductas de las personas a pesar de que se opere mediante la metodología simulacionista. Naturalmente estos conocimientos son adquiridos empíricamente a lo largo de la vida del individuo en sociedad. Pero es prudente señalar que la formulación de la Teoría de la Simulación en la versión de Gordon (1986) sólo reconoce estos hechos marginalmente.

2. Teoría de la Simulación: Goldman (1989)

Continúo ahora con la versión de la Teoría de la Simulación por parte de Goldman (1989).

Básicamente Goldman sostiene que:

They (the interpreters) ascribe mental states to others by pretending or imagining themselves to be in the other's shoes, constructing or generating the (further) state that they would then be in, ascribing that state to the other. In short we simulate the situation of others, and interpret them accordingly. (Goldman (1989), p. 81.)⁶²

Según Goldman, la simulación requiere imaginarse “estar en los zapatos del otro” y a partir de esta situación inferir el estado en el que estará el sujeto cuya conducta se quiere predecir. Este proceso de atribuir estados mentales y predecir conductas sugiere tomar como base determinados estímulos que el agente está experimentando junto con ciertos estados mentales que se supone tendría en su situación específica. La idea de ponerse en los zapatos de otro sugiere imaginarse como estando en la situación en la que el otro del mismo modo como sugiere Gordon (1986). En esta situación no es que se entre a la mente del otro sino que se supone cual sería su estado psicológico en esa situación y de acuerdo con eso se hace una elección sobre qué tipo de salida conductual o estado mental estaría presente. Este resultado es producido en concordancia con la suposición de que el otro tiene los mismos anhelos, gustos básicos y resultados de experiencias perceptuales. En el siguiente fragmento aclara:

From your perceptual situation, I infer that you have certain perceptual experiences or beliefs, the same ones I would have in your situation. I may also assume (pending information to the contrary) that you have the same basic likings that I have: for food, love, warmth, and so on. (Goldman (1989), p. 82)⁶³

⁶² “Ellos (los interpretes) atribuyen estados mentales a los otros al pretender o imaginarse a sí mismos en los zapatos de otros, construyendo o generando el estado (adicional) en el que ellos estarían, atribuyendo el estado al otro. En breve, simulamos la situación de los otros, y los interpretamos en concordancia”.

⁶³ “De tu situación perceptual, infiero que tienes ciertas experiencias perceptuales o creencias, las mismas

Esto quiere decir que al realizar el ejercicio de simulación suponemos cierta similitud con el objeto cuya conducta ha de predecirse.

Como podemos ver gran parte la formulación de Goldman es similar a la de Gordon (1986). Por ello, considero importante concentrarme ahora en los puntos donde Goldman y Gordon difieren. La concepción de Goldman introduce una división interesante y útil al interior de la simulación: la simulación guiada por proceso y la simulación guiada por teoría. Stich y Nichols (1992) consideran que una teoría es un cuerpo internamente almacenado de conocimientos sobre algún dominio específico. Esta noción será útil para seguir la formulación de la dicotomía realizada por Goldman. Primero, hay que tomar en cuenta que damos explicaciones y hacemos predicciones de los más variados sucesos. Podemos predecir la conducta de un perro irritable ante un infante corriendo y gritando cerca de él. Pero también explicamos y predecimos el comportamiento de máquinas, de los mercados internacionales, de los volcanes, etc. El partidario de la TT podría explicar nuestra capacidad para hacer predicciones y dar explicaciones sobre estos ámbitos apelando a la posesión de una teoría sobre cada campo. Goldman considera que la simulación es posible cuando un sistema tiene un isomorfismo relevante con lo que modela.⁶⁴ Naturalmente esta condición no es satisfecha en ninguno de los dos casos anteriores. Somos diferentes en muchos sentidos a un perro irritable y a máquinas, los mercados internacionales, y los volcanes. Sin embargo, podemos predecir y explicar sus conductas. Pero lo hacemos gracias a una teoría y no a una simulación. No podríamos modelar la conducta de muchos sistemas si no tenemos una teoría que nos guíe en este ejercicio. Pero ¿ocurre lo mismo al enfrentarnos a otros seres humanos? La respuesta es no, hay procesos que podemos simular sin necesidad de apoyarnos en cuerpos de conocimientos internamente representados. Goldman entonces efectúa su distinción. Una simulación guiada por teoría es aquella que necesita de un cuerpo de conocimientos sobre un dominio específico para construir o anticipar la secuencia de estados en los que está o

que yo tendría en tu situación. Puedo suponer adicionalmente (en espera de información contraria) que tienes los mismos gustos que yo tengo: por comida, amor, calidez, etc”.

⁶⁴ Goldman (1989), p. 85.

estará el sistema en cuestión. Hay una manera alternativa de realizar la simulación. Ocurre cuando el blanco de la predicción B tiene un isomorfismo relevante con el observador A y los procesos que guían la conducta B son los mismos que, básicamente, emplea el A para realizar la reconstrucción de la secuencia de estados mentales que sigue B. Esta es la simulación guiada por proceso.

En otras palabras, para que se de una simulación guiada por proceso, dado un par de agentes A (un intérprete) y B (un agente intencional), debe ocurrir que:

1. Los procesos cognitivos de A y B sean los mismos o relevantemente semejantes.
2. Los estados iniciales de A y B sean los mismos o relevantemente semejantes.

Como resultado el simulador obtendrá el mismo resultado que el agente observado.

Pero si el agente B tiene estados mentales diferentes la estrategia para A es tratar de fingir estar en esos estados. Entonces si el punto 2 de nuestra reconstrucción no se presenta tenemos que:

3. Alimentar al mecanismo de simulación con insumos hipotéticos que coincidan con los insumos que B recibió o recibe.

Así, el camino a seguir al percibir diferencias entre uno mismo y los demás es tratar de imaginar o aparentar tener los mismos estados iniciales que puede tener el interpretado. Estos estados iniciales imaginados producirán típicamente los mismos resultados que los del blanco de la predicción si se opera con los mismos procesos cognitivos.⁶⁵

Idealmente, si estamos en la misma situación del agente a predecir y operamos con los mismos procesos cognitivos entonces produciremos el mismo resultado, pero ¿qué ocurre cuando fallamos en nuestras predicciones? Hay casos donde el conocimiento que cada uno posee altera el curso de acción. Tal es el caso de un maestro de ajedrez y de

⁶⁵ Goldman (1989), p. 85.

un principiante. El novato difícilmente podrá predecir los movimientos del maestro. En este caso los insumos espacio-temporales pueden ser muy semejantes y el mecanismo cognitivo involucrado también, sin embargo, hay fallas predictivas. ¿Es normal que estas fallas se presenten? Goldman responde: “In any case, there is no assumption here that people are always successful or optimal simulators”.⁶⁶ La teoría de la Simulación no garantiza eficacia total. Goldman sugiere que si se desea tener éxito en la predicción y se encuentran diferencias relevantes entre las creencias y metas del sujeto a interpretar y las del intérprete no siempre se producirá el resultado correcto. Goldman no explica claramente por qué esto ocurre.⁶⁷

Ahora bien, Goldman describe 4 casos que apoyan la teoría de la simulación:

1. El caso del Señor Crane y el señor Tees. Dos hombres tienen diferentes vuelos programados a la misma hora y comparten el mismo automóvil que los lleva al aeropuerto. En el camino se atorán en el tráfico y se retrasan media hora. Cuando finalmente llegan al aeropuerto son notificados que el vuelo del señor Crane salió a tiempo pero que el del señor Tees salió hace tan sólo cinco minutos. Goldman se pregunta: ¿Quién está más molesto? La mayoría de la gente responde que el señor Tees. Si la Teoría de la Simulación está a la base de la PP entonces se explica fácilmente porque la gente acierta en casos como este. Claramente, se puso en los zapatos del otro.⁶⁸

2. El caso de *Sense and Sensibility*. Dos sujetos, Mary y Peter, van caminando y observan un paisaje. Mary le dice a Peter: “*Es el tipo de escena que haría que*

⁶⁶ Goldman (1989), p. 83 “En cualquier caso, aquí no está la suposición de que la gente siempre es simuladora óptima o exitosa”.

⁶⁷ Goldman llama intérprete a cualquier sujeto que atribuya actitudes proposicionales.

⁶⁸ Goldman (1989), p. 83.

Marianne Dashwood se desmayara". Mary está haciendo referencia a uno de los personajes del libro de Jane Austen *Sense and Sensibility*. Ella supone que Peter ha leído la novela y que la frase le parecerá significativa en ese contexto. La mejor explicación de que piense esto proviene de la Teoría de la Simulación. Nuevamente, Mary se puso en los zapatos de Peter y dado que a ella le parecía apropiada la escena para proferir esa frase, entonces pensó que Peter reaccionaría igual, recordando la obra y entendiendo el sentido de la emisión.⁶⁹

3. El caso del *Modus Ponens*. Supongamos que un sujeto A escucha dos frases que estructuralmente son del tipo $P > Q$ y P . Es muy probable que todos hagan predicciones correctas sobre el tipo de respuesta que dará A. Goldman sugiere que no tenemos una teoría sobre las inferencias de la lógica de primer orden. Lo que hacemos para predecir la conducta de A es tomar las dos frases como insumos para nuestro sistema de inferencias y echarlo a andar. El resultado lo atribuimos al sujeto A.⁷⁰ Otra vez, este es un caso de simulación, y claramente, un caso de simulación guiada por proceso.

