

Lectura de la mente en monos tití: respuestas al problema lógico a partir de evidencia empírica reciente

Pamela Barone

Centro de Investigaciones de la Facultad de Filosofía y Humanidades, Universidad Nacional de Córdoba.

Abstract

Partiendo desde una concepción naturalista de la filosofía, que articula desarrollos empíricos y teóricos, el objetivo de este trabajo consiste en analizar un protocolo experimental reciente sobre atribución intencional en monos tití para, a partir de allí, construir un argumento en contra de Povinelli, quien ha planteado el principal desafío para la investigación en lectura de la mente en animales no humanos. Argumentaré que un defensor de lectura de la conducta no podrá explicar el desempeño de los primates en dicho experimento y que sólo adjudicando algunas competencias mentalistas a los mismos, podremos dar cuenta de los datos.

El problema de Povinelli

Los experimentos de Premack y Woodruff (1978) dieron origen a un gran debate en el ámbito de la cognición social animal –campo interdisciplinario en el cual participan psicólogos, etólogos, filósofos, etc.– al cuestionar si la capacidad de los adultos humanos para comprender los estados mentales de otras personas se extiende a criaturas no humanas. Esta habilidad se conoce en la literatura como lectura de la mente.

Una posición en este debate la representan algunos autores que afirman, basándose en observaciones etológicas y estudios de laboratorio, que los primates no humanos cuentan con algunos rasgos de esta competencia para leer mentes (Byrne & Whiten, 1988; Cheney & Seyfarth, 1990; Hare et al., 2000, Call & Tomasello, 2008). Por otra parte, un segundo grupo de investigadores liderado por Povinelli ha planteado en varias oportunidades que los primates no humanos son incapaces de atribuir estados mentales a los demás y que todos los protocolos que intentan demostrar lo contrario sufren de un mismo problema metodológico que invalida sus resultados (Povinelli & Vonk, 2004; Vonk & Povinelli,

2006). Éste se denomina el *problema lógico* para la investigación en lectura de la mente en animales no humanos.

Povinelli parte de señalar que, en tanto nuestras mentes no son observables de manera directa, la atribución de estados mentales debe basarse en los rasgos observables de la conducta y del entorno del agente interpretado. Debido a que los enfoques experimentales dan por sentado que el animal razona e infiere la existencia de estados psicológicos apoyándose en los rasgos comportamentales observables del agente, siempre podremos ofrecer una explicación alternativa de los datos considerando que el animal en cuestión sólo es un lector de la conducta (Povinelli & Vonk, 2004). En este sentido, el animal razonaría de acuerdo a una ‘regla conductual’ según la cual ciertos rasgos observables en la conducta del agente o en su entorno llevan directamente a un comportamiento específico, sin necesidad de realizar el paso intermedio que implica inferir un estado mental (Lurz y Krachun, 2011).

Ningún protocolo experimental en boga que evalúe atribución mental en primates no humanos supera este problema: en cada estudio, podemos ofrecer una hipótesis complementaria en términos de lectura de la conducta que da cuenta de manera ajustada del comportamiento de los animales.

Evidencia empírica sobre atribución de objetivos en monos tití

En un experimento reciente, Burkart *et al.* (2012) evaluaron si los monos tití en una tarea de cognición social comprendían el comportamiento de otros agentes según la estructura intencional subyacente de sus acciones. Para ello, adoptaron el paradigma de violación de la expectativa originalmente desarrollado por Woodward para infantes humanos. Según argumentaré, pienso que este experimento está en mejores condiciones que otros protocolos para responder al problema lógico.

En la fase de habituación del estudio, los monos observaban un videoclip en el que un agente se acercaba a uno de dos objetos equidistante (éste era el *objeto-meta*). Luego, en la fase de prueba, se revertía la posición de los objetos y había dos condiciones. En la condición de violación del objetivo, el agente realizaba los mismos movimientos siguiendo la misma trayectoria que en la fase de habituación, alcanzando ahora el *objeto-no-meta*. En

la condición de no-violación del objetivo, en cambio, el agente se acercaba al *objeto-meta* realizando una nueva trayectoria (figura 1).

Los monos se repartían en tres grupos experimentales, que se diferenciaban entre sí por el tipo de agente que realizaba la acción: un congénere, un robot o una caja negra. Burkart *et al.* (2012) evaluaban los tiempos de mirada de los monos, los cuales indicaban las reacciones diferenciales de los sujetos a la novedad que se presentaba en las dos condiciones de prueba.



