

**Valor de la Recompensa: ¿Cómo y Para Qué se Usa el Concepto?\***

**Titulillo:** Valor de la Recompensa

**Enlace al doi:** <http://dx.doi.org/10.15446/rcp.v25n2.50405>

JONATHAN BURITICÁ

CRISTIANO VALERIO DOS SANTOS

Universidad de Guadalajara, Guadalajara, México

**Cómo citar este artículo:** Buriticá, J. & Dos Santos, C. V. (2016). Valor de la recompensa: ¿cómo y para qué se usa el concepto? *Revista Colombiana de Psicología*, 25(2), xx-xx. doi: 10.15446/rcp.v25n2.50405

La correspondencia relacionada con este artículo debe dirigirse al Dr. Jonathan Buriticá, e-mail: [jjburitica@cucba.udg.mx](mailto:jjburitica@cucba.udg.mx). Calle Francisco de Quevedo # 180, Col. Arcos Vallarta, Guadalajara, Jalisco, México. C. P. 44130.

ARTÍCULO DE REVISIÓN

RECIBIDO: 4 DE MAYO DEL 2015 - ACEPTADO: 21 DE MARZO DEL 2016

\* Parte del documento es la tesis doctoral del primer autor cuya realización fue financiada por CONACYT beca para realizar estudios de doctorado No. 334780/233159. Este manuscrito fue preparado gracias al financiamiento otorgado por la Secretaria de Educación Pública (SEP) de México proyecto de apoyo a NPTC UDG-PTC-1115.

El *valor de la recompensa* es un concepto utilizado para describir y/o explicar algunos de los efectos observados en diferentes programas de reforzamiento. Por ejemplo, para explicar que una alternativa es preferible a otras menos valiosas (Baum & Rachlin, 1969). También, como se mostrará enseguida, se ha usado para explicar por qué, en ciertas condiciones, los organismos parecen responder más por una consecuencia que por otras, aun cuando otras condiciones permanecen. Un problema asociado al uso del concepto es que se recurre más frecuentemente a él que las veces que se presenta una definición clara y concisa. Desde el sentido común, hablar del valor de las consecuencias parece intuitivamente útil y claro, pero en la literatura científica este concepto no se define estrictamente en muchos casos. Este trabajo es un intento por reunir definiciones del concepto y de los procedimientos asociados a la idea de valor de la recompensa y presentarlos en un marco conceptual relativamente sencillo, pero comprensivo de sus características principales. La revisión permite identificar quiénes y cómo usan el concepto, y la identificación puede ser el primer paso hacia otros análisis en torno a él. Esta revisión no pretende hacer un análisis lógico o filosófico profundo del concepto, sino describir dónde se pueden ubicar las diferentes manipulaciones con relación al cambio en la efectividad de la recompensa que estas variables producen.

La relevancia de esta discusión para la ciencia de la conducta, se presenta en el contexto de los conceptos a utilizar como herramientas para la investigación. Por ejemplo, desde perspectivas como la de Hull o, más contemporáneas, como la de Staddon o Killeen, el uso de variables intervinientes parece el siguiente paso lógico, mientras que en perspectivas como la de Skinner, se requiere el uso de otras formas de organizar los datos que no suponen niveles de análisis diferentes de la conducta. Como veremos, el concepto valor de la recompensa parece un ejemplo de lo que Staddon o Killeen consideran heurísticamente fructífero para el análisis de la conducta y que, en algún sentido, se aleja del ideal skinneriano en la medida en que parece implicar un significado, más allá de la pura abstracción. Esta revisión de la literatura en un sentido amplio, no se restringe a una perspectiva conductual particular (e.g., el análisis experimental de la conducta), e intenta describir y aclarar quiénes lo usan, cómo se ha utilizado y para qué se utiliza el concepto en la ciencia de la conducta y, específicamente, en la psicología experimental.

Para comenzar podemos preguntarnos: ¿Qué es el valor de la recompensa? ¿El valor de la recompensa es una característica o propiedad intrínseca? Parece ser una característica de la recompensa o depender de una suma de condiciones experimentales, contextuales o relativas a otras recompensas en la situación. Antes de adelantar en detalle esta descripción es importante afirmar que existe una literatura amplia sobre la razón por la que una recompensa actúa como reforzador, por ejemplo, teorías de regulación conductual como el principio de Premack u otros relacionados (para una revisión ver Clavijo, 1998). Sin embargo, en la presente revisión, no se aborda el tema, ya que nos interesa establecer en qué medida una recompensa es más o menos efectiva, y de antemano se asume que es un reforzador.

De acuerdo con Hursh y Silberberg (2008), históricamente, al hablar de valor de la recompensa, se hace referencia a su efectividad o fuerza, en términos de qué tantas respuestas un reforzador produce o mantiene. El concepto de valor de la recompensa existe en la literatura sobre investigación del comportamiento desde hace varias décadas. Por ejemplo, Elliot (1930) lo mencionó entre los determinantes de la ejecución efectiva en las tareas de laberinto: a mayor valor de la recompensa, se espera mejor ejecución en la tarea, por ejemplo, mayor velocidad de recorrido o más aciertos sobre el número total de intentos, etc.

Nuestra revisión del uso del concepto valor de la recompensa en la investigación sobre el comportamiento nos hace pensar que este se usa en dos sentidos. Por un lado, igual que Hursh y Silberberg (2008), Mazur (2001) entiende el valor de la recompensa como la eficacia de un evento (reforzador) para aumentar la probabilidad de respuesta. Por otro lado, el valor es similar al concepto económico de utilidad subjetiva; la utilidad se entiende como “una medida abstracta de la satisfacción o de la felicidad que reporta a un consumidor una cesta de bienes” (Mankiw, 2004, p. 284). Teniendo en mente estas dos definiciones podemos proponer que distintas variables como magnitud, concentración, duración de acceso, grado de privación y procedimientos, como asociar el reforzador con malestar inducido por drogas (e.g., cloruro de litio), o la preexposición a estas (e.g., metanfetamina), afectan la capacidad de la recompensa de actuar como reforzador, su efectividad y, tal vez, la satisfacción (utilidad) que genera su consumo.

Estas dos definiciones pueden ser incompatibles o, al menos, surgidas en distintos paradigmas conceptuales. Puede discutirse si estos dos usos del concepto son compatibles, complementarios y/o incompatibles. En principio parecen tener una lógica distinta. El primer uso parece empírico, mientras que el segundo parece inferido; esto implicaría que su papel en la explicación de la conducta es diferente. En un sentido empírico, proponemos que el concepto reúne los efectos de diferentes manipulaciones que, se sabe, tienen efectos en la tasa de respuesta o en la preferencia. En el segundo caso, se supone o, se infiere, que algo en la recompensa o en la evaluación que el sujeto hace de esta, afecta la tasa de respuesta o la preferencia. El primer uso parece similar a lo que en la literatura se ha llamado variable interviniente, pero el segundo parece similar a la idea de constructo hipotético. Más adelante discutiremos ampliamente este tema; por ahora pasamos a los procedimientos que han intentado medir el valor de la recompensa, entendido, de manera general, como eficacia del reforzador.