4. El caso de lo gracioso. Las personas somos bastante hábiles para predecir qué le resultará gracioso o divertido a los demás. Goldman sugiere que esta habilidad no es sostenida por algo como una teoría sobre el humor. Tiene más sentido pensar que al proferir algo que nos parece gracioso empleamos la simulación.⁷¹

Es importante notar que todos estos casos incluyen un alto nivel de dominio lingüístico. No sería extraño pensar que la teoría de la simulación está vinculada, de algún modo, con la facultad del lenguaje. Si lo anterior es el caso, podría resultar que especies

⁶⁹ Goldman (1989), p. 84.

⁷⁰ Goldman (1989), p. 84.

⁷¹ Goldman (1989), p. 84.

más simples que la nuestra no poseyeran la capacidad de simular al no poseer un lenguaje tan complejo.

Goldman se apoya en los estudios realizados por Wimmer y Perner (1983) y Baron-Cohen, et. al. (1985) como argumentos a favor de la teoría de la simulación. Dado que ambos argumentos ya fueron expuestos en la sección dedicada a Gordon (1986) no los repetiré en esta parte. Sólo comentaré dos rasgos de la Teoría de la Simulación que Goldman ofrece cuando se pregunta al respecto de la relevancia fenoménica de esta estrategia; no siempre somos conscientes del traslado hipotético que se supone debe realizarse para simular. En primer lugar, Goldman responde que la simulación no necesita ser introspectivamente vívida. Cuando un mecanismo está altamente desarrollado, normalmente se automatiza. Por tanto, cuando tratamos con mecanismos de este tipo es común que no exista una fenomenología sobresaliente.⁷² En segundo lugar, Goldman reconoce que un agente maduro puede construir ciertas generalizaciones a partir de la reiteración del resultado de las simulaciones. Es decir, el intérprete no tiene que correr la simulación en todas las ocasiones en las que quiere explicar o predecir la conducta de un agente. Puede recurrir al conocimiento adquirido inductivamente. Pero ¿acaso no es esto equivalente a decir que el usuario de la PP a veces se apoya en algo como una teoría? Goldman reconoce que en este caso hay que hacer una concesión al partidario de la TT. Efectivamente, al mentalizar, no siempre se emplea la teoría de la simulación. En ocasiones se recurre a principios generales. Sin embargo, las generalizaciones obtenidas están construidas por instancias donde sí se aplicó la simulación originalmente.⁷³ Goldman reconoce que su concesión dar pie a una variante de la TT. Hay dos diferencias claras con esta última: (a) que la TT no acepta a la simulación, y (b) que la Teoría de la Simulación se postula como el origen de las generalizaciones.⁷⁴

⁷² Goldman (1989), p. 88.

⁷³ Goldman (1989), p. 88.

⁷⁴ Goldman (1989), p. 88.

Hasta aquí la revisión de la concepción de Goldman (1989). Básicamente, la diferencia de su propuesta con la de Gordon (1986) radica en dos aspectos: 1. su distinción entre simulación guiada por proceso y guiada por teoría, y 2. La posibilidad de que haya una versión mixta que incorpore bases de conocimiento en el despliegue de la PP. Ambos sostienen que hay un mecanismo de razonamiento práctico que puede ser alimentado con insumos hipotéticos y que da como resultado un estado que no activa las salidas motoras.

3. Retos a la Teoría de la Simulación

En esta parte lanzaré algunos retos que tendría que superar un simulacionista.

Un problema con la simulación tiene que ver con que las predicciones y explicaciones de conducta son “cognitivamente penetrables”.

Esta objeción proviene directamente de Stich y Nichols (1992). Según los autores las predicciones que hacemos en la vida cotidiana sobre la conducta de los otros son más exactas en la medida en que poseemos mejores teorías sobre sus vidas mentales. Por consiguiente, el error en las predicciones es producto de teorías incompletas o equivocadas. El conocimiento que tengo importa para la eficacia predictiva.

Stich y Nichols (1992) pusieron a prueba a la Teoría de la Simulación mediante estudios experimentales basados en los hallazgos de Ellen Langer (1975). Básicamente los experimentos de Langer consistieron en vender a dos grupos de individuos boletos para una quiniela pues se acercaba un partido de football. A un grupo se le dio la oportunidad de elegir los boletos pero no a los miembros del otro grupo a quienes les fueron asignados. Langer descubrió que los sujetos que eligen tienden a pensar que las probabilidades de ganar son mayores, en cierto modo, se crea así la ilusión de control. Los boletos fueron vendidos en un dólar por igual. Langer más tarde les pidió a los sujetos que le vendieran los boletos de vuelta obteniendo resultados intrigantes. Los sujetos del grupo de elección vendieron los boletos en un rango de 6 a 8 veces su precio original. En cambio los sujetos del grupo de asignación sólo incrementaron el precio del boleto una o dos veces.

El trabajo de Stich y Nichols (1992) fue pedirles a algunos alumnos que predijeran la conducta de ambos grupos. Al contrario de lo que se podría pensar la gran mayoría falló en sus predicciones pues no esperaban que el notable incremento de los precios de venta por parte de los sujetos de elección. Si la teoría de la simulación fuera cierta, los alumnos tendrían que haber tenido un buen desempeño en la predicción de las conductas de los dos conjuntos de apostadores.

Stich y Nichols (1996) mencionan que bajo la Teoría de la Simulación es posible errar en las predicciones en dos escenarios: a) que el mecanismo de toma de decisiones del que predice la conducta sea diferente del mecanismo del blanco de la predicción; y b) que el generador de deseos y creencias hipotéticos no haya proporcionado las creencias y deseos hipotéticos adecuados al sistema de toma de decisiones.⁷⁵

Stich y Nichols (1992) afirman que sí el mecanismo de toma de decisiones del sujeto que predice la conducta del otro y aquel del blanco de la predicción son relevantemente semejantes entonces la predicción será correcta. Lo mismo ocurre si las creencias y deseos hipotéticos con que se alimentó el sistema de toma de decisiones o de razonamiento práctico son los mismos que los del blanco de la predicción. En el experimento de Langer-Stich y Nichols resulta que los sujetos satisfacían estos dos últimos requisitos en la simulación pero fallaban en sus resultados predictivos. Esto parece un grave problema para la teoría de la simulación. Sin embargo, la TT explica esto muy bien. Bajo la perspectiva de la TT los sujetos fallaron en sus predicciones pues no tenían conocimiento acerca del efecto Langer también conocido como “la ilusión de control” (que afirma que los sujetos que eligen tienen la ilusión de que hay mayor probabilidad de obtener el resultado deseado en comparación con quienes no eligen). Dado que la eficacia predictiva depende de la penetrabilidad cognitiva del mecanismo empleado en la predicción, si no se tiene la base de conocimientos correcta o está incompleta la predicción falla.⁷⁶

Otro reto que debe superar la Teoría de la simulación es dar cuenta de qué hacemos para explicar cuando las diferencias entre los sujetos relevantes. Esto ocurre cuando tenemos que explicar o predecir las conductas de agentes en el pasado, o incluso acciones de personajes históricos, a los cuales no podemos observar directamente y no conocemos el desenlace de sus acciones. En estos casos, es difícil imaginar la situación

⁷⁵ Stich y Nichols (1996), p. 49.

⁷⁶ Stich y Nichols (1992) y (1996).

espacio-temporal que enfrentó el blanco de nuestra predicción/explicación si no nos basamos en algún tipo de conocimiento sobre las costumbres, normas, y valores de la sociedad de esa época. Imaginemos que a un sujeto, que no sabe nada de Grecia ni de filosofía, se le pide que haga una predicción. Tiene que decir qué hará un sujeto en la siguiente situación:

1. Está condenado a muerte.
2. Es víctima de un juicio injusto.
3. Es un pensador muy brillante. Además, dentro y fuera de su ciudad es admirado por muchos.
4. Sus amigos han comprado a los guardias para que pueda salvarse y escapar.
5. Tiene una esposa y tres hijos.

El mecanismo de simulación seguramente producirá como resultado la predicción de que el sujeto escapará. Después de todo, si uno fuera condenado a muerte víctima de un juicio injusto con una carrera académica sobresaliente y con una esposa y tres hijos sería natural decidir escapar. Pero si conocemos algunos hechos sobre el sistema de creencias de este individuo, nuestra predicción cambiará radicalmente. Todos podemos entender por qué Sócrates decidió beber la cicuta y rechazó escaparse si conocemos la teoría que guiaba sus acciones. El simple hecho de ponernos imaginativamente en el lugar espacio-temporal de Sócrates y alimentar nuestro mecanismo de simulación con la invitación de Critón para escapar de prisión, y todas las condiciones antes señaladas, no produce la predicción correcta de su acción. En este caso parece que la simulación no es suficiente para acertar en las predicciones. Se necesita cierto conocimiento adicional para que se produzcan resultados correctos. Las diferencias entre los sujetos son tan grandes que es necesario emplear una teoría para dar sentido a las acciones de los otros.

