

## Los condicionales de utilidad y la teoría de los modelos mentales.

Miguel López-Astorga<sup>1</sup>  
([milopez@utalca.cl](mailto:milopez@utalca.cl))

Recibido: 14/05/2017

Aceptado: 22/06/2017

DOI: 10.5281/zenodo.833787

### **Resumen:**

Recientemente, se ha propuesto que la lógica no es el único factor que incide en el razonamiento condicional y que el valor y la utilidad son otros dos elementos importantes en este sentido. Al respecto, Bonnefon y Sloman añadieron en 2013 que la estructura de los condicionales es, del mismo modo, relevante en la actividad inferencial con condicionales de utilidad y trataron de demostrarlo por medio de varios experimentos. Sus resultados parecieron mostrar que, efectivamente, sus tesis eran correctas. No obstante, plantearon igualmente que su propuesta y, por tanto, sus resultados eran compatibles con una teoría contemporánea de amplio alcance, la teoría de los modelos mentales, y en este artículo se trata de desarrollar esta idea con el propósito de comprobar si es cierta.

**Palabras clave:** Condicional - Modelos Mentales - Posibilidades Semánticas – Razonamiento – Utilidad.

### **Abstract:**

Recently, it has been stated that logic is not the only factor that has an influence on conditional reasoning and that value and utility are other two important elements in this regard. In this way, Bonnefon and Sloman added in 2013 that the structure of conditionals is also relevant for the inferential activity with utility conditionals, and tried to prove it by means of several experiments. Their results seemed to show that, indeed, their theses were correct. However, They also claimed that their proposal and, therefore, their results were compatible with a contemporary wide scope theory, the mental models theory, and this paper is intended to develop this idea in order to check if it is right.

**Keywords:** Conditional - Mental Models - Semantic Possibilities – Reasoning – Utility.

---

<sup>1</sup> Doctor en Lógica y Filosofía de la Ciencia, Universidad de Cádiz, España. Profesor Asociado en Instituto de Estudios Humanísticos “Juan Ignacio Molina”, Universidad de Talca, Chile.

## Introducción

En el momento presente, se puede decir que una teoría ampliamente apoyada ofrece explicaciones interesantes acerca de los procesos de razonamiento humano. Esta teoría es conocida como la *Teoría de los modelos mentales* (p.ej., Hinterecker, Knauff y Johnson-Laird, 2016; Johnson-Laird, 2012, 2015; Johnson-Laird, Girotto y Legrenzi, 1999; Johnson-Laird, Khemlani y Goodwin, 2015; Orenes y Johnson-Laird, 2012; Quelhas y Johnson-Laird, 2017; Ragni, Sonntag y Johnson-Laird, 2016) y, en adelante, nos vamos a referir a ella utilizando las siglas MMT (procedentes de la expresión *Mental Models Theory*). MMT ha demostrado contar con un gran potencial para, por ejemplo, explicar diferentes problemas relativos al razonamiento condicional -entre ellos, las dificultades vinculadas con la famosa tarea de selección de las cuatro tarjetas de Peter Wason (Wason, 1966, 1968)- y parece ser una alternativa bastante fuerte al enfoque tradicional, que considera que la mente humana opera siguiendo reglas lógico-formales, que deriva de lo que en la lógica clásica se denomina *cálculo de deducción natural* (véase, además de Gentzen, 1934, 1935, p.ej., Deaño, 1999) y al que, quizás, podemos adscribir autores como, por citar algunos, Henlé (1962) o Rips (1994).

Por otra parte, recientemente, sobre la base de la idea de que la lógica no es el único elemento relevante que guía los procesos inferenciales, se ha considerado que la utilidad y el valor inciden igualmente en el razonamiento humano. Estas tesis han sido adoptadas en trabajos como el de Bonnefon y Sloman (2013), quienes, de la misma manera, piensan que, cuando los individuos razonan acerca de condicionales de utilidad, la estructura de tales condicionales ejerce también su influencia. A su juicio, dos tipos de condicionales son importantes a este respecto: los condicionales causales (*causal conditionals*) y los condicionales diagnósticos (*diagnostic conditionals*). Con el propósito de probar estos supuestos, Bonnefon y Sloman (2013) realizan diferentes experimentos y se puede decir que los resultados que obtienen parecen indicar que tales supuestos son correctos.

Sin embargo, el aspecto de su investigación que nos interesa especialmente en este trabajo es que Bonnefon y Sloman (2013) parecen suponer igualmente que ni sus planteamientos ni los resultados que obtienen con sus experimentos son necesariamente incompatibles con el margo general de MMT. A su juicio, lo único que se precisa para

lograr una clara compatibilidad entre las dos propuestas es que esta última teoría asuma dos axiomas introducidos por ellos y que van a ser indicados más abajo.