### **Medir el Valor de la Recompensa**

Una forma de determinar el valor de la recompensa ha sido establecer una equivalencia entre fuerza de la respuesta y valor: así, entre mayor fuerza de la respuesta, se considera que el valor de la recompensa es mayor (Hursh & Silberberg, 2008). Por ejemplo, Skinner (1932a, 1932b) midió la tasa de respuesta de consumo y palanqueo, en ratas privadas de alimento y encontró que la ejecución de los sujetos disminuía de acuerdo con una curva negativamente acelerada<sup>1</sup>. Lo que estos datos indican es que la fuerza de la respuesta disminuye con el consumo de los reforzadores y el paso del tiempo.

Esta aproximación presenta varios problemas. Primero, para medir el concepto de valor de la recompensa, se introduce el de fuerza de la respuesta. ¿Cómo se garantiza una medida confiable de la fuerza de la respuesta? Por ejemplo, una respuesta puede ocurrir a una tasa baja en adquisición y mantenimiento, pero demorar más para extinguirse (Blackman, 1968; Nevin, Tota, Torquato, & Shull, 1990). ¿Qué sería considerado más fuerte en ese caso? De acuerdo con Hursh y Silberberg (2008), los experimentos de Skinner (1932a, 1932b) mostraban otro problema con esta forma de establecer el valor de la

---

<sup>1</sup>De la forma  $N=Kt^n$ , donde  $N$  es igual a las respuestas acumuladas,  $t$  el tiempo transcurrido y  $K$  y  $n$  constantes.  $n$  es el parámetro relevante para determinar la forma de la curva. Skinner reporta que este valor se encontraba alrededor de .7.

recompensa: la tasa de respuesta dependía del tipo de contingencia programada, no solo del reforzador. Por ejemplo, un programa de razón variable genera una tasa de respuesta mayor que un programa de intervalo variable, con el mismo reforzador e intervalo entre reforzadores (Reynolds, 1975) por lo que en este caso la tasa no refleja el valor de la recompensa, que es el mismo en ambos programas.

Otra forma de medir el valor de la recompensa, en términos de fuerza de la respuesta, es la ley de igualación. Este principio fue propuesto por Herrnstein (1961, 1970) como una forma de mostrar que la tasa de respuesta depende de la tasa relativa de reforzamiento. En la investigación inicial, Herrnstein (1961) encontró que en un programa concurrente de dos intervalos variables, el porcentaje de respuestas en un componente igualaba el porcentaje de reforzadores obtenidos. También mostró que el número de respuestas en cada componente parece ser una función lineal del número de reforzadores obtenido, de acuerdo con la siguiente función:

$$p = ke \tag{1}$$

Donde  $p$  es la frecuencia absoluta de respuestas, y  $e$  la frecuencia de respuestas de comer; el intercepto de esta función sería cero y  $k$  la pendiente. Usando esta función, Herrnstein derivó que para dos componentes la tasa relativa debería tomar la siguiente forma:

$$\frac{p_1}{p_1 + p_2} = \frac{ke_1}{k(e_1 + e_2)} \tag{2}$$

Esta formulación sería consistente con la afirmación: “la tasa de respuesta es una representación de la fuerza de la respuesta, que es también función lineal de la frecuencia de reforzamiento” (p. 270). Así, Herrnstein (1961) justifica medir la fuerza de la respuesta usando programas concurrentes de intervalo variable.

Se han planteado diferentes propuestas para medir experimentalmente el valor de la recompensa, algunas más exitosas que otras. Por ejemplo, Verhave (1963) expuso algunas palomas a un procedimiento de razón fija ajustable en un programa múltiple con dos componentes (un procedimiento similar a los programas concurrentes de tipo Findley): un

componente era razón fija (RF) 100, y el otro RF 10. El cambio del componente RF 100 al componente RF 10 ocurría cuando el sujeto respondía en una tecla de cambio con RF inicial dos, que aumentó dos respuestas cada vez que los sujetos cambiaban de componente. Si el sujeto no respondía en la tecla de cambio por un tiempo determinado (*Allotated Time Criterion*, ATC), la razón disminuía dos respuestas. La variable dependiente de interés fue el punto en el que se estabilizaba la RF en la tecla de cambio. El autor llamó a esta razón valor de equilibrio y lo consideró un índice de atractivo (*seducibility*) del cambio. El valor de equilibrio de Verhave (1963) parece medir la preferencia relativa entre programas de RF, o del esfuerzo necesario para conseguir una recompensa. Así, cambiar del programa de RF 100 al de RF 10 solo es atractivo si el número de respuestas en la tecla de cambio está entre 62 y 68. Sin embargo, el procedimiento de Verhave (1963) parece depender del efecto de diferentes variables al mismo tiempo, lo cual dificulta determinar cómo afecta cada variable la medida. Por ejemplo, el tiempo criterio para disminuir la razón de cambio (ATC) podría afectar el intervalo entre reforzadores (IER) o la demora de reforzamiento. Además, el incremento en la razón progresiva de la tecla de cambio parece afectar los datos obtenidos en este tipo de procedimientos (ver Cabrera, Robayo-Castro, & Covarrubias, 2010; Stafford & Branch, 1998). Finalmente, las mediciones no parecen sensibles al cambio en la privación, lo que en principio, parece incongruente con la idea de que el valor de la recompensa depende parcialmente de la restricción o estado motivacional impuesto al organismo.

Otros procedimientos tienen una lógica similar al experimento de Verhave, aunque usan menos elementos. Hodos (1961) propuso medir la fuerza de la recompensa o su atractivo (*attractiveness*) utilizando un programa de razón progresiva. En estos programas, cada vez que el sujeto obtiene la recompensa, el número de respuestas requeridas para obtener la siguiente, aumenta de acuerdo con un criterio fijo (e.g., dos respuestas). Hodos (1961) encontró que el punto de quiebre o el número de respuestas en el que el sujeto abandona la tarea, dependía de la concentración de la recompensa. Así, en ratas no privadas (de alimento, agua u otras cosas), el punto de quiebre fue cercano a 70, cuando la concentración de leche condensada endulzada fue mayor, y aproximadamente cinco cuando la concentración fue baja o se les presentó solo agua. El punto de quiebre aumentó cuando

el peso corporal de los sujetos fue menor, cuando la privación fue mayor, y cuando la magnitud del reforzador fue mayor.