Figura 1. Agente-congénere en la fase de habituación y en las dos condiciones de la fase de prueba: violación del objetivo y no-violación del objetivo (extraída de Burkart et al., 2012)

Como resultado, los monos no reaccionaron del mismo modo a los tres agentes. Por un lado, les resultaba novedoso y miraban por mayor cantidad de tiempo, cuando el congénere y el robot realizaban una trayectoria inconsistente con su objetivo. Sin embargo, no les resultaba novedoso que este comportamiento lo realizara la caja negra, indicando que los sujetos codificaron los movimientos de la caja según la trayectoria realizada.

Ahora bien, ¿podría un sujeto, que sólo maneja reglas conductuales, responder de esta manera? En consecuencia, ¿podrían los críticos, como Povinelli y Lurz, afirmar que el experimento no permite distinguir entre un sujeto que es un lector de la mente de un mero lector de la conducta?

Explicación de los resultados según la hipótesis de lectura de la conducta

Veamos ahora qué tipo de reglas conductuales podrían formar los monos en el experimento de Burkart si ellos fueran meros lectores de la conducta. Por lo que alcanzo a vislumbrar, hay sólo dos opciones y ninguna de ellas nos permite explicar los resultados obtenidos.

Una primera alternativa es que la regla conductual contenga los movimientos específicos realizados por el agente. Una regla así indicaría que si el cuerpo del agente se dirige hacia un cierto objeto y comienza a realizar ciertos movimientos específicos M_1 en una trayectoria determinada, probablemente el agente seguirá esa trayectoria particular. Al formar esta regla, los monos codificarían la conducta a nivel de los movimientos específicos del agente como una conducta que le lleva a realizar cierta ruta.

Pero, ¿cómo se comportaría un mono tití que forma dicha regla en el experimento de Burkart? En la condición de violación del objetivo no se sorprenderá del comportamiento del agente debido a que éste realiza los mismos movimientos específicos que codificó en la fase de habituación. En cambio, en la condición de no-violación del objetivo, cuando el sujeto vea que el cuerpo del agente se dirige hacia un objeto pero comienza a realizar otros movimientos específicos M_2 , recorriendo para ello una nueva trayectoria, se sorprenderá y mirará por más tiempo.

Como hemos visto, los monos no reaccionan de modo uniforme para cada entidad dentro de una condición.

Una segunda alternativa, entonces, consiste en que el lector de la conducta especifique los componentes de la acción de un modo ‘más grueso’ sin considerar, como antes, cada movimiento fino específico que realiza una criatura. En este caso, la regla conductual indicaría que si el cuerpo de un agente se dirige hacia un objeto determinado y comienza a realizar una serie de movimientos dirigidos a ese objetivo, es probable que el agente alcance el objeto al final de la trayectoria. En este caso, el mono codificaría la acción del agente en función del resultado final hacia el cual se dirigen sus movimientos¹.

Ahora, ¿cómo se comportaría un mono tití con esta regla en el experimento de Burkart? En la condición de no-violación del objetivo, el sujeto esperaría que el agente realice una nueva trayectoria para alcanzar el mismo resultado final y, por lo tanto, no debiera sorprenderse. De modo opuesto, en la condición de violación del objetivo, como el movimiento del agente se mantiene igual a la fase de habituación, el sujeto miraría por un período de tiempo más largo ya que alcanza otro resultado final.

¹ Este tipo de regla concuerda con la ‘postura teleológica’ de Gergely y Csibra (2003). De acuerdo con su propuesta, los infantes humanos pueden representar, explicar y predecir acciones dirigidas a un objetivo sin atribuir estados mentales intencionales a la mente del actor.

En las dos alternativas planteadas, la regla conductual debe formarse en la fase de habituación en el momento en que los monos son expuestos a los agentes y, en cualquier caso, debería ser la misma para los tres agentes ya que el mono se percata de los mismos rasgos comportamentales y ambientales en los tres grupos². Entonces, si el sujeto estuviese siguiendo una regla conductual –ya sea la que codifica los movimientos específicos M_I del agente o bien su acción dirigida hacia un objetivo externo– tendría que reaccionar de manera uniforme para todas las entidades interpretadas dentro de una misma condición de prueba. Con esto me refiero a que no habría posibilidad para que el mono responda de modo diferenciado a cada agente potencial; es decir, reaccionar de una manera cuando se trate de un congénere, y de otra distinta cuando sea el caso del robot o de la caja negra por ejemplo. Cualquiera de las dos reglas conductuales le llevaría al sujeto a reaccionar por igual frente a todos los agentes dentro de una misma condición.