El problema es que esta idea solo está señalada en el estudio de Bonnefon y Sloman y, así, el objetivo principal de estas páginas es verificar si es correcta o no, esto es, si es cierto o no que no hay nada ni en su enfoque global ni en sus resultados que sea incoherente con las tesis básicas de MMT. De este modo, para proceder en tal dirección, primero, se revisarán detenidamente los supuestos teóricos y las características de los experimentos de Bonnefon y Sloman (2013) para, a continuación, en un segundo apartado, analizar la posibilidad de que MMT pueda realmente ofrecer también, sin mayores suposiciones adicionales que las indicadas, una explicación de tales resultados. Comenzamos, pues, con la investigación realizada por Bonnefon y Sloman (2013).

### Condicionales causales y condicionales diagnósticos

Bonnefon y Sloman (2013) asumen la tesis de que, si el antecedente o el consecuente (o ambos) de un enunciado condicional son útiles o tienen algún valor para un individuo, es posible predecir las conclusiones que ese individuo puede extraer a partir de dicho enunciado. Esto es fácil de comprender si pensamos en un ejemplo. Según Bonnefon y Sloman (2013), el enunciado “If I eat oysters, I become very sick”<sup>2</sup> (Bonnefon y Sloman, 2013, p. 194) nos lleva a pensar que el individuo que lo afirma no va a comer ostras. De esta manera, basándose en argumentos planteados en Bonnefon (2009), sostienen que puede ser oportuno aceptar ciertos axiomas ‘populares’ de decisión. En concreto, en su opinión, hay dos axiomas populares que son especialmente relevantes:

“**Folk Axiom 1** (Self-Interested behavior). Agents take actions that increase their own personal utility (and they do not take actions that decrease their own personal utility)”<sup>3</sup> (Bonnefon y Sloman, 2013, p. 195, negritas en el texto original).

---

<sup>2</sup> ‘Si como ostras, me pongo muy enfermo’.

<sup>3</sup> ‘**Axioma Popular 1** (comportamiento de interés en uno mismo). Los agentes realizan acciones que incrementan su propia utilidad personal (y no realizan acciones que disminuyen su propia utilidad personal)’.

“**Folk Axiom 2** (Self-Interested Attitude). Agents think that actions increasing their own personal utility should be taken by others, when others can take these actions (and they think that actions decreasing their own personal utility should not be taken by others, when others can take these actions)”<sup>4</sup> (Bonnefon y Sloman, 2013, p. 195, negritas en el texto original).

Estos axiomas, que van a ser denominados por nosotros en adelante, respectivamente, FA1 y FA2, son relevantes porque Bonnefon y Sloman parecen argumentar que, en realidad, MMT no puede por sí misma ofrecer una explicación para los condicionales de utilidad, y que esta teoría sólo puede proponer una explicación al respecto si asume FA1 y FA2. Más abajo mostraremos que esto, efectivamente, parece correcto. Sin embargo, lo importante en esta sección es describir adecuadamente las características de estos dos axiomas. FA1 es el axioma que permite pensar que el individuo que afirma la sentencia del ejemplo anterior (*Si como ostras, me pongo muy enfermo*) no va a comer ostras. Por otra parte, FA2 permite pensar que un individuo que afirma un condicional como “If she makes that deal, I am ruined”<sup>5</sup> (Bonnefon y Sloman, 2013, p. 194) cree que ella no debería aceptar el trato.

Empero, como hemos señalado anteriormente, Bonnefon y Sloman (2013) también consideran que la estructura de los condicionales es igualmente importante en este sentido. A su juicio, también lo hemos mencionado, contamos con dos tipos de condicionales relevantes para esta problemática: los condicionales causales y los condicionales diagnósticos. En su opinión, los condicionales causales son condicionales en los que el antecedente es una causa evidente del consecuente y los condicionales diagnósticos son condicionales en los que el antecedente es únicamente una señal o un síntoma de que el consecuente es verdadero. Por ejemplo, podemos interpretar que el condicional anterior (*Si ella acepta ese trato, estoy arruinado*) significa que la aceptación del trato por parte de ella provocará mi ruina. En este caso, el condicional sería causal. No obstante, también se puede interpretar que ese mismo condicional significa que ella sólo aceptará el trato si yo estoy en problemas económicos y financieros, ya que el trato es aceptado generalmente por personas

---

<sup>4</sup> ‘**Axioma Popular 2** (actitud de interés en uno mismo). Los agentes piensan que las acciones que incrementan su propia utilidad personal deben ser realizadas por los otros, cuando ellos pueden realizar tales acciones (y piensan que las acciones que disminuyen su propia utilidad personal no deben ser realizadas por los otros, cuando ellos pueden realizar tales acciones)’.

<sup>5</sup> ‘Si ella acepta ese trato, estoy arruinado’.

con dificultades económicas y financieras. En este segundo caso, el condicional sería diagnóstico.

Bonnefon y Sloman piensan que la utilidad sólo influye en los condicionales causales y que, cuando un condicional es diagnóstico, la utilidad no guía los procesos inferenciales en los que dicho condicional puede verse eventualmente incluido. Con el fin de demostrar este hecho, plantean cuatro experimentos. Sus dos primeros experimentos hacen referencia a FA1 y tanto el tercero como el cuarto se encuentran directamente relacionados con FA2. En concreto, en el primer experimento, los condicionales utilizados estaban vinculados a comentarios entre paréntesis y, según Bonnefon y Sloman (2013), eran tales comentarios los que posibilitaban interpretar los condicionales como causales o como diagnósticos. Los ejemplos que presentan son los siguientes:

Condicional causal: “If he buys this house, then is rich. (Because he can resell it for twice the price). Is he going to buy this house?”<sup>6</sup> (Bonnefon y Sloman, 2013, p. 197).