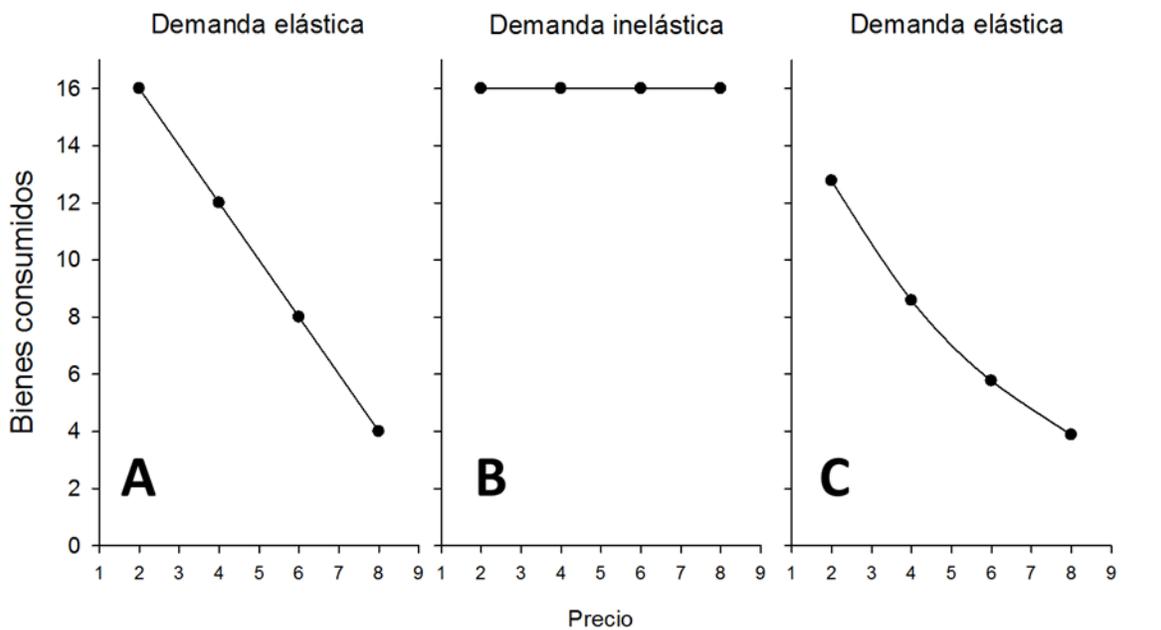
El procedimiento de razón progresiva también se ha utilizado para determinar la fuerza de la recompensa al usar reforzadores con diferentes texturas. Cabrera et al. (2010) intentaron establecer si el amaranto, un tipo de grano común de México y de origen prehispánico, y pellets estándares funcionaban igual como reforzadores. Encontraron que la pausa postreforzamiento obtenida fue mayor para el amaranto que para los pellets, aunque la tasa de carrera fue igual. El punto de quiebre fue mayor para pellets que para el amaranto; esto indicaría que los pellets tienen un valor de recompensa mayor que la misma magnitud de amaranto. Estas diferencias fueron estadísticamente significativas cuando utilizaron una progresión de uno, y aunque la tendencia se mantuvo usando una progresión de tres, las diferencias no fueron significativas en esa progresión. Esto último sugiere que el criterio de aumento de la razón parece afectar la medición del valor, como dicen Hursh y Silberberg (2008), lo cual puede ser un problema al usar este procedimiento para medir el valor de la recompensa.

Hursh y Silberberg (2008) proponen otra forma de establecer el valor de la recompensa, mediante herramientas conceptuales y metodológicas de la economía conductual. Ellos definen explícitamente el concepto valor de la siguiente forma:

En términos de como la palabra se usa comúnmente, el “valor” de un bien es la función conjunta de características hedónicas positivas y negativas. Las características positivas en valor son todos los atributos apetitivos que tiene el bien, como sus posibles efectos biológicos positivos en el consumidor. Las características negativas en valor son todos los rasgos del precio que se requieren para conseguir el bien. Por ejemplo, el tiempo o las respuestas que se necesitan para conseguir cada bien. (p. 189)

Christensen, Silberberg, Hursh, Huntsberry y Riley (2008) reportan un experimento que sigue esta propuesta. Los autores intentaron establecer el valor de recompensa de pellets de comida y administración intracraneal de cocaína. La literatura sobre farmacología muestra que, en algunos casos, la cocaína parece más efectiva como reforzador que la comida, mientras que se observa lo contrario en otros casos, incluso usando otras drogas como heroína (Elsmore, Fletcher, Conrad, & Sodetz, 1980). Christensen et al. (2008) entrenaron a ratas Long-Evans a responder por pellets y luego por cocaína (1 mg/kg) en

programas RF 10. En seguida, utilizaron diferentes RF que aumentaron de 3 a 560 en diferentes fases con componentes reforzados con pellets o cocaína alternados de manera aleatoria. Hallaron que el número de reforzadores consumidos en los componentes con comida era mayor que en los componentes de cocaína en las RF bajas. Cuando utilizaron medidas relativas, observaron que el consumo de comida se mantuvo constante en diferentes valores de RF, en comparación con el consumo de cocaína, que disminuyó en RF altas. Concluyeron que el valor de recompensa fue mayor para la comida que para la cocaína. En el análisis matemático de los resultados, los autores ajustaron el modelo propuesto por Hursh y Silberberg (2008), donde un parámetro representa el grado de disminución del consumo del reforzador, a medida que aumenta la RF. El valor de este parámetro, ya sea que los reforzadores se presenten al tiempo o por separado, sugiere que la comida es más valiosa que la cocaína.



**Figura 1.** Ejemplos de curvas de demanda.

Lo que representa el modelo de Hursh y Silberberg (2008) es la sensibilidad del consumo de la recompensa al costo-beneficio o, en términos de la economía conductual, la elasticidad de la demanda. Una curva de demanda representa el consumo de un bien dado un precio establecido en el mercado. Por ejemplo, un agente puede consumir 16 unidades de un bien, cuando el precio es dos, pero consumir solo 12 cuando el precio es cuatro, y

consumir cuatro si el precio aumenta a ocho (Figura 1, panel A). En este caso, el consumo del bien depende del precio, es decir la demanda del bien es elástica. Cuando la demanda del bien no es elástica, el aumento del precio no afecta el consumo del bien; así, cambios en el precio no afectan el número de unidades consumidas por el agente (ver panel B, Figura 1). Un caso más ajustado a lo que sucede realmente se ve en el panel C de la Figura 1, donde la relación entre consumo y precio se describe como una función exponencial simple. De acuerdo con Madden (2000), la demanda es elástica cuando un cambio del precio del bien en 1% produce cambios en el consumo mayores al 1% y es inelástica cuando el cambio en 1% del precio produce cambios en el consumo menores al 1%.

Si el valor de la recompensa se entiende como elasticidad de la demanda, la razón de cambio entre precio y consumo puede verse como una representación cuantitativa de ese valor: mientras menor sea la razón de cambio, mayor es el valor, ya que el sujeto está dispuesto a pagar un precio alto para mantener el consumo del bien. En una función lineal, la razón de cambio es la pendiente de la recta. Lo que realmente se observa en este tipo de situaciones es un cambio gradual no necesariamente lineal, por lo que medir el valor de la recompensa en dichas funciones curvas plantea ciertos retos. El problema de medir el valor del reforzador, en términos de la elasticidad de la demanda, es que este valor cambia a lo largo de la curva consumo-precio. Es decir, la razón de cambio para diferentes puntos es diferente, por lo que no se puede asignar un valor único a la recompensa; de hecho, en cada punto hay una representación diferente de valor. La alternativa, según Hursh y Silberberg (2008) es, como mencionaba, un modelo con un parámetro que representa la elasticidad para toda la curva. Ese parámetro es la representación del valor *esencial* de la recompensa.