No es necesario ir tan lejos. Las diferencias de género son lo suficientemente claras para arrojar luz sobre la insuficiencia de la simulación. Si una mujer simula la conducta de otra mujer, es más probable que acierte en sus predicciones/explicaciones que si un hombre simula a una mujer. En este último caso tendría que postular una serie de deseos y creencias hipotéticas para alimentar al mecanismo de simulación, y así, acertar en sus predicciones y explicaciones. En este caso, parece necesario un cuerpo de conocimientos sobre un dominio específico.

Una última cuestión antes de pasar al siguiente capítulo. Gordon (1986) y Goldman (1989) han sugerido que la Teoría de la Simulación es más simple que la TT. Esta afirmación se sustenta que los simulacionistas no tienen que postular un cuerpo de conocimientos, una base de datos, o una teoría para explicar a la Psicología Popular. Un mecanismo de toma de decisiones es suficiente para realizar dos tareas, producir acciones, y producir predicciones/explicaciones. En cambio la Psicología Popular bajo la concepción antagónica tiene que postular una teoría. La conclusión es que resulta más conveniente elegir una teoría más simple a una más complicada si ambas enfrentan exitosamente a los fenómenos.

Para aclarar este punto recordemos las etapas en el proceso de la simulación que habíamos observado en el caso de Gordon. Antes de correr la simulación es necesario:

1. Detectar a un objeto X como agente intencional.

En este punto no queda claro qué criterio se utiliza para considerar a un objeto como agente. El simulacionista habla de isomorfismo. Pero cómo detectamos este isomorfismo no queda claro.

2. Ver los movimientos de X como conducta.

Aquí parece surgir la misma duda del caso anterior. No es claro el criterio para saber cuando un suceso físico puede verse como una conducta intencional.

3. Intención de comprender, predecir, o explicar las acciones.

En el inicio de la simulación las cosas se empiezan a complicar pues hay que:

4. Cambiar imaginativamente perspectivas espacio-temporales con el agente en cuestión.

Esto resulta difícil si la distancia espacio-temporal es muy grande.

5. Cambiar papeles institucionales.

En este caso parece ser necesario conocer primero los papeles institucionales para después poder cambiarlos. Este es un conocimiento que no se obtiene de manera simulacionista, sino registrando regularidades que constituyen generalizaciones.

6. Racionalizar la conducta del agente buscando y alterando sus posibles creencias y deseos preservando coherencia.

Esta etapa implica poder manejar una serie de conjuntos diferentes de creencias en muy poco tiempo.

7. Determinar aspectos de origen que pudieran dar sentido a la conducta del agente.

Buscar correlaciones entre contextos, historias de vida y conductas. Aquí nuevamente es necesario tener un conocimiento bastante amplio de generalidades culturales previamente almacenadas.

8. Crear hipótesis alternativas sobre su conducta.

El número de las hipótesis competitivas no tiene restricción en principio y puede ser tan grande como los recursos cognitivos lo permitan. Esto pone en duda la economía del proceso.

9. Eliminar hipótesis descabelladas si están demasiado alejadas de las tendencias naturales del observador.

Hacer una depuración de hipótesis que no pueden entrar a la competencia por explicar o predecir una conducta. El criterio parece ser que si están muy alejadas de la idiosincrasia propia entonces se deben descartar. Un problema con esto surge cuando las diferencias con el blanco de la predicción/explicación son lo suficientemente relevantes que resulta necesario recuperar estas hipótesis previamente descartadas. El gasto cognitivo parecería ser inmenso en este punto.

10. Probar las hipótesis buscando coherencia.

Ya que hemos conseguido un conjunto de creencias y deseos para el agente en cuestión y una hipótesis sobre su posible conducta, entonces es necesario ponerla a prueba. Si falla la combinación entonces habría que cambiar subconjuntos de deseos o de creencias por alternativas.

11. En caso de encontrar un par de hipótesis con adecuación empírica, juzgar nuevamente según la propia tendencia y repetir el paso 10.

En ocasiones, este criterio de elección de hipótesis tiene que desecharse pues puede ser obvio que la otra persona se aleja de nuestras propias tendencias. Entonces hay que buscar un criterio que nos permita modificar este punto a conveniencia.

12. Al hallar una hipótesis correcta, crear un mundo simulado que corresponda al agente y que tenga adecuación empírica.⁷⁷

Este punto parece sugerir que creamos un conjunto de relaciones causales entre insumos, creencias, deseos, sensaciones, emociones, etc. y conductas que, suponemos, el agente va a respetar. Bajo esta lectura, claramente se sigue que obtenemos algo como una teoría particular sobre ese individuo.

⁷⁷ Gordon (1986), pp. 64 y 65. Gordon no hace la formulación explícita pero puede fácilmente seguirse de sus ejemplos.

Vista de este modo, la simulación no parece ser más económica que la TT y tampoco parece estar carente de teoría. En muchas de sus etapas emplea conjuntos de datos sobre dominios específicos, como relaciones entre contexto y conducta, historia personal y conducta, detección de movimiento y atribución de intencionalidad, etc. Por tanto, la Teoría de la Simulación no parece ser tan simple como lo sostienen sus partidarios. Más aún, si tomamos en cuenta que en sociedad nos enfrentamos con más de un agente a la vez cuya conducta queremos explicar o predecir, el proceso se multiplica increíblemente y la carga cognitiva parece inmanejable para el simulacionista.⁷⁸

En este capítulo hemos visto una concepción que busca dar cuenta de la Psicología Popular conocida como La Teoría de la Simulación. Exploré dos versiones de esta propuesta. Por un lado, tenemos a la simulación radical de Gordon (1986) que excluye cualquier uso de teorías. Por otro lado, está la simulación de Goldman (1989) que crea la división entre simulaciones guiadas por proceso y las simulaciones guiadas por teoría.

La simulación *off-line* se presenta como una alternativa a la TT pues elimina cualquier teoría del núcleo de la Psicología Popular. Funciona desacoplando al mecanismo de razonamiento práctico de sus entradas y salidas naturales para ser alimentado con insumos hipotéticos. El resultado es una explicación o una predicción de la conducta de un agente.

Al final del capítulo presenté algunos puntos controvertibles en torno a la Teoría de la Simulación. Uno de ellos fue la idea de la penetrabilidad cognitiva de Stich y Nichols (1992). Después hice una observación sobre las diferencias espacio-temporales y cognitivas entre el blanco de la predicción y el agente intencional. Finalmente, argumenté que la Teoría de la Simulación no es tan simple como pretenden sus defensores.

⁷⁸ Este último punto se verá con más detalle en el siguiente capítulo.

Capítulo III

Hacia una Teoría Híbrida

Introducción

Para justificar la posible unificación de las dos concepciones diferentes es preciso que aisladamente ambas tengan deficiencias que no permitan dar una explicación de todos los fenómenos involucrados y relevantes en el uso de la Psicología Popular. Los seis argumentos que veremos a continuación (y que lejos de ser concluyentes, buscan sólo esbozar la posibilidad de una teoría híbrida) están basados en la insuficiencia explicativa como punto de partida para la incorporación de ambos enfoques. Este capítulo no tiene como propósito crear una teoría híbrida, sino que el objetivo es mostrar como los problemas que tienen que manejar las teorías aisladamente sugieren la integración. Por último, hablare brevemente sobre cómo podría ser el primer paso para integrar a las dos concepciones.

1. Argumentos

El primer argumento tiene como propósito mostrar que ni la Teoría de la Simulación ni la TT separadas son suficientes para dar cuenta de la FP. La conclusión es que se pueden emplear ambas según el contexto.

El partidario de la Teoría de la Simulación supone que su alternativa es la correcta cuando de dar cuenta de la FP se trata. El punto en esta parte es descubrir si algún tipo de simulación se efectúa sin teoría y hallar cuál es. Davies y Stone (1996) proponen cuatro ejemplos clave en los cuales se puede investigar lo anterior empleando la estrategia de la simulación mental para predecir la conducta de una persona C:

- (i) ¿cómo se sentirá C después de tomar medio litro de whisky?
- (ii) ¿cómo lucirá la ilusión Müller-Lyer para C?
- (iii) ¿cómo se sentirá C y qué decidirá hacer si está suspendido en un precipicio sujeto de una cuerda sin poder encontrar un punto de apoyo y con sus manos empezando a resbalar?
- (iv) ¿sacará C la conclusión de que algo es blanco a partir de su creencia de que la nieve es blanca?⁷⁹

Si tengo a un sujeto *D* que es similar en aspectos relevantes a un sujeto *C* y lo expongo a las mismas situaciones para observar cómo actúa, es muy probable que obtenga un resultado apropiado para el caso de *C*. Algunos modos en los cuales *D* debe ser similar a *C* son: (i) las maneras en las que el alcohol afecta su constitución corporal; (ii) la manera en que trabaja su sistema visual; (iii) el modo en que experimenta y actúa según sus emociones; (iv) el modo en que razona.⁸⁰

Ahora bien, si ese individuo *D* no está a mi alcance (como normalmente ocurre), tengo a otro individuo muy similar a *C*, a saber: a mi mismo. De este modo, si quiero predecir cómo actuará *C* simplemente debo seguir la estrategia de exponerme a las situaciones antes mencionadas. Luego, el resultado de este ejercicio se lo atribuyo a *C* como una predicción de su conducta en tales circunstancias. Lo anterior sería una instancia de simulación en la realidad. Atribuyo los resultados de mis experiencias en las situaciones (i)-(iv) a *C*.⁸¹ En este caso, no estoy empleando en principio una teoría, a menos que la hipótesis de semejanza entre *C* y yo la presuponga.