Condicional diagnóstico: “If he buys this house, then he is rich. (Because it costs a fortune). Is he going to buy this house?”<sup>7</sup> (Bonnefon y Sloman, 2013, p. 197).

En el texto original, en el ejemplo correspondiente al condicional causal figura una nota después de la palabra *price*<sup>8</sup> que indica que la redacción no es del todo correcta porque se procuró que las estructuras en las dos condiciones experimentales fueran similares. Empero, el aspecto más importante de este experimento para este trabajo es que Bonnefon y Sloman (2013) predijeron que la utilidad sólo iba a ser relevante en la condición relativa al condicional causal y que, por tanto, sus participantes sólo iban a responder positivamente en esa condición (tenían que responder, en las dos condiciones, en base a una escala que oscilaba entre la opción *Certainly not*<sup>9</sup> en un extremo y la opción *Certainly*<sup>10</sup> en el otro).

---

<sup>6</sup> ‘Si él compra esta casa, entonces es rico (porque puede revenderla por el doble de su precio. ¿Comprará él esta casa?’

<sup>7</sup> ‘Si él compra esta casa, entonces es rico (porque cuesta una fortuna). ¿Comprará él esta casa?’

<sup>8</sup> Precio.

<sup>9</sup> Ciertamente no.

<sup>10</sup> Ciertamente.

Su segundo experimento fue planteado con la intención de asegurar que las condiciones del primer experimento no estaban excesivamente focalizadas en la probabilidad del consecuente (esto es, en el caso de los ejemplos anteriores, en la probabilidad de que el personaje descrito sea rico). Por esta razón, los comentarios aparecían antes de las sentencias condicionales en este experimento. Sus ejemplos son éstos:

Condicional causal: “The fact that he buys this house would make him rich as a result. In other words: If he buys this house, then he is rich”<sup>11</sup> (Bonnefon y Sloman, 2013, p. 199).

Condicional diagnóstico: “The fact that he buys this house would indicate that he is rich. In other words: If he buys this house, then he is rich”<sup>12</sup> (Bonnefon y Sloman, 2013, p. 199).

Obviamente, en este experimento, también Bonnefon y Sloman (2013) predijeron que sus participantes iban a responder afirmativamente sólo en la condición del condicional causal, ya que en tal condición se daban las dos características necesarias: presentaba un condicional de utilidad y dicho condicional tenía una estructura causal (de nuevo, se utilizó una escala de *Certainly not* a *Certainly*).

Por otra parte, el tercer experimento fue diferente. Como comentamos, los dos experimentos anteriores se centraron en FA1. Sin embargo, como igualmente apuntamos, este tercer experimento hizo referencia a FA2 y con él se trató de identificar las creencias que los participantes atribuían a otras personas. Los siguientes ejemplos muestran cómo los textos fueron manipulados en esta ocasión:

Condicional causal: “The fact that she buys the house would make him rich as a result. In other words: If she buys a house, then he is rich. Does he think she should buy a house?”<sup>13</sup> (Bonnefon y Sloman, 2013, p. 201).

---

<sup>11</sup> ‘El hecho de que él comprara esta casa lo haría rico como resultado. En otras palabras: Si él compra esta casa, entonces es rico’.

<sup>12</sup> ‘El hecho de que él comprara esta casa indicaría que es rico. En otras palabras: Si él compra esta casa, entonces es rico’.

<sup>13</sup> ‘El hecho de que ella comprara la casa lo haría rico como resultado. En otras palabras: Si ella compra una casa, entonces él es rico. ¿Cree él que ella debe comprar una casa?’

Condiciona! diagnóstico: “The fact that she buys a house would indicate that he is rich. In other words: If she buys a house, then he is rich. Does he think she should buy a house?”<sup>14</sup> (Bonnefon y Sloman, 2013, p. 201).

Nuevamente, Bonnefon y Sloman (2013) esperaban que sus participantes solamente dieran respuestas positivas en la condición del condicional causal, pues esa era la condición con un condicional de utilidad y una estructura causal (por supuesto, emplearon igualmente en este experimento una escala de *Certainly not* a *Certainly*).

Finalmente, el cuarto experimento, aunque estaba centrado también, como reflejamos, en FA2, tenía como objetivo principal comprobar de manera concluyente que la utilidad es un factor de suma importancia y que, cuando se usa un condicional causal sin una utilidad clara, las conclusiones de los participantes son muy diferentes (ya que, si no existe una referencia clara a una situación valorada, los participantes no necesariamente consideran que esa situación es probable). Con estos propósitos, Bonnefon y Sloman (2013) plantearon cuatro condiciones en este experimento. Dos condiciones presentaban condicionales causales, pero en una de ellas el condicional remitía a una situación de clara utilidad y en la otra no. Por su parte, en las otras dos condiciones figuraban condicionales diagnósticos, pero, del mismo modo, en una de ellas la referencia a la utilidad era obvia y en la otra no. Sus ejemplos son:

Condiciona! causal con utilidad clara: “Her accepting the deal would cause him to be ruined. In other words, if she accepts the deal, then he will be ruined. Does he think she should accept the deal?”<sup>15</sup> (Bonnefon y Sloman, 2013, p. 202).