El modelo de Hursh y Silberberg (2008) no es el único planteado para medir el valor de la recompensa. Reilly (2003) utiliza una derivación de un modelo general, desarrollado por Killeen (1994), que intenta dar cuenta de la conducta en función de componentes motores, de memoria y motivacionales en diferentes programas de reforzamiento. Reilly (2003), usando ratas machos Sprague Dawley, se propuso conocer los efectos de la D-anfetamina en la ejecución en programas de reforzamiento y, especialmente, si estos efectos pueden atribuirse a cambios en componentes motores, cognitivos o al valor del reforzador. Para esto, primero estableció que su método registraba los cambios en el valor de la recompensa, mientras mantenía constantes otras variables relevantes. Utilizó un programa

múltiple de cinco componentes de RF, cuyos valores iban de 5 a 150. Se utilizaron diferentes magnitudes de reforzador (uno y tres pellets), diferentes tipos de pellets (solo azúcar, azúcar y concentrado, solo concentrado), y diversas concentraciones de D-anfetamina. Al ajustar el modelo, Reilly (2003) observó que el parámetro que representa el valor de la recompensa fue mayor cuando entregó más pellets y cuando entregó solo azúcar, disminuyó con la mezcla de pellets de azúcar y concentrado y, finalmente, fue más bajo con pellets de solo concentrado. La D-anfetamina no parece tener un efecto en concentraciones bajas, pero parece disminuir el valor de la recompensa en concentraciones altas. En el modelo utilizado por Reilly, el parámetro representa la cantidad de conducta, o grado de activación, que produce el reforzador; en alguna medida es similar al parámetro del modelo de Hursh y Silberberg (2008), ya que parece representar qué tanto el sujeto sigue respondiendo, a medida que aumenta la RF, es decir, qué tanto defiende el organismo el consumo del reforzador al incrementar la RF (el precio). Para un caso similar puede verse el trabajo de Belke y colaboradores (Belke, 2006; Belke & Pierce, 2009; Belke, Pierce, & Duncan, 2006; Belke & Wagner, 2005).

Hasta ahora hemos mencionado que el valor de la recompensa se asocia a la idea de eficacia del reforzador, y mencionamos algunas formas que se han propuesto para medirlo; entre dichas formas están la tasa de respuesta o consumo del reforzador, los puntos de quiebre en programas de razón progresiva, la ley de igualación y diferentes modelos matemáticos, en los que un parámetro representa la resistencia del sujeto por mantener el consumo del reforzador. No obstante, el término valor de la recompensa o eficacia del reforzador es algo vago. Con la finalidad de aclarar la idea, introduciremos la distinción entre constructo hipotético y variable interviniente, la cual nos permite dar ejemplos en los que el concepto de valor de la recompensa se ha utilizado para describir y/o explicar experimentos y resultados, aun cuando no se mide directamente el valor. Algunas de las maneras de medir el valor se asocian al concepto de variable interviniente, por ejemplo, los puntos de quiebre en los programas de razón progresiva o el mantenimiento de la tasa de respuesta en los programas de RF.

### **Valor de la Recompensa como Constructo Hipotético**

Existe literatura que utiliza el concepto de valor de la recompensa como constructo hipotético, en el sentido de ser algo que existe más allá del nivel de observación analizado (MacCorquodale & Meehl, 1948). Esta definición sugiere que el valor de la recompensa existe, en alguna medida, como algo diferente de los procedimientos que se utilizan, por ejemplo, al hablar de la satisfacción que genera el consumo de un bien o de la representación neural de la recompensa. Un caso de este tipo se observa en Sugrue, Corrado y Newsome (2004), quienes sostienen que el “valor [de la recompensa] es un concepto en sí mismo subjetivo. Para estudiar el valor [de la recompensa] es necesario primero operacionalizarlo en términos de las variables que pueden ser observadas directamente” (p. 1786). De acuerdo con Sugrue et al., los determinantes observables del valor son características de la recompensa, como su magnitud y su frecuencia, pero el valor en sí mismo es la representación neural, asociada a estas manipulaciones.

Otro caso del valor de la recompensa como constructo hipotético es el de Rangel, Camerer y Montague (2008), quienes propusieron, como parte de su modelo neurobiológico de elección, que hay sistemas neurológicos de representación del valor. Un ejemplo de este razonamiento en investigación experimental son algunos estudios, como el de Chib, Rangel, Shimojo y O’Doherty (2009), quienes sugieren que bienes de diferentes tipos son representados en el cerebro en áreas similares, como si existiera una moneda (*currency*) común que representa el valor y puede ser utilizada para comparar diferentes bienes y así elegir uno. En dos experimentos de Chib et al. (2009) con participantes humanos, se estableció inicialmente un precio para diferentes bienes, a través de una subasta. Luego, mientras hacían una imagen por resonancia magnética funcional (fMRI), presentaron dos alternativas a los sujetos: el precio promedio de los bienes de la subasta o un ítem equivalente a ese precio, y otros bienes, seleccionados al azar, del conjunto de bienes subastados. A continuación, asociaron la disposición a pagar por un ítem con la activación de áreas del cerebro, y encontraron que esta activación fue alta en el mismo conjunto de áreas, la corteza prefrontal ventromedial, para los diferentes ítems presentados. Los resultados sugieren que la corteza prefrontal ventromedial codifica un valor común de diferentes categorías de bienes que los participantes eligen en el experimento.

En esta literatura se presenta controversia en cuanto a qué áreas del cerebro pueden asociarse a la representación de la recompensa (Gallagher, McMahan, & Schoenbaum,

1999) o si diferentes tipos de valoración ocurren en la misma área (Montague & Berns, 2002). En todo caso, la idea del valor de la recompensa como una representación fisiológica, es muy frecuente y parece generar hipótesis y datos interesantes para la ciencia de la conducta en general (ver Montague & Berns, 2002; Padoa-Schioppa & Assad, 2008; Sugrue et al., 2004).

### **Valor de la Recompensa como Variable Interviniente**

Otras definiciones o usos del concepto, similares al de variable interviniente en el sentido propuesto por MacCorquodale y Meehl (1948), las utilizan algunos autores que no hacen explícitamente el trabajo de vincular diferentes variables dependientes e independientes, por lo que, tal vez, no se pueda hablar de variable interviniente en muchos de esos casos, en sentido estricto. En una variable interviniente se resume un conjunto de relaciones entre cambios en variables dependientes, al menos dos, originados en la manipulación de distintas variables independientes, también, al menos dos. Aunque este resumen de efectos de relaciones entre variables pudiera ser similar en principio, a la idea de constructo hipotético ya operacionalizado, la gran diferencia entre ambos parece ser la suposición, en el segundo caso, de que el concepto existe en un nivel de observación o descripción más allá del nivel donde se observan las relaciones funcionales entre variables dependientes e independientes.

El valor de la recompensa se ha utilizado como etiqueta para agrupar procedimientos que disminuyen o aumentan la tasa de respuesta y, en algún sentido, como constructo para explicar resultados, preferencias o tasas de respuestas, sin que el valor sea objeto directo de medición. A continuación, exponemos tres casos: uno asociado a la investigación sobre aprendizaje asociativo en procedimientos pavlovianos y dos en procedimientos instrumentales.

Holland y Rescorla (1975) intentaron establecer si la devaluación del estímulo incondicionado (recompensa) afectaba la respuesta condicionada a estímulos condicionados (EC) de primer y segundo órdenes, mediante un diseño de comparación entre grupos. Inicialmente, todos los sujetos fueron expuestos al apareamiento EC (luz) - comida. Luego dos grupos fueron entrenados en EC (tono) - comida y otros dos en EC de segundo orden (tono) - EC (luz). Entonces, un grupo que solo recibió entrenamiento EC - comida y otro que

recibió entrenamiento EC - EC de segundo orden, recibieron apareamiento entre la comida y el malestar inducido por poner a los sujetos en una plataforma giratoria; los otros dos grupos solo fueron expuestos a los giros en la plataforma. Finalmente, probaron en extinción la respuesta al tono, que en unos casos es EC y en otros es EC de segundo orden. Los resultados mostraron que el grupo entrenado con el tono como EC y la comida asociada al malestar tuvo tasas de respuesta menores que los otros grupos. En el segundo estudio que reportan, en lugar de parear el EC con malestar, disminuyen la privación del reforzador y observan resultados similares. Estos resultados sugieren que el tratamiento de devaluación, entrenamiento reforzador-malestar, o disminución de la privación, solo fue efectivo en el condicionamiento de primer orden.