Supongamos que no quiero experimentar todas esas situaciones conmigo mismo. Aún tengo otra opción. Es posible hacerlo con un sujeto imaginario. No obstante, al poner

⁷⁹ Davies y Stone (1996), p. 57.

⁸⁰ Davies y Stone (1996), pp. 58-59.

⁸¹ Davies y Stone (1996), p. 58.

a una persona imaginaria, me estoy basando en una teoría que me indica cómo actuarán las personas típicamente en esas situaciones. Si elimino la idea de otra persona y me imagino a mí mismo en esa situación podría ser que algunas predicciones fuesen generadas. Si al imaginarme en esa situación paso por los mismos procesos psicológicos que el sujeto cuya conducta deseo predecir entonces tengo una esperanza de que el resultado no esté basado en una teoría.

Hemos visto cuatro situaciones en las cuales se puede emplear la teoría de la simulación. Es prudente preguntarnos en cuáles de ellas se realiza una simulación imaginaria o real. En primer lugar, si quiero predecir cómo se sentirá un individuo después de beber medio litro de whisky basándome en la estrategia simulacionista, tendría que producirse en mi imaginación por lo menos la sensación de estar mareado. En el segundo caso, sólo tendría una predicción correcta del estado psicológico de *C* si después de imaginarme la ilusión Müller-Lyer obtengo como resultado que una línea me parece más larga que la otra. En el tercer caso, llegaría a un estado de miedo o pánico real o imaginado. En el último caso, inferiría un enunciado existencial.

No obstante, imaginarme beber medio litro de whisky no me conduce a un estado imaginado (y menos real) de estar mareado. La razón es que los mecanismos involucrados en producir la sensación de estar mareados no participan en el proceso cuando se lleva a cabo la acción de beber whisky en la imaginación. La única manera de saber qué estado se producirá es apoyándome en una teoría acerca del whisky y de sus efectos sobre los estados psicológicos típicos en los humanos. Si alguien me da a probar un líquido que no conozco, por decir algo, Erttert, y se me pregunta qué efecto producirá en un individuo *C* al beberlo, si no tengo el conocimiento teórico pertinente acerca de esa sustancia y sus efectos, difícilmente podría hacer una predicción sobre los estados psicológicos que experimentaría *C* después de su consumo. Así que, en este caso, la simulación en la imaginación no puede darse libre de teoría pues se necesita un cuerpo de conocimientos

que especifique la relación entre la ingesta de la bebida y los estados psicológicos del que bebe.

Los mecanismos encargados de la percepción visual de las flechas paralelas como de diferente tamaño en la ilusión Müller-Lyer no se activan si trato de repetir ese proceso en mi imaginación. Si pienso en dos flechas paralelas pero una con las puntas hacia dentro y otra con las puntas hacia afuera no puedo suscitar en la imaginación el mismo resultado ilusorio que se produce cuando se da la percepción actual de las líneas. Los mecanismos visuales que participan normalmente en la experiencia perceptual de la ilusión no están presentes cuando se realiza el mismo ejercicio en la imaginación.

Los dos casos anteriores parecen sugerir una conclusión pesimista para el partidario de la teoría de la simulación. No obstante, con algunas emociones parecen producirse efectos semejantes a los reales cuando se hacen simulaciones en la imaginación. Hay personas que a partir de sólo imaginarse un insecto pueden producir el mismo estado psicológico de miedo que se produciría si lo tuvieran frente a ellos. En el caso (iii) (cuando me imagino estar a punto de caer a un precipicio al estar sujetado de una cuerda y tener las manos resbalosas), Davies y Stone sostienen que se puede en realidad alcanzar estados de miedo y de pánico como los que se experimentarían cuando se está en esa situación. No sólo eso, sino que al parecer el resultado alcanzado imaginativamente no se apoya en ninguna teoría empírica acerca de las sensaciones que normalmente se producen en situaciones de riesgo como esa. A mi juicio es difícil aceptar que no se utilice algún conocimiento teórico en ocasiones como esta. Como Stich y Nichols sugieren, para que se active el proceso de simulación requerimos, en primer lugar, un conjunto de conocimientos teóricos acerca del mundo y, en segundo lugar, un conjunto de conocimientos acerca de cómo los estados del mundo afectan los estados de creencia de los seres humanos.⁸² De otro modo, sería difícil imaginar cómo se puede trasladar el sujeto en la imaginación a una situación ficticia y para simular las emociones correspondientes.

⁸² Stich y Nichols (1992), p. 140.

En el último caso que plantearon Davies y Stone, es intuitivamente plausible pensar que un proceso de razonamiento en la imaginación puede conducirnos a la misma conclusión que nos llevaría un proceso de razonamiento en la realidad. Creer que la nieve es blanca efectivamente puede hacerme creer que algo es blanco. La idea es que en los casos de razonamiento o de inferencias de este tipo no empleo una teoría de fondo que respalde mis conclusiones sino que entro al mismo proceso inferencial que se efectuaría en caso del razonamiento real. Aplico las mismas reglas inferenciales al enfrentarme a un caso real que a un caso imaginario.

Los autores también sostienen que a partir de la inferencia efectuada en la imaginación se puede obtener el enunciado de que si las premisas fueran verdaderas entonces la conclusión también lo sería.⁸³

De acuerdo con esto, parece que al menos un tipo de predicción no se apoya en teoría. La simulación guiada por proceso en el caso de la predicción de la conclusión en una inferencia no se basa en un conocimiento de fondo o en una teoría que relacione insumos con estados psicológicos y con salidas conductuales. Por tanto, aún hay espacio para la simulación en la Psicología Popular. La situación, según la concepción de Davies y Stone (1996), amerita una investigación empírica mediante la cual se pueda construir una taxonomía de las instancias de predicción que involucran una teoría de fondo y las predicciones que sólo descansan en el proceso mismo cuyo resultado se quiere predecir.

Si lo anterior es el caso, entonces es plausible suponer que el usuario de la Psicología Popular usa uno u otro mecanismo según se lo exija el contexto.

El segundo argumento tiene el propósito, nuevamente, de mostrar que individualmente las visiones que hemos venido estudiando son incompletas y que

⁸³ Davies y Stone (1996), p. 60.

integrándolas de un modo particular se puede avanzar explicativamente. Sin embargo, la conclusión del mismo será expresada en un enunciado condicional y no en uno categórico.

La TT nos puede servir para predecir y explicar las conductas de los demás pero ¿es necesaria para el caso propio? Recordando lo que Robert Gordon nos recomienda, es preciso tomar en cuenta las diferencias entre predecir la conducta propia y la de los otros, la conducta inmediata y la mediata, y finalmente, la conducta según condiciones existentes o según condiciones hipotéticas.⁸⁴ Como bien sabemos, el nivel de exactitud en cada una de estas clases de predicciones es diferente. Nuestras predicciones de conducta inmediata propia son normalmente infalibles. Introspectivamente no usamos una teoría o un conjunto de leyes psicológicas. Lo que se manifiesta es el empleo de un razonamiento práctico cuya conclusión es precedida por la emisión de la predicción. En el caso de predecir nuestra conducta en una situación hipotética ocurre algo similar. Consideramos nuestros estados mentales actuales y pensamos qué curso de acción seguiríamos si los estímulos fueran diferentes. Cuando nos preguntan “¿Qué harías si P?” simplemente respondemos a partir del caso propio, de cómo hemos reaccionado en situaciones similares y cómo han resultado las cosas a partir de ello; incluso hay ocasiones en las que no tomamos en cuenta información previa al respecto pues no la tenemos. Supongamos que alguien nos pregunta: “¿Qué harías si te informaran que tienes 1 mes de vida solamente?”, es plausible que esa situación no haya ocurrido y por ello que no tengamos antecedentes de haberla considerado. Sin embargo, podemos dar una respuesta casi automática. ¿Acaso recurrimos en estos casos de predicción a una teoría de lo mental? ¿No sería más plausible que estuviéramos empleando simplemente nuestro mecanismo de razonamiento práctico alimentado con entradas hipotéticas para predecir la conducta propia? ¿Necesitamos acaso una teoría de lo mental para saber qué vamos a hacer? Aclaro que parece ser diferente la predicción de la conducta propia y la comprensión del mundo mental propio de la predicción y comprensión de los demás como veremos con más detalle a continuación.