Condiciona! causal sin utilidad clara: “Her accepting the deal would cause him to have different prospects. In other words, if she accepts the deal, then he will have different prospects. Does he think she should accept the deal?”<sup>16</sup> (Bonnefon y Sloman, 2013, p. 202).

---

<sup>14</sup> ‘El hecho de que ella comprara una casa indicaría que él es rico. En otras palabras: Si ella compra una casa, entonces él es rico. ¿Cree él que ella debe comprar una casa?’

<sup>15</sup> ‘Su aceptación del trato provocaría que él se arruinara. En otras palabras, si ella acepta el trato, entonces él se arruinará. ¿Cree él que ella debe aceptar el trato?’

Condiciona! diagnóstico con utilidad clara: “Her accepting the deal would indicate that he is ruined. In other words, if she accepts the deal, then he must be ruined. Does he think she should accept the deal?”<sup>17</sup> (Bonnefon & Sloman, 2013, p. 202).

Condiciona! diagnóstico sin utilidad clara: “Her accepting the deal would indicate that he has different prospects. In other words, if she accepts the deal, then he must have different prospects. Does he think she should accept the deal?”<sup>18</sup> (Bonnefon y Sloman, 2013, p. 202).

Como podemos notar, la expresión *different prospects*<sup>19</sup> introducía incertidumbre en las condiciones sin utilidad clara, ya que los participantes desconocían cuáles eran esas perspectivas. De esta manera, Bonnefon y Sloman (2013) predijeron que los participantes en este experimento sólo iban a responder de un modo positivo en la condición en la que el condicional era causal y la referencia a la utilidad era clara, pues ésta era la única condición con los dos elementos necesarios: utilidad y estructura causal (en este experimento, se recurrió, del mismo modo, a una escala de *Certainly not* a *Certainly*).

Efectivamente, sus resultados parecieron apoyar sus supuestos, puesto que, en los cuatro experimentos, sus participantes tendieron a ofrecer una respuesta afirmativa únicamente en las condiciones predichas por ellos, esto es, en las condiciones con condicionales causales y de utilidad. No obstante, como avanzamos más arriba, los resultados que obtuvieron pueden considerarse también consistentes con MMT, siempre que esta última teoría asuma FA1 y FA2. De hecho, como también hemos apuntado, esto es lo que parecen defender Bonnefon y Sloman (2013). En el siguiente apartado mostramos que, ciertamente, una versión de MMT que incluya FA1 y FA2 puede predecir los resultados de los cuatro experimentos descritos en las líneas precedentes.

---

<sup>16</sup> ‘Su aceptación del trato provocaría que él tuviera perspectivas diferentes. En otras palabras, si ella acepta el trato, entonces él tendrá perspectivas diferentes. ¿Cree él que ella debe aceptar el trato?’

<sup>17</sup> ‘Su aceptación del trato indicaría que él está arruinado. En otras palabras, si ella acepta el trato, entonces él tiene que estar arruinado. ¿Cree él que ella debe aceptar el trato?’

<sup>18</sup> ‘Su aceptación del trato indicaría que él tiene perspectivas diferentes. En otras palabras, si ella acepta el trato, entonces él tiene que tener perspectivas diferentes. ¿Cree él que ella debe aceptar el trato?’

<sup>19</sup> Perspectivas diferentes.



## Utilidad, condicionales y modelos mentales

Aunque MMT es bastante conocida, quizás, puede ser apropiado comenzar este punto comentando algunas de sus principales tesis, especialmente, las vinculadas con el razonamiento condicional. MMT plantea que los seres humanos razonan por medio de modelos que elaboran a partir de la información que reciben. De esta manera, ante unos datos concretos, las personas toman en consideración el modelo (o los modelos) que puede (o que pueden) ser coherente (o coherentes) con tales datos. En el caso de un condicional semejante al de la condición causal propuesta en el primer experimento de Bonnefon y Sloman (2013), esto es, semejante a *Si él compra esta casa, entonces es rico*, los tres modelos posibles son los siguientes:

A: Un modelo en el que tanto el antecedente como el consecuente son verdaderos, esto es, un modelo como éste:

Él compra la casa

Él es rico

B: Un modelo en el que el antecedente es falso y el consecuente es verdadero, esto es, un modelo como éste:

Él no compra la casa

Él es rico

C: Un modelo en el que tanto el antecedente como el consecuente son falsos, esto es, un modelo como éste:

Él no compra la casa

Él no es rico

Es importante mencionar que MMT no permite pensar, como otro modelo posible, en la situación correspondiente a la combinación restante, esto es, la situación correspondiente a un escenario en el que el antecedente es verdadero y el consecuente es falso (es decir, en el que él compra la casa y no es rico). Esta no es una combinación

contemplada por MMT porque hace referencia a una situación en la que el condicional es falso (si él compró la casa y, no obstante, no es rico, no es cierto que *Si él compra esta casa, entonces es rico*) y, según MMT, los individuos no se representan como modelos escenarios que, una vez que es aceptada la información, pueden ser calificados como imposibles.