Por consiguiente, un lector de la conducta realizaría predicciones uniformes: esperaría que los tres agentes se comportaran de la misma manera, ya sea de acuerdo al contenido de una regla (la topografía del movimiento) o de la otra (la acción dirigida a un objetivo). Sin embargo, ninguna de las dos alternativas propuestas es consistente con los resultados obtenidos en el estudio de Burkart.

La atribución intencional como explicación de los datos

Una tercera posibilidad, aquella que querría defender en lo que sigue, consiste en que los lectores tengan diferentes reglas cuyos contenidos varíen dependiendo de cómo han categorizado al agente en cada caso. Esto significa que los monos títi no representan todas las entidades del mismo modo, sino que esperan algunos tipos de comportamientos y no otros en función del tipo de agente que se trate.

Entonces, podríamos sugerir que el sujeto primero reconoce que una criatura es un agente intencional y luego, espera que éste se comporte flexiblemente de acuerdo con su objetivo y que realice la acción apropiada para lograrlo. En este caso, el contenido de la regla incluye el objetivo interno del agente. Pero en esta oportunidad, el objetivo es del agente y no el objetivo de la acción, ni del entorno.

² De esto podemos estar seguros debido a que se proyectan videoclips de los tres tipos de agentes, lo cual garantiza que los movimientos están estandarizados.

Por otra parte, si las entidades no fueran representadas como agentes intencionales y por eso los monos no les atribuyeran objetivos internos, el contenido de la regla variaría. En esta situación alternativa, el lector codificaría mediante una regla conductual los movimientos específicos no-intencionales del agente.

En consecuencia, una única regla no le sería suficiente al sujeto para predecir el comportamiento de los tres agentes. Por el contrario, el contenido de la regla variaría dependiendo de cómo fue representado cada agente. Esta alternativa es la mejor explicación del comportamiento de los monos tíí ya que permite dar cuenta de las reacciones variadas de los lectores ante los tres agentes al considerar que:

- i) primero, los sujetos diferencian entre agentes intencionales y no intencionales; y
- ii) luego, construyen reglas diferenciadas: una de ellas es una regla mental y se aplica sólo para atribuirle objetivos internos a los agentes intencionales.

Como contrapartida a la sugerencia de Povinelli, los monos estarían realizando un paso cognitivo adicional al rastreo de indicadores observables y la predicción subsiguiente de la acción. Los datos permiten apoyar la conclusión de que los sujetos captan los mismos rasgos observables de la conducta y el ambiente, pero no realizan las mismas predicciones para todos los agentes porque las criaturas tienen expectativas distintas dependiendo del tipo de agente con el cual interactúan.

Ahora bien, cabe preguntarse de qué depende que algunas criaturas sean interpretadas como agentes intencionales (el congénere y el robot) mientras otras no estén englobadas en esta categoría (como la caja negra). A partir de los resultados del experimento, podemos sugerir que a fin de que los monos consideren que una entidad es un agente intencional –y por ello, puedan atribuirle objetivos–, dicha entidad debería presentar algún grado básico de similitud a un mono, como tener extremidades y una cabeza (Kupferberg *et al.*, 2013). Ni la presencia de ojos ni el movimiento autopropulsado parecen ser por sí mismos requisitos necesarios y suficientes para considerar al agente como intencional.

Una característica que sí presentan los agentes que se comportan siguiendo un objetivo interno es flexibilidad conductual. En consecuencia, si un agente posee un

propósito, el sujeto esperará que éste se comporte flexiblemente, ajustando su conducta a los cambios ambientales de un modo que le permita alcanzar su objetivo.

En el experimento de Burkart, se presenta un cambio ambiental durante las condiciones de prueba: se invierten de lugar los objetos. En futuros diseños experimentales podríamos introducir otras variaciones ambientales que impliquen que el agente no alcance su objetivo por diversas razones: porque se ve impedido por constricciones del entorno (por ejemplo, se interpone un obstáculo en su camino), porque aparecen demandas en el ambiente que le exigen postergar la satisfacción del objetivo, porque sufre un accidente y en el primer intento por cumplir su objetivo no lo logra, etc.

En este tipo de situaciones, sólo una criatura que represente el objetivo específico del agente intencional podrá reaccionar de la manera adecuada, esperando y entendiendo que un agente que por diversos impedimentos actuales no puede lograr su objetivo, seguirá realizando acciones diversas, guiadas por ése propósito, hasta cumplir con el mismo. Adicionando estas condiciones de evaluación, obtendríamos evidencia convergente a favor de la afirmación de que un animal atribuye objetivos.