⁸⁴ Gordon (1986), p. 61.

Para el caso de predicción propia sería plausible, en principio, no tener que postular una teoría de lo mental como tal. Niños con síndrome de Asperger, que tienen la particularidad de no rehuir el contacto con otros seres humanos, parecen estar conscientes de los estados mentales propios pero “ciegos” ante la vida mental de los demás. La Asociación de Autismo de América (ASA) señala:

Children with autism are frequently seen as aloof and uninterested in others. This is not the case with Asperger's Disorder. Individuals with Asperger's Disorder usually want to fit in and have interaction with others; they simply don't know how to do it. They may be socially awkward, not understanding of conventional social rules, or may show a lack of empathy. They may have limited eye contact, seem to be unengaged in a conversation, and not understand the use of gestures. (Autism Society of America.)⁸⁵

La ineficacia en sus nexos sociales normalmente está acompañada de una plática egocéntrica que desemboca normalmente en la ausencia de lazos amistosos y en el consecuente rechazo por parte de los interlocutores. Se podría inferir, entonces, que la predicción de la conducta propia y la de los demás parece no ser monopolio de un mecanismo mental único. Si algunas predicciones de conducta no se efectúan exclusivamente mediante una teoría de la mente entonces debemos aceptar otro mecanismo que explique esta diversidad.

Como podemos imaginar, Robert Gordon sugeriría un mecanismo de razonamiento práctico para explicar la diversidad anterior.⁸⁶ Aún más, él extiende el empleo del

⁸⁵ Extraído de http://www.autism-society.org/site/PageServer?pagename=about_what_is_asperger. “Niños con autismo son vistos frecuentemente como distantes y desinteresados en los otros. Este no es el caso de los niños con el desorden de Asperger. Los individuos con el desorden de Asperger usualmente quieren integrarse y tener interacción con los otros; Ellos simplemente no saben como hacerlo. Pueden ser socialmente torpes, no entendiendo las reglas sociales convencionales. O bien pueden mostrar una falta de empatía. Pueden tener contacto visual limitado, parecen no estar involucrados en la conversación y no entienden el uso de gestos”.

⁸⁶ Gordon (1986).

razonamiento práctico simulado a los demás como base de la atribución de estados mentales. Sin embargo, si son ciertos los fenómenos que parecen presentarse en niños con síndrome de Asperger entonces se necesita algo más que el razonamiento práctico simulado para tener la FP. Ellos presentan un mecanismo de razonamiento práctico y bajo algunas interpretaciones están conscientes de su vida mental pero no pueden establecer el puente para considerar como semejantes a los demás.

Un buen candidato para fungir como vínculo entre las atribuciones mentales propias y las ajenas podría ser la teoría de la mente. Para extender las atribuciones propias a los demás se necesita considerar a los otros sujetos como miembros de la misma clase. Pero ¿acaso esta situación no favorecería la idea de una teoría de la mente que integrara a ambas concepciones? La respuesta, me parece, debe quedar como un enunciado condicional: si la predicción de la conducta propia no depende del todo de una teoría de la mente entonces ésta no es el único mecanismo que empleamos en el despliegue de la FP.

Sin embargo, es una cuestión abierta todavía, y sujeta a un análisis empírico, si los sujetos con síndrome de Asperger efectivamente tienen una vida mental normal en la cual explican y predicen la conducta propia empleando los conceptos de la FP. En caso de que la respuesta fuera positiva, la teoría de la simulación se anotaría una victoria pues la predicción de la conducta propia sería diferente de la predicción de la conducta ajena y esto forzaría a una integración de las dos concepciones para explicar los dos aspectos de la predicción (el propio y el ajeno). Si la respuesta fuera negativa, la necesidad de una teoría de la simulación se desvanecería pues no sería determinante como recurso explicativo de ningún fenómeno psicológico.

El tercer argumento para apoyar la integración de teorías está basado en unos estudios de Kahneman y Tversky que Goldman (1989) menciona. La conclusión del argumento es que la Teoría de la Simulación puede servir para indicar el grado de

intensidad que tiene la conducta o el estado mental resultante predicho gracias a la TT. En este sentido postular la integración contribuiría a dar cuenta de la exactitud de las predicciones y las explicaciones. La simulación sería un mecanismo que afinaría los resultados obtenidos. Cabe señalar que el propósito de Goldman en su argumento no es apoyar una visión híbrida. Sin embargo, considero plausible hacer una lectura que lo haga compatible con esta empresa.

Goldman sostiene que la Teoría de la Simulación puede dar una respuesta exacta a casos como el siguiente:

Mr Crane and Mr. Tees were scheduled to leave the airport on different flights, at the same time. They traveled from town in the same limousine, were caught in the traffic jam, and arrived at the airport 30 minutes after the scheduled departure time of their flights. Mr. Crane is told that his flight left on time. Mr. Tees is told that his was delayed, and just left five minutes ago. Who is more upset? (Goldman 1989, p. 83)

El 96% de los sujetos en el estudio respondieron que Mr. Tees estaría más molesto. Goldman sugiere que esto es fácilmente explicable para el simulacionista pues los participantes se imaginaron cómo se sentirían en los zapatos de los personajes y respondieron en concordancia.⁸⁷

La idea de Goldman a favor de la simulación es que los partidarios de la TT no podrían explicar este tipo de respuestas particulares que efectivamente se presentan en la vida cotidiana. Sin embargo, no niega la posibilidad de que la TT pueda dar una respuesta más general al estudio de Kahneman y Tversky. La TT fácilmente puede predecir qué estados mentales se producirán en los sujetos del ejemplo al no conseguir lo que desean. La exactitud en la intensidad del estado mental atribuido en esa predicción no es absolutamente necesaria para la coordinación interpersonal. Es suficiente saber qué

⁸⁷ Goldman (1989), p. 83.

estado mental se producirá en la otra persona. Si fuera necesario dar un resultado más preciso podríamos recurrir a cualquiera de las dos siguientes estrategias: recurrir a un conocimiento de situaciones semejantes o la Teoría de la Simulación. La primera estrategia nos daría una respuesta adecuada siempre que hayamos experimentado eventos similares. Si esto es el caso entonces podremos ser capaces de dar una respuesta precisa como la que solicitan Kahneman y Tversky. En caso contrario, sería preciso recurrir a una alternativa: La Teoría de la Simulación. Dado que no podemos emplear conocimiento sobre el caso en cuestión en la simulación pues por hipótesis no lo tenemos entonces sólo nos queda la opción de la simulación guiada por proceso. Si simulamos estar en la situación del otro y efectuamos el mismo proceso que ocurriría en él entonces podremos alcanzar los mismos estados mentales a los que él llega. Pero para echar a andar la simulación guiada por proceso en este punto es preciso haber realizado una primera aproximación al estado en el que puede estar por medio de nuestra teoría de la mente. ¿Cómo podríamos realizar la simulación off-line si no sabemos desde qué punto psicológico partir? Sugiero que no es posible a menos que se haga uso de la teoría de la mente para plantear la situación cuya intensidad se desea prever. Una vez que ya sabemos qué estado mental se va a producir gracias a nuestra teoría psicológica entonces podríamos modular la salida obtenida en virtud de una simulación guiada por proceso. Pero para alimentar el mecanismo de simulación es claro que necesitamos tanto una teoría de la mente como principios generales empíricos. La teoría de la mente al mostrar las relaciones causales típicas produce una primera salida general (no-específica) y los principios generales empíricos modifican o desvían la salida adecuadamente. Si llegáramos a un lugar donde existieran super-espartanos como los que Putnam (1965) imagina entonces tendríamos que tomar en cuenta los principios generales que desvían la salida conductual. Recordemos que estos sujetos de la imaginación putnamiana sienten dolores pero han eliminado por un proceso de adiestramiento cualquier salida conductual. Por ello, la observación de su conducta no constituye un factor suficiente para inferir estados mentales. Si quisiéramos predecir la conducta de dolor de un super espartano sólo basándonos en simulación, tendríamos que tomar en cuenta que él no reacciona del

mismo modo que nosotros ante los dolores. Sin esto en mente los resultados serían erróneos pues obviaríamos las diferencias relevantes. En este sentido necesitamos principios generales empíricos. Una vez que recopilamos la información sobre las interacciones causales típicas de un tipo de sujetos podemos hacer la simulación. Pero la variación de las combinaciones de estados mentales está acotada por nuestra teoría de la mente. Podemos explicar que si el super-espartano fue golpeado en el pie y no emitió queja fue porque reconoció ese estado como de dolor y entonces inhibió su salida típica. Cuando este sujeto recibe el golpe se presentan estados producto de disparos de las fibras C y esto le permita reconocer que su situación no debe desembocar en una conducta que en su sociedad se considera censurable. Pero ha aprendido a censurar tal reacción pues sabe que es fruto de tales causas típicas. Sus interacciones mentales están mediadas por una teoría de la mente.