Por otra parte, es igualmente relevante destacar que, de acuerdo con MMT, el modelo A es el único sencillo y explícito. Los otros dos modelos, B y C, son modelos implícitos que sólo pueden explicitarse en virtud de un esfuerzo cognitivo mucho mayor que el requerido para A. Dado que se puede suponer que tal esfuerzo cognitivo mayor implica necesariamente la acción de la memoria de trabajo, para esta teoría, en general, las personas razonan solamente sobre la base de modelos como A, esto es, sobre la base de modelos en los que tanto el antecedente de un condicional como su consecuente son verdaderos. No obstante, los modelos mentales son sensibles a la información contextual y, por ello, en virtud de tal información, un modelo explícito se puede bloquear, de la misma manera que un modelo implícito puede hacerse explícito.

Así, si FA1 y FA2 son aceptados, tener en cuenta estas características de los modelos mentales puede ser suficiente para mostrar que MMT es consistente con los resultados de Bonnefon y Sloman (2013). Los argumentos que exponemos a continuación permiten comprobar esta última afirmación.

Como dijimos, los participantes en la condición causal del primer experimento de Bonnefon y Sloman (2013) tendieron a responder que el personaje descrito iba a comprar la casa. Este hecho no es sorprendente desde la óptica de MMT. El modelo inicial, inmediato, explícito y más sencillo (esto es, A) es un modelo en el que, precisamente, él compra la casa. Los otros dos modelos, como hemos apuntado, se encuentran, en principio, implícitos, y es mucho más difícil acceder a ellos. Además, se puede afirmar que el contexto que acompaña al condicional en esta condición puede bloquearlos.

B hace referencia a una situación en la que el personaje consigue ser rico sin necesidad de adquirir la casa. Esta situación puede ser difícil de comprender para los participantes, ya que saben que no existen muchas oportunidades de ser rico en las sociedades actuales. De este modo, el escenario descrito por B puede ser inconsistente con el conocimiento general de los participantes. No obstante, igualmente, también puede ser

igualmente inconsistente con el texto que aparece en esta condición, pues, en el mundo cerrado descrito por dicho texto, la compra de la casa parece ser la única posibilidad de ser rico.

Con respecto a C, si tenemos presente FA1, podemos decir que el participante va a considerarlo una oportunidad perdida. Y es que cabe la posibilidad de pensar que la situación de C es improbable, puesto que, en ella, el personaje puede ser rico y rechaza la oportunidad.

Por tanto, desde la perspectiva de MMT, es obvio por qué los participantes en la condición causal del primer experimento de Bonnefon y Sloman (2013) tendieron mayoritariamente a considerar la posibilidad de que el personaje compre la casa como muy probable. El escenario de esta condición experimental remite a un modelo explícito y sencillo y, aunque existen otros dos modelos posibles, estos se hallan implícitos, siendo mucho más complejo el acceso a ellos. Además, tales modelos implícitos no parecen muy oportunos si se tienen en cuenta, al mismo tiempo, el texto incluido en esta condición, el conocimiento general de los participantes y FA1.

La situación cambia completamente en la condición diagnóstica de este primer experimento. En ella, A no es ni el único modelo explícito ni el más simple, pues los participantes pueden suponer dos situaciones posibles: que el personaje es rico y que no lo es. Es cierto que, en este escenario, el hecho de que él compre la casa es sólo posible si él es rico, y que esto último puede llevar a aceptar A, pero el participante no sabe con certeza si el personaje es rico o no. Por ello, C, una situación en la que el personaje ni es rico ni compra la casa, también es perfectamente posible. De la misma manera, podemos suponer igualmente que el hecho de que el personaje sea rico no necesariamente conduce a que él adquiera la casa, pues puede darse el caso de que él, efectivamente, sea rico y, sin embargo, no esté interesado en ella (por ejemplo, puede pensar que no es adecuada para él o, sencillamente, que no es de su preferencia). En este sentido, B es también claramente posible.

Por consiguiente, en base a MMT, se puede interpretar que, en esta condición del primer experimento de Bonnefon y Sloman (2013), A, el modelo inicialmente simple, no es el único modelo evidente. Los condicionales diagnósticos introducen cierta incertidumbre con respecto a sus consecuentes (en el ejemplo comentado, con respecto al hecho de que el

personaje sea rico) y, debido a ello, el recurso a los tres modelos posibles parece viable. De esta forma, no es llamativo que los participantes, en condiciones como esta, piensen que los antecedentes (en el ejemplo comentado, el hecho de que él compre la casa) son inciertos.