La idea que defienden Holland y Rescorla (1975) es que el condicionamiento de primer orden crea una asociación entre EC y una representación del reforzador. Esta incluye una evaluación o valoración que, probablemente, tiene en cuenta sus características apetitivas y aversivas<sup>2</sup>. En el caso del condicionamiento de segundo orden, la asociación no incluye la representación del reforzador, por lo que, los cambios en las características de este no afectan la tasa de respuesta evocada por el EC de segundo orden. Durante la prueba, luego del tratamiento de devaluación, asociación con malestar o disminución de la privación, la tasa de respuesta evocada por el EC disminuye, ya que el valor de la recompensa, su representación de características positivas, disminuye con la devaluación.

En un experimento similar en condicionamiento instrumental, Balleine y Dickinson (1991) intentaron mostrar que el efecto de devaluación de la recompensa por asociación con cloruro de litio (LiCl) se debe a la experiencia con los efectos anticipatorios negativos de la exposición a esta. Es decir, para que se observe una reducción en la tasa de respuesta, como consecuencia de asociar la recompensa con LiCl, los sujetos tienen que experimentar los efectos negativos anticipatorios a su consumo. En este caso, el sabor de la recompensa está asociado a la toxicosis provocada por el LiCl y su consumo parece producir reacciones

---

<sup>2</sup>Puede argumentarse que hablar de representación de la alternativa es utilizar un constructo hipotético, por lo que este ejemplo debería ir en otra sección, ya que se habla de un nivel de análisis diferente al estrictamente conductual, y tal vez sea el caso. Sin embargo, incluimos el ejemplo en esta sección, ya que la operacionalización realizada y el uso de diferentes variables, que tienen un efecto similar, es consistente con la lógica de variable interviniente. Además, la idea de asociación utilizada por Rescorla es, posiblemente, una elaboración de la idea de fuerza del hábito (*habit strength*) propuesta por Hull, que es un ejemplo representativo de variable interviniente en la literatura conductual.

“emocionales” negativas que reducen su valor. Primero, entrenaron la respuesta a la palanca en ratas machos adultas Lister Hooded, usando programas de intervalo aleatorio. Luego, dos grupos experimentales fueron inyectados con LiCl intra-peritonealmente. Uno de estos grupos fue expuesto nuevamente al agua azucarada. En las pruebas en extinción, se observó que el grupo re-expuesto al agua azucarada mostró una menor tasa de respuesta, comparada con los grupos control, que recibieron LiCl, pero no fueron re-expuestos al agua azucarada, o con los grupos que no recibieron LiCl. En este caso, la devaluación de la recompensa solo se observó en el grupo intoxicado, que la experimentó de nuevo. Los autores sugieren que el efecto de devaluación se da después de que el sujeto aprende que la recompensa produce respuestas emocionales negativas, o de expectativa de la toxicosis producida por el LiCl.

En este tipo de procedimientos, no solo se ha devaluado la recompensa, sino que también se ha aumentado su valor. Nordquist et al. (2007) trataron de hallar evidencia sobre dos hipótesis del origen de la dependencia abusiva de drogas. Una de estas hipótesis sostiene que la exposición a drogas, como la D-anfetamina aumenta el valor de incentivo de la recompensa. Según esta hipótesis, el consumo de D-anfetamina sensibiliza al sujeto a la recompensa, aumentando su motivación, medida en tasa de respuestas, para conseguirla. En un experimento, Nordquist et al. compararon la ejecución de dos grupos de ratas en programas de razón aleatoria. Primero inyectaron intra-peritonealmente al grupo experimental (machos Wistar) una solución de 2.5 ml/kg de D-anfetamina por cinco días, y al grupo control una solución de .5 ml/kg de salina. Luego de dos semanas de finalizado el tratamiento con D-anfetamina, entrenaron palanqueo en un programa de reforzamiento continuo, y luego disminuyeron la probabilidad de reforzamiento de la respuesta a .05. La tasa de respuesta del grupo experimental fue significativamente mayor que la del grupo control durante los primeros 10 minutos de la sesión. Nordquist et al. asumen que las manipulaciones afectaron la capacidad de la recompensa de evocar la respuesta, es decir que aumentaron su valor.

En los procedimientos pavlovianos (Holland & Rescorla, 1975) e instrumentales (Balleine & Dickinson, 1991; Nordquist et al., 2007), se considera que el valor de la recompensa cambia cuando disminuye o aumenta la tasa de respuesta, en comparación con condiciones de control. La variable interviniente está anclada a dos cosas, a variables

independientes (los procedimientos) y a una variable dependiente, la tasa de respuesta y, probablemente también, a la ingesta. Para Holland y Rescorla (1975), el malestar inducido en la rueda, y la disminución de la privación devalúan la recompensa, ya que generan una disminución de la tasa de respuesta en extinción. En el caso de Balleine y Dickinson (1991), la devaluación, el cambio en la variable interviniente, ocurre cuando los sujetos son expuestos a dos procedimientos: inyección de LiCl y re-exposición al agua azucarada, pero no sucede al presentar solo un procedimiento a las ratas, y se mide como una disminución de la tasa de respuesta en extinción. Para Nordquist et al., la pre-exposición a la D-anfetamina aumenta el valor de la recompensa, por lo que el grupo de sujetos pre-expuesto tiene una tasa de respuesta mayor, comparada con el grupo de control. La idea de valor de la recompensa resume los efectos combinados de diferentes procedimientos experimentales y el efecto en las medidas dependientes y, en ese sentido, actúa como variable interviniente (MacCorquodale & Meehl, 1948; Zuriff, 1985).

La definición más estricta de valor de la recompensa como variable interviniente en la literatura es, tal vez, la de Baum y Rachlin (1969). Estos autores introducen en la ley de igualación, la idea de que la relación se establece entre el valor y la tasa de respuesta; o el tiempo relativo que el sujeto gasta/está en cada opción. Para ellos, el valor es una variable interviniente que resume el efecto de diferentes variables. Lo primero que hacen Baum y Rachlin es mostrar que la ley de igualación aplica a medidas continuas de respuesta (duración) y no solo a medidas discretas (e.g., frecuencia de picoteos). Habiendo hecho esto, presentan las siguientes generalizaciones en las que afirman que el comportamiento en cada opción ( $T$ ) iguala ciertas características del reforzamiento: su frecuencia ( $r$ ), magnitud ( $a$ ) e inmediatez ( $i$ ); esta última es el recíproco de la demora:

$$\frac{T_1}{T_2} = \frac{r_1}{r_2}$$

$$\frac{T_1}{T_2} = \frac{a_1}{a_2}$$

$$\frac{T_1}{T_2} = \frac{i_1}{i_2}$$

Esto puede resumirse de la siguiente forma, y generalizarse para incluir más características del reforzador, que se denotan como  $x$ :

$$\frac{T_1}{T_2} = \frac{r_1 a_1 i_1}{r_2 a_2 i_2}$$

$$\frac{T_1}{T_2} = \prod_{j=1}^n \frac{x_{1j}}{x_{2j}}$$

Entonces los autores definen el valor de una opción de la siguiente forma:

$$V_1 = \prod_{j=1}^n x_{1j}$$

En otros términos, el valor es el producto de características de la recompensa, como su frecuencia, magnitud o inmediatez, y de otras características que pudieran agregarse empíricamente. Siguiendo la definición, Baum y Rachlin (1969) deducen:

$$\frac{T_1}{T_2} = \frac{V_1}{V_2} \tag{3}$$

Esto quiere decir que el comportamiento asignado a una opción iguala el valor de esa opción.