Regresemos al papel de la simulación en el ejemplo de Goldman. Es posible afirmar que si se necesita mayor precisión en la predicción sería en principio plausible emplear un mecanismo de razonamiento práctico simulado sobre los resultados ya obtenidos a partir de la teoría de la mente y las generalizaciones empíricas registradas. Cuando tenemos un resultado para dos sujetos y necesitamos evidenciar sus diferencias entonces entramos en un proceso de simulación para determinar la diferencia. Pero sería innecesario entrar en este estado de simulación cada vez pues normalmente los resultados típicos bastan. Por tanto, la simulación sólo sería un medio de afinar las salidas predictivas y explicativas o de determinar su intensidad en caso de controversia.

El cuarto argumento, nuevamente de Goldman (1989), puede ser útil para llamar la atención sobre la necesidad de una teoría híbrida. En este caso, se concentra en la conducta verbal. Según Goldman, las personas que entran en una charla hacen suposiciones sobre la información que puede ser recuperable para su interlocutor en la discusión. Es decir, suponemos que la otra persona puede hallar evocativa una frase pues

nos resulta evocativa a nosotros. Cierta información que en el caso particular puede detonar una reacción en uno mismo la empleamos para que produzca una reacción en el otro.⁸⁸ Pero nuestra esperanza sobre su actuar es de hecho una predicción sobre su conducta. Este tipo de predicción seguramente no se basa en una teoría como tal, por tanto, la opción que queda es nuestro mecanismo de razonamiento práctico off-line. Nos ponemos en el lugar del otro y ponemos a funcionar el sistema. Así predecimos su conducta.

El lector se podría preguntar en qué sentido puede este argumento apoyar una teoría híbrida. A lo cual respondo que nuevamente en el caso de la predicción de conducta verbal suponemos, gracias a nuestra teoría de la mente, que ciertas percepciones producirán ciertos estados mentales. Estas percepciones (escuchar una frase) poseen cierto contenido (y este contenido en este caso es proposicional). Para predecir que nuestro interlocutor reaccionará ante cierta frase tenemos que suponer que las personas reaccionan ante instancias proposicionales que comprenden. Esto ya lo considera la TT previamente. La reacción exacta, sin embargo, podría ser producto de una simulación. Por tanto, al entrar en una conversación empleamos nuestra teoría de la mente pues consideramos que la otra persona formará creencias a partir de las experiencias perceptuales que reciba, conocimiento sobre regularidades empíricas pues según a la clase de sujetos que pertenezca (un círculo literario tal vez) reaccionará ante nuestra emisión, y en caso de ser necesario, se puede emplear el sistema de razonamiento práctico simulado para modular el efecto deseado o calcularlo.

El quinto argumento, extraído de Goldman (1989), se concentra en los procesos inferenciales. La idea es que los seres humanos poseemos ciertos rasgos cognitivos inferenciales inconscientes (reglas de inferencia deductiva eliminativas como la simplificación) que empleamos en el momento de extraer conclusiones. Al ser

⁸⁸ Goldman (1989), p. 84.

inconscientes el usuario no puede apelar a ellas al hacer predicciones aunque de hecho las use. Por tanto, al predecir la conducta inferencial de otra persona es necesario que utilice otro mecanismo que en este caso sería la simulación guiada por proceso. Goldman supone que la teoría de la mente no cubre estos casos de predicción inferencial. Si lo anterior es el caso, entonces existe la posibilidad de que la simulación sea empleada como mecanismo para generar predicciones en estos casos límite. No obstante, es claro que sería innecesario emplear el mecanismo de simulación cada vez que se presente una situación que involucre extraer conclusiones a partir de ciertas premisas. La razón es que la predicción puede rápidamente ser almacenada como una salida típica para inferencias y así no ejecutar nuevamente el proceso de simulación. Si la teoría de la mente no explica estos casos entonces hay lugar para que la Teoría de la Simulación complete la caracterización de la competencia psicológica. En caso de que esto ocurra tendríamos una teoría de la mente (para atribuir el resultado del caso propio al blanco de nuestra predicción), un conjunto de principios generales empíricos, y un mecanismo de razonamiento práctico guiado por proceso (que genere la conclusión a partir de las premisas).

Un último argumento proviene del artículo *The Domain of Folk Psychology* de José Luis Bermúdez (2003). Bermúdez sostiene que hay dos maneras de ver el dominio de la FP: una estrecha y una amplia. La manera estrecha consiste en limitar el uso de la FP a instancias de predicción/explicación donde el sujeto haga uso consciente y explícito de los conceptos mentalistas. La manera amplia considera que a todas las interacciones sociales subyace la suposición de que el otro posee estados mentales y que la explicación/predicción de conducta y estados mentales se produce mediante el despliegue de los conceptos de la FP.⁸⁹ Según Bermúdez el dominio de la FP es estrecho. Su conclusión se basa en una idea que parece haber pasado desapercibida para muchos filósofos: la conducta social implica coordinación con más de una persona a la vez. Ya sea

⁸⁹ Bermúdez (2003), p. 4.

que estemos jugando un juego, comiendo en un restaurante, asistiendo a una clase, o divirtiéndonos en una fiesta, nuestra conducta es coordinada con las de los otros de una forma casi perfecta. Incluso modificamos nuestras acciones cuando estamos en grupo según observamos sus reacciones. Hacemos un monitoreo constante de sus emociones a partir de sus gestos y en general de su conducta. Para Bermúdez esa coordinación grupal no la puede explicar ni la TT ni la Teoría de la Simulación. La TT no puede hacerlo pues se necesitaría una capacidad sobrehumana para reconocer simultáneamente que las conductas de todas las personas son subsumidas por tal o cual principio general. Más aún, al relacionarnos socialmente nuestra conducta depende muchas veces de la conducta y estados mentales variables de los otros. Necesitaríamos considerar qué cosas piensan los sujetos sobre sí mismos, sobre nosotros, y sobre los demás para prever y explicar sus conductas y poder relacionarnos adecuadamente esperando que sus deseos y creencias no cambien, o que no lo hagan drásticamente. Las combinaciones posibles parecen ser inmanejables. Lo mismo debe enfrentar la Teoría de la Simulación. Sería implausible sostener que empleamos un mecanismo de simulación para coordinarnos socialmente pues las combinaciones de estados mentales posibles superarían nuestras capacidades fácilmente. En una situación uno a uno parece sugerente pensar que haciendo una simulación off-line podemos interactuar con otra persona de manera adecuada. Pero cuando nos relacionamos con 20 o 30 personas simultáneamente vemos que la idea inicialmente atractiva resulta difícil de aceptar. Es poco probable que podamos simular las combinaciones de estados mentales que acaecen al mismo tiempo en 20 personas para coordinarnos socialmente.⁹⁰ La salida a esto, según Bermúdez, es recurrir a una noción ya clásica en inteligencia artificial: la noción de un marco. La mayoría de las interacciones sociales en las que nos involucramos cotidianamente están estereotipadas. Esto significa que en virtud de que se presentan patrones de conducta constantes se puede anticipar el curso de los eventos fácilmente. Esto es palpable cuando entramos a una tienda o a un restaurante. Tenemos expectativas muy claras sobre el curso de los eventos en estas circunstancias. Estas situaciones de rutina no necesitan que hagamos uso de una teoría de

⁹⁰ Bermúdez (2003), pp. 11 y 12.

la mente o de una simulación. Sólo seguimos el guión que hemos aprendido a través de nuestras experiencias previas. Estos marcos los detalla Marvin Minsky en su texto *A Framework for Representing Knowledge*. Minsky define al marco como una estructura de datos para representar una situación estereotipada.⁹¹ Entre los elementos del marco hay información acerca de cómo usarlo, qué puede esperar uno que suceda y qué hacer si no sucede eso que uno espera. Hay ciertas partes del marco que están fijas y otras que cambian. Minsky señala:

We can think of a frame as a network of nodes and relations. The top levels of a frame are fixed, and represent things that are always true about the supposed situation. The lower levels have many *terminals* – slots that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet. (The assignments themselves are usually smaller sub-frames.) Simple conditions are specified by *markers* that might require a terminal assignment to be a person, an object of sufficient value, or a pointer to a sub-frame of a certain type. More complex conditions can specify relations among the things assigned to several terminals. (Minsky 1974, pp. 111-112)⁹²

Bermúdez considera que el empleo de los marcos es más generalizado de lo que pensamos pues hay gran cantidad de actividades sociales que exigen repetición. Por lo anterior, sería raro tener que realizar simulaciones repetidas o emplear nuestra teoría de la mente cada vez si la situación es muy similar a otras. Pero ¿qué hacer cuando la

⁹¹ Minsky (1974), p. 112.