Podemos plantear estos mismos argumentos con respecto al segundo experimento. Dado que la riqueza es muy valorada en el presente, independientemente de las modificaciones realizadas por Bonnefon y Sloman (2013), A continúa siendo el modelo obvio en la condición causal, pues la compra de la casa es, para el personaje, una oportunidad de ser rico, esto es, una oportunidad que FA1 no permite ignorar. Por otra parte, B y C siguen siendo modelos que no se pueden considerar fácilmente, ya que, como en el caso del primer experimento, se puede pensar que estos modelos son inconsistentes con el conocimiento general de los participantes, el contexto cerrado descrito por el texto y FA1. Así, al igual que en el caso anterior, puede ser difícil imaginar tanto otra oportunidad de ser rico como una situación en la que el personaje rechaza su única posibilidad de ser rico.

Y, desde el marco establecido por MMT, también se puede entender que la condición diagnóstica es similar a la del primer experimento. De nuevo, la incertidumbre acerca del consecuente (no se puede saber si el personaje es rico o no) permite tomar en consideración los tres modelos. Es posible que él sea rico y compre la casa, pero también lo es que él no sea rico y no la compre. Igualmente, puede, incluso, que él sea rico y no compre la casa (porque cabe la posibilidad de que no sea de su agrado). Por estas razones, se puede sostener que MMT también explica los resultados obtenidos por Bonnefon y Sloman (2013) en su segundo experimento.

Por lo que se refiere a los experimentos tercero y cuarto, si los revisamos en función de MMT, podemos ofrecer, de la misma manera, argumentos muy semejantes a los presentados para el primer y el segundo experimento. En la condición causal del tercer experimento, dado que el condicional es *Si ella compra una casa, entonces él es rico*, los tres modelos posibles necesitan algunas modificaciones. De este modo, los tres modelos serían, en este caso, estos:

A': Ella compra una casa

Él es rico

B': Ella no compra una casa

Él es rico

C': Ella no compra una casa

Él no es rico

A' es el modelo explícito y evidente en esta condición porque, como hemos mencionado, ser rico es valorado en el presente y, si nos atenemos a este ejemplo, parece apropiado suponer que él desea ser rico y que, en virtud de FA2, espera que ella compre una casa. Por su parte, B' es difícil de aceptar, puesto que se puede pensar que las oportunidades de ser rico son limitadas y que no son obvias otras posibilidades para lograr serlo, al menos, en el escenario cerrado que acompaña a este experimento. Finalmente, debido a que, como hemos indicado, es evidente que él desea ser rico y espera que ella compre la casa, C' puede entenderse de nuevo como una oportunidad perdida. Así, estos argumentos constituyen una explicación de por qué los participantes de esta condición tendieron a afirmar que el personaje piensa que ella debería comprar la casa.

Pero en la condición diagnóstica del tercer experimento la incertidumbre aparece una vez más. El participante no sabe si el personaje es rico o no (esto es, no sabe si el consecuente del condicional es verdad o no) y, por ello, A', B' y C' son modelos igualmente aceptables. Él puede ser rico y ella comprar una casa (porque él puede pagarla). Del mismo modo, él puede ser rico y ella no comprar una casa (porque, por ejemplo, a él o a ella no le convence ninguna). Y, por supuesto, él puede no ser rico y, por esta razón, ella no comprar una casa. Vemos, así, que la incertidumbre puede provocar que el participante no prefiera ningún modelo en particular y que no tenga seguridad con respecto a si el personaje piensa que ella debería comprar una casa o no.

Finalmente, por lo que se refiere al cuarto experimento, podemos decir que el modelo adoptado por los participantes en la condición con utilidad clara es evidente. Los modelos correspondientes al nuevo condicional (*Si ella acepta el trato, entonces él se arruinará*) son los siguientes:

A'': Ella acepta el trato

Él está arruinado

B'': Ella no acepta el trato

Él está arruinado

C'': Ella no acepta el trato

Él no está arruinado

El modelo explícito e inmediato, de acuerdo con MMT, es A''. Sin embargo, el participante puede obviar este modelo porque él (o ella) sabe que nadie quiere estar arruinado y este modelo apunta a una situación indeseable incompatible con FA2. Por otra parte, B'' no parece ser una situación posible, ya que se puede interpretar, a partir del contexto de esta condición, que, si ella no acepta el trato, él no se arruinará. De esta forma, C'' parece ser el único modelo válido, pues hace referencia a una situación en la que ni él está arruinado ni ella acepta el trato (esto es, a una situación coherente con FA2). Por estas razones, desde los supuestos de MMT, no hay duda en lo relativo a por qué los participantes se inclinaron a responder, en esta condición, que el personaje piensa que ella no debería aceptar el trato.

El caso de la condición diagnóstica con utilidad clara es también fácil de comprender sobre la base de MMT. Su texto apunta del mismo modo a A'', B'' y C'', y es igualmente obvio por qué, en esta condición, los participantes desconocían la opinión del personaje con respecto a la aceptación del trato por parte de ella, pues no sabían si el personaje estaba realmente arruinado. El participante pudo pensar en dos situaciones posibles: él está arruinado y él no lo está. El hecho de que ella acepte el trato sólo parece posible si él verdaderamente está arruinado, y ello puede llevar a aceptar A''. No obstante, el participante no sabe con certeza si el personaje está arruinado o no, por consiguiente, C'', una situación en la que él no está arruinado y ella no acepta el trato, también parece perfectamente viable. Por último, el hecho de que él esté arruinado no implica necesariamente el hecho de que ella acepte el trato, ya que, incluso estando él arruinado, el trato puede ser visto como no deseable (por ejemplo, porque él –o ella- puede tener cierto orgullo y creer que aceptar el trato es denigrante). De esta manera, B'' también es totalmente admisible.