Un ejemplo de cómo puede ser utilizado este razonamiento es el trabajo de Orduña, Valencia-Torres, Cruz y Bouzas (2013). Ellos usaron ratas como sujetos en programas múltiples encadenados intervalo variable (IV) – tiempo fijo (TF) o intervalo fijo (IF); el último link funcionaba como demoras de reforzamiento para el programa IV. Cada componente tenía programada una magnitud de reforzamiento diferente. Los autores presentan la siguiente formulación de la ley de igualación:

$$\frac{B_1}{B_2} = k \left( \frac{R_1}{R_2} \right)^5 \tag{4}$$

Donde  $k$  y  $s$  son el sesgo y la sensibilidad al reforzador, respectivamente, y ambos factores dependen de la situación de elección, pero no de la tasa de reforzamiento ( $R_1/R_2$ ). Una ventaja, entre otras, de esta expresión es que puede presentarse en forma logarítmica, de manera que ambos parámetros son fáciles de estimar, usando una regresión lineal simple:

$$\log\left(\frac{B_1}{B_2}\right) = s \log\left(\frac{R_1}{R_2}\right) + \log k \quad (5)$$

Si se tiene presente la Ecuación 3 y el hecho de que esta resume los efectos de la demora, la magnitud y la frecuencia se pueden extender en la Ecuación 5, para incluir esos elementos y, así, establecer la sensibilidad a la demora y la magnitud; esa expresión se vería de la siguiente forma:

$$\log\left(\frac{B_1}{B_2}\right) = sr \log\left(\frac{R_1}{R_2}\right) + sd \log\left(\frac{1/D_1}{1/D_2}\right) + sm \log\left(\frac{M_1}{M_2}\right) + \log k \quad (6)$$

Aquí  $sr$ ,  $sd$ , y  $sm$  son parámetros que indican la sensibilidad a la tasa de reforzamiento, a la demora (inmediatez) y a la magnitud. De acuerdo con Orduña et al. (2013), de esta forma se pueden integrar, en un solo índice de valor, las diferentes características del reforzador; también implica que cada característica es independiente de las otras.

Los resultados de Orduña et al. (2013) muestran que la sensibilidad a la demora es dependiente de la magnitud de la recompensa: la demora tiene mayor efecto cuando se presenta en una magnitud grande, que cuando se presenta en una magnitud pequeña. Para determinar esto Orduña et al. (2013) calcularon la sensibilidad a la demora, en componentes con diferentes magnitudes (uno vs. cuatro pellets) y detectaron que esta era mayor en el componente de mayor magnitud. Esto sugiere que la suposición inicial multiplicativa entre los componentes que determinan el valor no se cumple en todos los casos; algo similar ha sido reportado en palomas (Ong & White, 2004). La formulación necesita ser revisada para incluir estos efectos de interacción entre magnitud y demora. Sin embargo, la interacción parece justificar el uso del concepto valor de la recompensa como variable interviniente, ya que el efecto no puede describirse en solo términos de la demora o de la magnitud.

Una de las limitaciones al usar la ley de igualación para escalar el valor de la recompensa parece ser que no tiene en cuenta el efecto del precio y el ingreso sobre la preferencia (Hursh & Silberberg, 2008). Es decir, niveles diferentes de ingreso (e.g., oportunidades para conseguir el reforzador) o el precio de los bienes (reforzadores) aparentemente cambian la preferencia por ellos, afectando su valor y, este hecho, de acuerdo con los autores, no tendría cabida en alguna formulación de la ley de igualación. Una variable importante al momento de considerar el costo de una recompensa es el tiempo de espera para recibirla, o la demora en su obtención. Esta variable, en la actualidad, se ha convertido en un tema de especial relevancia para la investigación en el área, debido a las múltiples implicaciones que su efecto tiene tanto en la práctica como en lo teórico. Por tal razón, la demora de la recompensa se tratará en otra sección, mostrando casos en los que se ha conceptualizado como componente de una variable interviniente.

### **Valor, Demora de la Recompensa y Descuento Temporal**

De acuerdo con Lattal (2010), Thorndike resume en la ley del efecto los principales antecedentes históricos del estudio experimental de la demora de la recompensa. La ley sostiene que:

[D]e las respuestas hechas en la misma situación, aquellas que son acompañadas, o seguidas rápidamente, por satisfacción del animal, si otras condiciones se mantienen constantes, estarían más firmemente conectadas a la situación, por lo que cuando esta situación vuelva a ocurrir es más probable que ocurra la respuesta (Thorndike, 1911, citado en Lattal, 2010, pp. 136-137).

Es decir, la efectividad de la recompensa disminuye a medida que se aleja temporalmente de la respuesta, y se rompe la contigüidad respuesta-consecuencia.

Una explicación posible, aunque no la única, de la disminución de la efectividad de la recompensa demorada puede ser la disminución de la tasa global de reforzamiento. Cuando se introduce la demora en un periodo dado, la tasa de reforzamiento disminuye en comparación con el mismo periodo sin demora, y si la tasa de respuesta es dependiente de la tasa de reforzamiento, una disminución en esta produciría una disminución en la tasa de respuesta. Un argumento similar es presentado por Gallistel y Gibbon (2002), quienes sostienen que no existe efecto de la demora, solo un cambio del tiempo entre reforzadores

estimado por el sujeto, lo que cambiaría el momento en el que las respuestas empiezan a aparecer y, en esa medida, la tasa de respuesta.

Aunque la disminución de la tasa global de reforzamiento podría dar cuenta de los resultados obtenidos al introducir una demora, Lattal (2010) sostiene que hay evidencia del efecto independiente de cambios en la tasa de reforzamiento. Por ejemplo, Sizemore y Lattal (1977, 1978) usando palomas en programas tándem IV - TF, observaron disminución de la tasa de respuesta cuando la demora fue mayor a 3 segundos, en comparación con programas IV de tiempos entre reforzamiento similares a los producidos por el programa tándem sin demora (e.g., tándem [IV 60] [TF 10] vs. IV 70). En los programas tándem, se puede decir que el efecto se debe a la ausencia de contigüidad entre respuesta y consecuencia, ya que en la condición control se mantiene el intervalo entre reforzadores programado y se obtiene una tasa de respuesta mayor.

La demora de la recompensa o la inmediatez de la entrega, es una característica que afecta su valor, entendido desde la ley de igualación (ver Baum & Rachlin, 1969). Chung y Herrnstein (1967) lo demostraron usando palomas y programas concurrentes con dos opciones. Una opción tenía una demora constante entre fases, 0, 8 o 16 segundos a diferentes sujetos. En la otra opción, las demoras cambiaron entre fases y estuvieron en un rango de 1 a 30 segundos. Los sujetos mostraron tasas relativas de respuesta altas en los componentes con demora corta, y bajas en la demora mayor. Los sujetos igualaron la proporción de respuestas a la proporción de inmediatez con la que se entregó el reforzador: mayor inmediatez, mayor tasa de respuesta. Baum y Rachlin (1969) se basaron en esta evidencia para afirmar que existe una relación inversamente proporcional entre valor y demora (ver Ecuación 3).