Consideraciones finales

El problema de Povinelli ha impactado profundamente en la investigación sobre atribución mental en primates no humanos. Buscando superar este desafío, se han propuesto diversas estrategias que otorgan a algunos animales no humanos ciertas competencias mentalistas. Una salida argumentativa consiste en evaluar comparativamente la hipótesis de lectura de la conducta y la hipótesis de lectura de la mente según ciertos criterios epistémicos y/o metodológicos (Fitzpatrick, 2009; Fletcher & Carruthers, 2012). Otros autores, optan por idear experimentos que sí sean capaces de superar el problema lógico (Lurz & Krachun, 2011; Lurz, 2011).

Aquí, me he concentrado en una alternativa distinta: analizar un protocolo experimental reciente, que no ha sido aún considerado en los debates en torno al problema lógico, y que sirve como evidencia empírica a favor de la tesis de que los primates no humanos poseen algunas habilidades para leer mentes ajenas.

Según he argumentado, quien intente explicar el estudio de Burkart *et al.* (2012) caracterizando a los monos tití como lectores de la conducta, no podrá dar cuenta de la

formación de las reglas diferenciadas que deberían poseer estos sujetos a fin de tratar, por un lado, con el congénere y el robot al comprender sus acciones según el objetivo del agente; y por el otro, con la caja negra al entender su patrón comportamental en función de los movimientos específicos que realiza.

Sólo esta diferenciación, y la posterior formación de reglas diferentes les permiten a los lectores anticipar diferencialmente los eventos del estudio según el tipo de criatura que las realiza. Una de las reglas anteriores ya no es una regla conductual sino que se forma luego de discriminar entre aquellos sujetos que pueden considerarse como agentes intencionales y entre el tipo de criaturas que no se las puede considerar de ese modo. Lo cual resulta indispensable para entender las acciones del primero según sus objetivos y las acciones del segundo en función de los movimientos específicos realizados.

Referencias bibliográficas

- Burkart, J.; Kupferberg, A.; Glasauer, S.; & van Schaik, C. (2012), "Even Simple Forms of Social Learning Rely on Intention Attribution in Marmoset Monkeys (*Callithrix jacchus*)", *Journal of Comparative Psychology*, vol. 126, No. 2: 129–138.
- Byrne, R. & Whiten, A. (1988), *Machiavellian intelligence: Social expertise and evolution of intellect in monkeys, apes, and humans*, Oxford: Clarendon Press.
- Call, J. & Tomasello, M. (2008), "Does the chimpanzee have a theory of mind? 30 years later", *Trends in Cognitive Sciences*, Vol.12, No.5: 187-192.
- Cheney, D. & Seyfarth, R. (1990), *How monkeys see the world*, Chicago: The Chicago University Press.
- Fitzpatrick, S. (2009), "The primate mindreading controversy: a case study in simplicity and methodology in animal psychology", en R. Lurz (ed.), *The philosophy of animal minds*, Cambridge: Cambridge University Press, pp. 258–276.
- Fletcher, L. & Carruthers, P. (2012), "Behavior-reading versus mentalizing in animals", en J. Metcalfe and H. Terrace (eds.), *Agency and Joint Attention*, New York: Oxford University Press, pp. 82-99.
- Gergely, G. & Csibra, G. (2003), "Teleological reasoning in infancy: The naïve theory of rational action", *Trends in Cognitive Sciences*, 7: 287–292.
- Hare, B., Call, J., Agnetta, B. & Tomasello, M. (2000), "Chimpanzees know what conspecifics do and do not see", *Animal Behaviour*, 59: 771–785.
- Kupferberg, A., Glasauer, S., & Burkart, J. M. (2013), "Do robots have goals? How agent cues influence action understanding in non-human primates", *Behavioural Brain Research*, 246: 47-54.
- Lurz, R. & Krachun, C. (2011), "How Could We Know Whether Nonhuman Primates Understand Others' Internal Goals and Intentions? Solving Povinelli's Problem", *Review of Philosophy and Psychology*, v. 2, n°3: 449-481.
- Lurz, R. W. (2011), *Mindreading animals: the debate over what animals know about other minds*, MIT Press.
- Povinelli, D. & Vonk, J. (2004), "We Don't Need a Microscope to Explore the Chimpanzee's Mind", *Mind & Language*, Vol. 19, No. 1: 1–28.
- Premack, D. & Woodruff, G. (1978), "Does the chimpanzee have a theory of mind?", *The Behavioral and Brain Sciences*, 4: 515-526.
- Vonk, J., & Povinelli, D. J. (2006), "Similarity and difference in the conceptual systems of primates: The unobservability hypothesis", en T. Zentall & E. Wasserman (eds.), *Comparative cognition: Experimental explorations of animal intelligence*, New York: Oxford University Press, pp. 363-87.