Así, MMT no permite discriminar ante los modelos y decantarse por uno en esta condición. El consecuente es incierto (como hemos señalado, no se sabe si él está arruinado o no) y, debido a ello, los tres modelos posibles pueden ser tenidos en cuenta. Y esto explica, sin duda, por qué, en este escenario, los participantes piensan que el antecedente del condicional (*ella acepta el trato*) es igualmente incierto.

No es complejo tampoco comprender los resultados de la condición causal sin utilidad clara desde los supuestos de MMT. Puesto que el consecuente del condicional de

esta condición no es exactamente el mismo que el de las condiciones previas, los modelos requieren de nuevo pequeñas modificaciones. En este sentido, los modelos adecuados serían ahora los siguientes:

|                               |                                     |
|-------------------------------|-------------------------------------|
| A''': Ella acepta el trato    | Él tiene diferentes perspectivas    |
| B''': Ella no acepta el trato | Él tiene diferentes perspectivas    |
| C''': Ella no acepta el trato | Él no tiene diferentes perspectivas |

En este caso, no se puede decir que FA2 bloquea A''', pues no se sabe si el hecho de tener diferentes perspectivas incrementa la utilidad personal o no, lo que significa que A''' es un modelo válido. Por otra parte, C''' también es aceptable porque, como el participante desconoce si es positivo o no tener diferentes perspectivas, cabe la posibilidad de que el personaje no desee disponer de nuevas perspectivas (y, por tanto, de que no quiera que ella acepte el trato). B''', sin embargo, puede ser perfectamente rechazado por el participante, puesto que describe una situación que es improbable que ocurra en el mundo cerrado presentado en esta condición (se indica que el trato conduce a diferentes perspectivas y no se mencionan otras circunstancias que puedan ofrecer nuevas perspectivas). Empero, aunque B''' sea rechazado, dos modelos, A''' y C''', siguen siendo aceptables, por lo que no es extraño que en esta condición los participantes respondieran que no sabían si él quería o no que ella aceptara el trato.

Los modelos correspondientes a la condición diagnóstica sin utilidad clara son también A''', B''' y C''', pero el problema de esta condición es que, de nuevo, el consecuente es incierto. No se sabe si él tiene o no diferentes perspectivas y, por esta razón, A''' y C''' pueden ser indiscutiblemente aceptados. De esta forma, dado que B''' también es válido aquí (él puede tener diferentes perspectivas y el trato no ser aceptado por cualquier motivo), no es complicado tampoco comprender por qué los participantes en esta condición no supieron si el personaje pensaba o no que ella debía aceptar el trato.

Así, los argumentos precedentes nos muestran que MMT puede explicar los resultados obtenidos por Bonnefon y Sloman (2013) si dicha teoría asume tanto FA1 como FA2. Esto, sin duda, nos conduce a diversas conclusiones.

## Conclusiones

Como se ha indicado más arriba, MMT es una teoría de amplio alcance y de gran éxito en el presente. Entre sus fortalezas está el hecho de que no solo puede explicar los mismos resultados experimentales que otras teorías rivales, sino también interpretar, y hasta predecir, fenómenos que ni siquiera pueden ser entendidos desde otras perspectivas teóricas (véase, p.ej., Orenes y Johnson-Laird, 2012). Desde esta óptica, se puede pensar que los argumentos presentados en este trabajo constituyen una evidencia más en este sentido y, por tanto, otra prueba de que MMT es el enfoque que, en la actualidad, parece tener mayor apoyo empírico.

Por otra parte, las posibilidades prácticas de esta teoría no se limitan a sus potencialidades para entender el modo en que los seres humanos razonan. Su marco ha sido también utilizado como recurso metodológico con diversos propósitos. Entre ellos, está el intento de comprender en mayor medida a través de ella teorías y sistemas de la antigüedad que todavía presentan hoy puntos oscuros y para los que no existen interpretaciones unánimes y consensuadas por parte de los especialistas. Así, se ha usado, por ejemplo, para plantear una acercamiento diferente a ciertos esquemas inferenciales de la llamada *lógica india* (p.ej., López-Astorga, 2016). Del mismo modo, y aunque los partidarios de MMT suelen rechazar habitualmente la necesidad de buscar formas lógicas en cuanto representaciones de la estructura profunda de las sentencias, la teoría también ha sido utilizada en áreas relacionadas con la lingüística y la filosofía del lenguaje para tratar de establecer mecanismos claros, más o menos directos, de traducción del lenguaje natural al lenguaje formal de la lógica, esto es, mecanismos en virtud de los que las sentencias expresadas en la comunicación cotidiana en cualquiera de los idiomas hablados en el presente puedan ser relacionadas con formas lógicas (p.ej., López-Astorga, 2015).