Otro conjunto de investigaciones que han abordado el tema de la demora de la recompensa, asociado a la idea de valor, se encuentran en el área de descuento temporal (Madden & Johnson, 2010; Odum, 2011a; Rachlin, 2006). En esta área no se mide el valor de la recompensa, sino que se asume que el valor disminuye a medida que se demora su entrega. Se han utilizado varios procedimientos para determinar el descuento temporal y el valor de la recompensa, la mayoría de los cuales busca establecer equivalencias en la

preferencia entre alternativas demoradas y alternativas de menor magnitud entregadas inmediatamente (para una revisión ver Madden & Johnson, 2010).

Uno de los procedimientos para investigar el descuento temporal intenta establecer la equivalencia de valor entre una alternativa de recompensa grande demorada y una pequeña inmediata. Mazur (1987) encontró que hay demoras en la alternativa grande que el animal prefiere con igual frecuencia que una cantidad pequeña, pero inmediata. Mazur presentó a algunas palomas dos opciones en ensayos discretos, una de magnitud pequeña con demora fija (PDF) y otra de magnitud grande con demora ajustable (GDA). Luego de un bloque de cuatro ensayos, dos forzados y dos libres, la demora de la alternativa GDA aumentó un segundo si era elegida en los dos ensayos libres, y disminuyó un segundo si se elegía la PDF, o se mantuvo igual si se elegían ambas opciones (una y una). Entre fases, la opción PDF tenía una demora diferente, lo que permitió hacer varias equivalencias con el valor de la alternativa GDA.

Los puntos de indiferencia obtenidos por Mazur (1987), o los valores de la demora GDA, en los que ambas opciones eran preferidas con igual frecuencia, parecen ajustarse al siguiente modelo:

$$V = \frac{A}{(1 + kD)} \quad (7)$$

Donde  $v$  es el valor de la recompensa,  $A$  su magnitud,  $D$  su demora y  $k$  una constante que representa el cambio en la curva cuando la demora aumenta. Odum (2011b) propone que  $k$  puede ser una variable de rasgo de los sujetos, lo cual explicaría que  $k$  sea relativamente constante, cuando se usan diferentes procedimientos (e.g., consecuencias hipotéticas vs. reales), o dentro de una categoría de personas como fumadores, entre otros (ver también Odum & Baumann, 2010). De acuerdo con Mazur, agregar 1 al denominador evita ciertas dificultades, por ejemplo que  $v$  se aproxime a infinito cuando  $D$  tiende a cero o que sea una indeterminación cuando  $D$  es igual a cero, que sería consecuencia de una ecuación recíproca simple, por ejemplo:  $v=A/kD$ .

Aunque la Ecuación 7 ha funcionado bien en diferentes situaciones (ver Mazur, 2001) algunos autores han objetado el modelo. Myerson y Green (1995) proponen un

modelo hiperboloide o cuasi-hiperbólico, que pretende dar cuenta de los resultados individuales, que parecen no estar bien representados en el modelo de Mazur (1987) agregándole un parámetro adicional. Myerson y Green proponen lo siguiente:

$$V = \frac{A}{(1 + kD)^s} \quad (8)$$

Los autores afirman que el parámetro  $s$  puede reflejar diferencias en la estimación de cantidades y demoras cuando se acoge la suposición de Mazur (1987) sobre la relación directamente proporcional de la magnitud e inversamente proporcional de la demora con el valor.

La introducción de la demora de la recompensa tiene varios efectos: disminuye la tasa de respuesta en procedimientos de operante libre (Sizemore & Lattal, 1977, 1978), disminuye la tasa relativa de respuesta y la preferencia por el componente en programas concurrentes (Chung & Herrnstein, 1967), y disminuye la preferencia por reforzadores de magnitud grande en procedimientos de ensayos discretos (Madden & Johnson, 2010; Mazur, 1987). En esta medida, la introducción de demoras parece tener efectos similares a otros procedimientos mencionados antes que disminuyen el valor de la recompensa, lo cual sugiere que la demora devalúa la recompensa.

### **Consideraciones sobre el Valor de la Recompensa**

De acuerdo con MacCorquordale y Meehl (1948), variables intervinientes como la fuerza del hábito, propuesta por Hull, solo pueden calificarse en términos de conveniencia. En la medida en que son el resumen de un hecho empírico, la relación entre variables dependientes e independientes, no tiene sentido considerarlas no científicas, a menos que se nieguen los hechos resumidos en la variable. Sin embargo, cabe la pregunta de ¿qué tan conveniente es usar estas variables intervinientes? El siguiente razonamiento parece válido: si estas variables intervinientes solo resumen el efecto de un conjunto de variables ¿no sería lo mismo simplemente sostener que  $x$  variable tuvo un efecto similar al de  $y$  variable?

Para Zuriff (1985) una razón para utilizar variables intervinientes es su utilidad como herramienta para deducir efectos de la manipulación de variables similares en diferentes procedimientos experimentales. Por ejemplo, suponiendo que el grado de

privación genera cambios en el valor de la recompensa, se puede deducir que el reforzador es más valioso cuando la privación es más aguda. Si se considera que la demora en la entrega de la recompensa también afecta su valor, se esperaría que la introducción de la demora genere un efecto similar al de disminuir el grado de privación del sujeto o viceversa. En este caso, el valor de la recompensa funciona como puente conceptual entre los dos procedimientos y permite dos cosas: hacer predicciones sobre los efectos de la variable en un procedimiento en el cual no se ha utilizado aún, y utilizar explicaciones propuestas en un caso para todos los casos.

Por ejemplo, si suponemos que el mecanismo por el que la demora genera su efecto es que se produce reforzamiento de otras conductas diferentes a la conducta objetivo (Schaal, Shahan, Kovera, & Reilly, 1998), podemos pensar que tal vez, algo similar ocurre en el caso de la disminución de la privación o el cambio en la concentración de azúcar. Luego de hacer esto, podemos buscar en los datos formas de probar esta explicación o pensar experimentos que nos permitan determinar esto. En resumen, usar el concepto de valor de la recompensa permite, al igual que otras variables intervinientes, hacer varias cosas: (a) resumir sintéticamente los efectos de varios procedimientos, (b) hacer predicciones sobre el efecto de las variables independientes en nuevos procedimientos y (c) sugerir análisis y experimentos posibles.

En cuanto al valor de la recompensa como constructo hipotético el caso es un poco diferente. En un sentido resume los efectos de diferentes variables independientes, suponiendo que todas ellas ejercen influencia sobre la misma cosa, en este caso el valor. Este uso sugiere la existencia del valor como algo independiente de la conducta, por ejemplo, como una representación<sup>3</sup> neurofisiológica del reforzador. La intención no es únicamente resumir resultados de investigaciones, sino sugerir hipótesis acerca de cómo explicar el comportamiento y, especialmente, relacionarlo con otros niveles de análisis, por ejemplo, con procesos fisiológicos particulares que suceden en el cerebro. Esto permite acciones similares a las que permite la variable interviniente, por ejemplo: (a) resumir efectos de varios procedimientos, (b) hacer predicciones sobre el efecto de variables independientes en nuevos procedimientos y áreas de análisis, (c) proponer nuevos

---

<sup>3</sup>En el sentido de estar en lugar de.

experimentos y (d) proponer mecanismos de acción de las variables más allá del nivel de análisis propuesto.