⁹² “Podemos pensar en un marco como una cadena de nodos y relaciones. Los niveles superiores del marco están fijados, y representan cosas que son siempre verdaderas de la situación supuesta. Los niveles más inferiores tienen muchas terminales – ranuras que deben ser llenadas por instancias específicas o datos. Cada terminal puede especificar condiciones que sus asignaciones deben satisfacer. (las asignaciones son usualmente sub-marcos más pequeños.) Las condiciones simples son especificadas por marcadores que podrían requerir que la asignación a una terminal sea una persona, un objeto de suficiente valor, o un indicador a un sub-marco de cierto tipo. Condiciones más complejas pueden especificar relaciones entre las cosas asignadas a varias terminales”.

situación no puede ser asimilable a un marco? Cuando no hay un marco a seguir entonces empleamos nuestra teoría de la mente para saber como operar. Por tanto, el dominio de la FP debe concebirse de manera estrecha.⁹³

Bermúdez propiamente no habla de una Teoría Híbrida. Sin embargo, a partir de lo anterior podemos concebir una teoría que incorpore estos tres elementos que hemos venido mencionando. Una Teoría de la Mente, principios generales empíricos (aquí se puede ubicar a los marcos), y por último, un componente de simulación guiada por proceso. Si el argumento de Bermúdez es correcto, entonces hay ciertas interacciones sociales que no suponen necesariamente el despliegue de la teoría de la Mente ni de la simulación. En este caso, puede decirse que la FP está formada por tres elementos que se emplean simultánea o alternadamente.

Hemos revisado seis argumentos que tenían como objetivo sugerir que la Psicología Popular no está sustentada por un mecanismo único y que para dar cuenta correctamente de su funcionamiento es una buena apuesta buscar una integración.

⁹³ Bermúdez (2003), pp. 24-26.

2. Posible Estrategia de Integración

En esta sección, quiero describir cómo podría ser el núcleo de una teoría híbrida que de cuenta de la FP. Me inclino a pensar que existe en el fondo de la FP algo como una teoría. Imagino que el simulacionista no aceptaría, *prima facie*, que esto es el caso. Gran parte del desacuerdo entre las dos concepciones principales involucradas en el debate radica en este punto. Sin embargo, pido una oportunidad para la noción de teoría que voy a esbozar y sostener.

La conducta de cualquier ser vivo animado puede ser analizada en elementos simples. Normalmente está constituida por cadenas de movimientos. Cada eslabón de la cadena puede entenderse como una palabra y toda ella como un enunciado. Como en el lenguaje natural, hay cadenas largas y cadenas cortas. Pero todas respetan algo parecido a la sintaxis. Hay cierto orden que debe seguirse para que una acción se constituya como tal y no sea simplemente un conjunto de movimientos sin razón. Si me quiero rascar la cabeza tengo que hacer las cosas en cierto orden. No tendría caso mover los dedos repetidamente antes de que mi brazo haya realizado el movimiento adecuado para que mi mano esté sobre mi cabeza. Tan sólo imaginar la escena puede producir desconcierto. La semántica de la cadena puede ser dada por el sentido de la acción. Mi rascarme la cabeza puede entenderse como el significado de mover mi brazo y mis dedos de cierto modo. En este caso (como en la mayoría), sin sintaxis cuesta trabajo encontrar el significado de la oración. Siguiendo con la analogía, muchas cadenas de movimientos diferentes pueden constituir la misma acción, o por lo menos ser entendidas como la misma acción. Entre una acción y otra hay muchas sutilezas que las diferencian entre sí, pero varias pueden formar parte de un mismo tipo de acción. Un ejemplo de esto se puede ver en el caso del “comer”. Podemos reconocer en la mayoría de los casos, y en gran parte de los seres vivos, que se están realizando la acción de comer. Parece que podemos reconocer esa serie de movimientos como instancias de un tipo especial de comportamiento diferente a

caminar o dormir. Nuevamente, esta conducta parece tener una estructura. Por lo menos tiene un inicio, un desarrollo, y un fin trazables. No tendría sentido abrir la boca, cerrarla, masticar, y finalmente acercar la cuchara con la sopa a la boca. Parece demasiado trivial decirlo pero nadie se equivoca en el orden que debe seguir cada movimiento. Ni siquiera los bebés fallan en estas tareas tan sencillas. ¿Hay acaso una “gramática” interna que dicta qué pasos seguir para tener éxito en nuestras acciones y satisfacer nuestros deseos? No parece insensato pensar que así es. Por lo menos, si quisiéramos crear un modelo computacional de la mente humana, o más simple, de cualquier mamífero, tendríamos que programar las reglas adecuadas para que este orden se cumpliera. Las formas de actuar del modelo estarían especificadas por este conjunto de principios que contendrían los posibles movimientos a realizar. La posibilidad de que fueran aprendidas por nuestro modelo no parece ser plausible. La razón es que sin instrucciones para ordenar, identificar, almacenar, y combinar la información, sería imposible integrarla de algún modo al sistema. Debe haber una regla que permita encontrar las similitudes, los patrones, las frecuencias, etc. y construir algo más a partir de los resultados obtenidos para poder recordar, inferir, agrupar, y transformar la información. Incluso puede decirse que los seres más simples tienen instrucciones implícitas sobre qué hacer frente a estímulos relevantes para su supervivencia y reproducción. Ahora bien, una arquitectura cognitiva muy económica (como la humeana) tiene que postular, al menos, un mecanismo que permite individuar objetos y eventos, hallar sus semejanzas, y crear relaciones entre ellos. Esto equivale a la tendencia asociacionista que se reconoce en ese modelo. Regresando a la analogía que habíamos trazado, si las acciones son un tipo de lenguaje, entonces debería existir algo como un vocabulario básico. Este léxico debería estar acompañado de reglas para combinar los elementos atómicos en frases. Dentro de su vocabulario estarían acciones como dormir, comer, golpear, coger algo, etc. Esto es lo que entenderé en una primera instancia como teoría.

Ahora trataré de elaborar un poco más esta noción aunque no dejará de ser un bosquejo. El modelo cognitivo que encuentro más atractivo es el que postula, en vez de

un sistema todo propósito, mecanismos diferenciables de dominio específico como componentes. Los rasgos de estos mecanismos están sujetos a debate. Unos consideran esencial que estén informacionalmente encapsulados y que su funcionamiento sea obligatorio (Fodor 1983) y otros piensan que esas características no son necesarias (Sperber 2002). Independientemente de esto, hay fenómenos que hacen plausible esta concepción. En la visión por ejemplo, encontramos fallas específicas en la integración de movimiento o en la codificación y detección del color cuando percibimos algo pero un funcionamiento adecuado al representar las propiedades restantes de los elementos de la escena (tamaños, matices, etc.). En el lenguaje, por otro lado, podemos encontrar fenómenos como la afasia de Broca, desorden lingüístico que impide la construcción de frases bien articuladas. También hay fenómenos como la prosopagnosia que consiste en la imposibilidad de reconocer los rostros a pesar de poder identificar sus partes. Los individuos con este desorden no son capaces de integrar los diversos elementos de un rostro para crear una representación identificable. Sperber (2002) considera que la capacidad de reconocer los rostros es posible gracias un módulo. Podemos comprometernos con la modularidad de la mente o no, y si lo hacemos, con alguna noción de módulo particular, pero lo que es difícil de ignorar es que existe una división cognitiva.

En lo que respecta a la teoría que tengo en mente, debería estar constituida por:

1. Una serie de mecanismos que establezcan relaciones jerárquicas entre movimientos y que operen con
2. Algoritmos para construir cadenas alternativas de movimientos.
3. Un sistema de reconocimiento de agencia que al detectar coincidencia con algunas de las cadenas construibles dentro del sistema identifique al objeto como un agente.

4. Conexión con los resultados del procesamiento visual.

5. Conexión con sistemas cognitivos superiores.

El empleo de la Psicología Popular por cualquier usuario normal, supone la posesión de todos estos elementos como partes de su sistema cognitivo. El autista presenta igualmente todos ellos, menos el 3. Este mecanismo tendría que estar relacionado con el proceso realizado por las neuronas espejo.⁹⁴ Las neuronas espejo son células nerviosas encargadas de identificar y replicar conductas a partir de la información obtenida por la percepción visual. Sus patrones de activación son sorprendentemente semejantes cuando se realiza una acción y cuando sólo es observada. En ambos casos hay actividad en la región premotora F5. Rizzolatti sugiere que este grupo de neuronas posibilitan la comprensión del movimiento como acción.⁹⁵ Dejaré la parte neurofisiológica de lado pues excedería los propósitos de este trabajo hacer una descripción explícita de todos los detalles involucrados en este tema. Lo que es importante notar es que se puede confiar en que hay un mecanismo con base neuronal involucrado en el reconocimiento de conductas y que no está presente en sujetos con autismo.⁹⁶ Esto sugiere que la capacidad de mentalizar tiene un componente identificable cerebralmente. Considero que este mecanismo (identificado con el número tres en mi lista) más los otros elementos mencionados del 1-5 pueden entenderse como una teoría.

Tanto los partidarios de la TT y como los teóricos de la simulación, aceptarían que existe algo como lo que acabo de proponer. En este punto creo que se puede empezar a construir lo que puede llamarse una teoría híbrida. El simulacionista está comprometido con un mecanismo de reconocimiento y replicación de acciones. El partidario de la TT aceptaría un mecanismo constituido por algoritmos vocabularios y criterios de identificación y construcción de conductas. La noción tentativa que he propuesto puede

⁹⁴ Rizzolatti, et al. (1996).