Por lo demás, un punto importante que debe ser igualmente considerado es que, quizás, un aspecto del trabajo de Bonnefon y Sloman (2013) y de la argumentación presentada aquí puede ser discutido. Como se ha dicho y explicado, parece que MMT solo puede ser verdaderamente compatible con los resultados de Bonnefon y Sloman (2013) si extiende sus supuestos y admite, de esta manera, a FA1 y FA2 como elementos fundamentales de su marco global. Esto, a partir de lo expuesto en las páginas precedentes,



se torna, aparentemente, indiscutible. Empero, también cabe otra perspectiva sobre este asunto. MMT vincula posibilidades a las sentencias, como también se ha apuntado, atendiendo al significado de los términos incluidos en ellas y a los contextos en que son expresadas. Por tanto, como se puede observar en la literatura al respecto citada, la pragmática es, de igual modo, una dimensión fundamental de esta teoría. Y tal dimensión puede revelar que asumir FA1 y FA2 puede ser una acción redundante bajo sus tesis, una acción que únicamente hace explícito algo que ya está implícito en su planteamiento. Ciertamente, la pragmática nos puede indicar que los individuos se suelen inclinar, en muy diversas situaciones, por aquello que les es útil, de la misma manera que, en muchos casos, usualmente prefieren que los demás hagan aquello que les beneficia. Y, dados estos hechos, proponer que MMT debe contar con FA1 y FA2 puede ser considerado superfluo.

En cualquier caso, lo que parece estar claro es, como también se ha dicho, que MMT es un enfoque de suma relevancia en el presente en muy diversos ámbitos académicos. Por consiguiente, parece oportuno no solo continuar desarrollando sus tesis y realizando experimentos para comprobar sus predicciones, sino también tratar de utilizarla, como ya se ha hecho con respecto a algunas temáticas e igualmente se ha reflejado más arriba, como recurso metodológico en áreas diferentes a la del razonamiento y la actividad inferencial humana. Y es que cabe la posibilidad de que lo que esta teoría realmente nos esté ofreciendo sea un marco general en el que se pueda encuadrar toda la actividad intelectual humana en general.

**Referencias bibliográficas**

- Bonnefon, J.-F. (2009). A theory of utility conditionals: Paralogical reasoning from decision-making leakage. *Psychological Review*, 116, 888-907.
- Bonnefon, J.-F. y Sloman, S. A. (2013). The causal structure of utility conditionals. *Cognitive Science*, 37, 193-209.
- Deaño, A. (1999). *Introducción a la lógica formal*. Madrid, Spain: Alianza Editorial.
- Gentzen, G. (1934). Untersuchungen über das logische Schließen I. *Mathematische Zeitschrift*, 39(2), 176-210.
- Gentzen, G. (1935). Untersuchungen über das logische Schließen II. *Mathematische Zeitschrift*, 39(3), 405-431.
- Henlé, M. (1962). On the relation between logic and thinking. *Psychological Review*, 69, 366-378.
- Hinterecker, T., Knauff, M., & Johnson-Laird, P. N. (2016). Modality, probability, and mental models. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 42(10), 1606-1620.
- Johnson-Laird, P. N. (2012). Inference with mental models. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning* (pp. 134-145). New York, NY: Oxford University Press.
- Johnson-Laird, P. N. (2015). How to improve thinking. In R. Wegerif, L. Li, & J. C. Kaufman (Eds.), *The Routledge International Handbook of Research on Teaching Thinking* (pp. 80-91). Abingdon, UK, & New York, NY: Routledge.

Johnson-Laird, P. N., Girotto, V., & Legrenzi, P. (1999). Modelli mentali: Una guida facile per il profano. *Sistemi Intelligenti*, XI(1), 63-84.

Johnson-Laird, P. N., Khemlani, S., & Goodwin, G. P. (2015). Logic, probability, and human reasoning. *Trends in Cognitive Sciences*, 19(4), 201-214.

López-Astorga, M. (2015). The disjunction introduction rule: Syntactic and semantic considerations. *Pragmalingüística*, 23, 141-149.

López-Astorga, M. (2016). The Hindu Syllogism, iconic representations, and human thought. *Revue Roumaine de Philosophie*, 60(2), 351-358.

Orenes, I. & Johnson-Laird, P. N. (2012). Logic, models, and paradoxical inferences. *Mind & Language*, 27(4), 357-377.

Quelhas, A. C. & Johnson-Laird, P. N. (2017). The modulation of disjunctive assertions. *The Quarterly Journal of Experimental Psychology*, 70(4), 703-717.

Ragni, M., Sonntag, T., & Johnson-Laird, P. N. (2016). Spatial conditionals and illusory inferences. *Journal of Cognitive Psychology*, 28(3), 348-365.

Rips, L. J. (1994). *The Psychology of Proof: Deductive Reasoning in Human Thinking*. Cambridge, MA: Massachusetts Institute of Technology (MIT) Press.

Wason, P. C. (1966). Reasoning. In B. Foss (Comp.), *New Horizons in Psychology* (pp. 135-151). Harmondsworth (Middlesex), UK: Penguin.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, 273-281.