⁹⁵ Rizzolatti, et. al. (2006).

⁹⁶ Con algunos tipos de autismo.

cumplir con ambos requisitos. Incluso la simulación guiada por proceso implicaría una teoría en este sentido que estoy usando.

3. Conclusión

La conclusión de este trabajo es muy breve. La pregunta que se trató de responder en el texto fue ¿cuál es la base de la Psicología Popular? La respuesta se articuló a partir de dos posturas antagónicas: La Teoría-Teoría y la Teoría de la Simulación. La primera enfrenta problemas cuando hay que hacer predicciones muy detalladas. El caso presentado sobre los grados de intensidad de una emoción parece ser difícilmente manejado si no hay simulación guiada por proceso. Otro punto interesante se presenta cuando uno quiere predecir sus acciones. No es claro fenoménicamente que se esté recurriendo a una teoría. La Teoría de la Simulación enfrenta un problema grave al predecir y explicar conductas cuando se enfrenta a varios agentes a la vez. El simulacionista pretende que el mecanismo que postula replique los procesos cognitivos que se efectúan en el agente. Además argumenta que al emplear un mecanismo en lugar de una teoría, su concepción es más simple. Sin embargo, no es más simple tener que correr una docena de simulaciones a la vez cuando estamos en una charla grupal, en un juego de equipo o simplemente conduciendo nuestro automóvil en las calles de la ciudad. Si tuviéramos que replicar cada proceso que se realiza en la mente de los otros y manejar todas las variables involucradas en esto, dudo que pudiéramos relacionarnos socialmente.

Los argumentos esgrimidos nos llevaron a descubrir lo intrincado de los mecanismos empleados en el uso de la Psicología de creencias y deseos. Una respuesta unilateral parecía dejar de lado ciertos aspectos importantes y por ello no fue posible pronunciarse a favor de una u otra. Concluí que una visión integradora podría arrojar más luz sobre la Psicología Popular. Esto implicaría que no se emplea un único mecanismo para mentalizar pues la FP está constituida por varios componentes que sirven como recursos optimizadores de su aplicación. Los argumentos presentados para llegar a esta conclusión, mostraron que hay ciertas particularidades en la atribución de conductas y de estados mentales que resultan problemáticas si no se incorpora a distintos sistemas en la base de la ejecución psicológica. Estos tres elementos son: la Teoría de la Mente, la

Simulación Guiada por Proceso, y principios generales empíricos (en los que se insertan los marcos). Juntos constituyen una posible Teoría Híbrida.

Todas las interacciones sociales presuponen el reconocimiento del otro como un agente intencional. A partir de esto, se puede hacer la atribución. Cuando esto exige mayor precisión se podría recurrir a la simulación guiada por proceso. Por último, cuando la situación es demasiado estable o involucra a demasiadas personas se suele pasar a un nivel donde se categoriza la situación como de un tipo específico. De este modo, se puede manejar mediante marcos la explicación y predicción de la conducta de los otros seres humanos e incluso de uno mismo.

Referencias

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the Autistic Child have a "Theory of Mind"? *Cognition*, 21, 37-46.

Bermúdez, J. L. (2003) "The Domain of Folk Psychology", <http://www.stir.ac.uk/departments/arts/philosophy/cnw/webpapers/jose17.htm>, y en O'Hear, A. (2003), *Mind and Persons*, Cambridge: Cambridge University Press.

Block, N. & Fodor, J. (1972). "What Psychological States Are Not". *Philosophical Review*, 81.

Block, N. (ed.) (1980). *Readings in Philosophy of Psychology*. Cambridge, MA: Harvard University Press.

Botterill G. (1996) "Folk Psychology and Theoretical Status" en Carruthers & Smith (eds.) (1996), *Theories of Theories of Mind*, Cambridge: Cambridge University Press.

Botterill G. & Carruthers P. (1999), *The Philosophy of Psychology*, Cambridge: Cambridge University Press.

Carruthers P. & Smith P. (eds.) (1996), *Theories of Theories of Mind*, Cambridge: Cambridge University Press.

Chomsky N. (1980), *Rules and Representations*. New York: Columbia University Press y Oxford: Basil Blackwell Publisher.

Churchland, P.M. (1988). *Matter and Consciousness*, Cambridge, MA: MIT Press.

Davidson D. (1994), *Ensayos sobre acciones y sucesos*, Barcelona/México: Crítica/UNAM.

Davidson, D., (1970), "Mental Events", en Lawrence Foster and J. W. Swanson (eds.), *Experience and Theory*, London: Duckworth.

Davidson, D. (1971) "Agency", en Robert Binkley, Richard Bronaugh, & Ausonia Marras (eds.), (1977) *Agent, Action, and Reason*, Toronto: University of Toronto Press.

Davies, M. & Stone, T. (1995), *Folk Psychology, The Theory of Mind Debate*. Oxford: Blackwell Publishers.

Davies, M. & T. Stone. (1996). "The Mental Simulation Debate: a Progress Report". en Peter Carruthers & Peter K. Smith (eds.), *Theories of theories of mind*.

Dretske, F. (1969), *Seeing and Knowing*, Chicago: The University of Chicago Press.

Dretske, F. (1981), *Knowledge and the Flow of Information*, Cambridge, Mass.: The MIT Press.

Eraña, L. (2006), "Las Teorías Intuitivas son Módulos?", *Revista de Filosofía de la Universidad de Costa Rica*, vol. XLIV, número 111.

Fodor J. (1968), "The Appeal to Tacit Knowledge in Psychological Explanation", *Journal of Philosophy*, 65, 627-640.

Fodor J. (1987), *Psychosemantics*. Cambridge, MA: MIT Press

Fodor J. & Lepore E. (1992), *Holism: a Shopper's Guide*, Oxford: Blackwell Publishers.

- Fodor J. (2000), *The Mind Doesn't Work That Way; The Scope and Limits of Computational Psychology*, Cambridge, MA: MIT Press
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593-609.
- Goldman, A., 1989, "Interpretation Psychologized." *Mind and Language* 4, 161-185; Reimpreso en Davies M. & Stone T. (eds.), (1995), *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell Publishers.
- Gopnik, A., (1996), "The Scientist as Child", *Philosophy of Science*, 63, 485-514.
- Gordon, R., 1986, "Folk Psychology as Simulation", *Mind and Language* 1, 158-171; Reimpreso en Davies, M. & Stone T., (eds.), (1995), *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell Publishers.
- Heider, F., & Simmel, M. (1944), "An Experimental Study of Apparent Behavior" en *The American Journal of Psychology*, Vol. 57, No. 2, 243-259.
- Lewis D. (1972), "Psychophysical and Theoretical Identifications", en Block (1980), ed., *Readings in Philosophy of Psychology*. Cambridge, MA: Harvard University Press.
- Lewis, D., (1974), "Causation", *The Journal of Philosophy*, Vol. 70, No. 17, 556-567.
- Minsky, M. (1974), "A Framework for Representing Knowledge", en Haugeland J. comp. (1997), *Mind Design II*, Cambridge, MA: The MIT Press.
- Nichols, S., Stich, S., Leslie, A., & Klein, D. (1996), "The Varieties of Off-Line Simulation", en Carruthers & Smith (eds.) (1996), *Theories of Theories of Mind*, Cambridge: University Press.

Premack, D., & Woodruff, G., (1978). Does the Chimpanzee have a 'Theory of Mind'? *Behavioral and Brain Sciences*, 1, 516-526.

Putnam, Hilary. (1967) "The Nature of Mental States.", en *Art, Mind and Religion*. Editado por W.H. Capitan & D.D.Merrill. Pittsburgh: University of Pittsburgh Press. Reimpreso en Block (1980).

Ravenscroft, I. (2004) "Folk Psychology as a Theory", *The Stanford Encyclopedia of Philosophy* (Winter 2004 Edition, Feb.-23-2004), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2003/entries/folkpsych-theory/>

Ravenscroft, I. (2005) *Philosophy of Mind: A Beginner's Guide*, Oxford: Oxford University Press.

Rizzolatti, G., & Sinigaglia, C. (2006), *Las Neuronas Espejo: Los Mecanismos de la Empatía Emocional*, Barcelona: Paidós.

Segal, G. (1996). "The Modularity of Theory of Mind". En Carruthers, & Smith (eds.) (1996), *Theories of Theories of Mind*, Cambridge: Cambridge University Press.

Sperber D. (2002), "Modularity and Relevance: How can a Massively Modular Mind be Flexible and Context-Sensitive?" En Carruthers P., Laurence S. & Stich S. (eds.) (2005) *The Innate Mind: Structure and Content*. Oxford: Oxford University Press.

Stich, S. & Nichols, S. (1992): "Folk Psychology: Simulation or Tacit Theory?". *Mind and Language* 7, 35-71.

Stich S. (1983), *From Folk Psychology to Cognitive Science*. Cambridge, MA: MIT Press.

Tye, M. (2000), *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103-128.