### **Algunas Implicaciones**

El concepto se usa de diferentes formas y en diferentes tradiciones de investigación. Sin embargo, al menos intuitivamente parece referir el mismo tipo de cosas. A nivel operacional parece claro que las mismas variables independientes pueden agruparse tanto en el concepto como variable interviniente o como factores que lo afectan cuando se lo entiende como constructo hipotético. Un análisis crítico de qué tanto se ha avanzado o se puede avanzar usando el concepto de una u otra forma, o si este se usa de manera inconsistente, está fuera del alcance del presente artículo, por lo que, concluir acerca de si el concepto funcionaría en el futuro para avanzar la ciencia de la conducta sería más especulativo que otra cosa. Un análisis filosófico o lógico del tema podría ayudar en este sentido.

Una forma en la que este análisis podría hacerse requeriría el planteamiento explícito del marco conceptual de referencia. Sería necesario especificar qué se entiende por ciencia y qué objetivos pretende alcanzar la empresa científica, entre otras cosas. Por ejemplo, para Skinner (1950) la ciencia de la conducta debe intentar establecer relaciones funcionales entre variables, dependientes e independientes, establecidas en el mismo nivel de observación. Esto puede lograrse sin utilizar otros niveles de análisis, ya que el objetivo de la ciencia de la conducta, para este autor, es la predicción y el control de la conducta, y ambas pueden hacerse sin apelar a constructos o entidades que existen en un nivel de análisis distinto al de la conducta y las variables de las que depende. Desde esta perspectiva, usar el concepto de valor de recompensa como variable interviniente parece consistente y adecuado, ya que Skinner (1950) propone que la teoría apropiada para la ciencia de la conducta puede ser “una representación formal de los datos reducida a un número mínimo de términos. Una construcción teórica puede lograr mayor generalidad que una recolección de hechos. Aunque tal construcción no se referiría a otro nivel de observación [...]” (p. 216). En este sentido, la teoría apropiada para la ciencia de la conducta es exactamente lo que proponemos que hace el concepto de valor de la recompensa entendido como variable interviniente.

El caso del constructo hipotético podría ser interesante para otro tipo de científico de la conducta, por ejemplo uno con un interés más cercano a la fisiología de la conducta. Este segundo científico podría estar más interesado en las implicaciones que el constructo hipotético de valor de la recompensa le sugiere acerca de cómo diferentes manipulaciones de variables independientes deberían afectar los procesos fisiológicos asociados al control de la conducta. Sin embargo, el marco de referencia de este científico hipotético debería ser presentado explícitamente, para determinar si el constructo valor de la recompensa resulta consistente o no con tal forma de pensar la ciencia de la conducta. Lo mismo aplicaría para otros científicos de la conducta, interesados en entender el fenómeno desde perspectivas sociales, y que se inclinarían por constructos al nivel de observación de grupos, culturas o sociedades.

El objetivo de esta revisión fue presentar cómo se ha medido y definido el valor de la recompensa y también esbozar para qué puede servir el concepto. Entendido como constructo hipotético o como variable interviniente el concepto implica diferentes cosas, aunque, en principio, parece útil en la investigación psicológica experimental. La anterior conclusión se sostiene en que aún en la actualidad el concepto sigue dando lugar a nuevas investigaciones y datos en diferentes perspectivas y formas de definir la ciencia de la conducta. Sin embargo, una conclusión fuerte acerca de la utilidad del concepto requeriría una comparación entre propuestas teóricas y usos del concepto, que solo se esboza en este artículo, y que podría ser una tarea filosófica pendiente.

### Referencias

- Balleine, B. & Dickinson, A. (1991). Instrumental performance following reinforcer devaluation depends upon incentive learning. *The Quarterly Journal of Experimental Psychology Section B: Comparative and Physiological Psychology*, 43, 279-296. doi: 10.1080/14640749108401271
- Baum, W. M. & Rachlin, H. (1969). Choice as time allocation. *Journal of the Experimental Analysis of Behavior*, 12, 861-874. doi: 10.1901/jeab.1969.12-861
- Belke, T. W. (2006). Responding for sucrose and wheel-running reinforcement: Effect of prerunning. *Behavioural Processes*, 71(1), 1-7. doi: 10.1016/j.beproc.2005.08.003

- Belke, T. W. & Pierce, W. D. (2009). Body weight manipulation, reinforcement value and choice between sucrose and wheel running: A behavioral economic analysis. *Behavioural Processes*, 80, 147-156. doi: 10.1016/j.beproc.2008.11.006
- Belke, T. W. & Wagner, J. P. (2005). The reinforcing property and the rewarding aftereffect of wheel running in rats: A combination of two paradigms. *Behavioural Processes*, 68(2), 165-172. doi: 10.1016/j.beproc.2004.12.006
- Belke, T. W., Pierce, W. D., & Duncan, I. D. (2006). Reinforcement value and substitutability of sucrose and wheel running: Implications for activity anorexia. *Journal of the Experimental Analysis of Behavior*, 86(2), 131-158. doi: 10.1901/jeab.2006.98-05
- Blackman, D. (1968). Conditioned suppression or facilitation as a function of the behavioral baseline. *Journal of the Experimental Analysis of Behavior*, 11(1), 53-61. doi: 10.1901/jeab.1968.11-53
- Cabrera, F., Robayo-Castro, B., & Covarrubias, P. (2010). The 'Huatli' alternative: Amaranth as reinforcer in operant procedures. *Revista Mexicana de Análisis de la Conducta*, 36, 71-92.
- Chib, V. S., Rangel, A., Shimojo, S., & O'Doherty, J. P. (2009). Evidence for a common representation of decision values for dissimilar goods in human ventromedial prefrontal cortex. *The Journal of Neuroscience*, 29(39), 12315-12320. doi: 10.1523/jneurosci.2575-09.2009
- Christensen, C. J., Silberberg, A., Hursh, S. R., Huntsberry, M. E., & Riley, A. L. (2008). Essential value of cocaine and food in rats: Tests of the exponential model of demand. *Psychopharmacology*, 198(2), 221-229. doi: 10.1007/s00213-008-1120-0
- Chung, S. H. & Herrnstein, R. J. (1967). Choice and delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 10, 67-74. doi: 10.1901/jeab.1967.10-67
- Clavijo, A. (1998). Regulación de la conducta y teoría del refuerzo: conceptos básicos. En R. Ardila, W. López-López, A. M. Pérez, R. Quiñones, & F. Reyes (Eds.), *Manual de análisis experimental del comportamiento* (pp. 115-135). Madrid: Biblioteca Nueva.

- Elliott, M. H. (1930). Some determining factors in maze-performance. *The American Journal of Psychology*, *42*(2), 315-317. doi: 10.2307/1415287
- Elsmore, T. F., Fletcher, G. V., Conrad, D. G., & Sodetz, F. J. (1980). Reduction of heroin intake in baboons by an economic constraint. *Pharmacology Biochemistry and Behavior*, *13*(5), 729-731. doi: 10.1016/0091-3057(80)90018-0
- Gallagher, M., McMahan, R. W., & Schoenbaum, G. (1999). Orbitofrontal cortex and representation of incentive value in associative learning. *The Journal of Neuroscience*, *19*(15), 6610-6614.
- Gallistel, C. R. & Gibbon, J. (2002). *The symbolic foundations of conditioned behavior*. Mahwah, NJ: Lawrence Erlbaum.
- Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267-272. doi: 10.1901/jeab.1961.4-267
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*, 243-266. doi: 10.1901/jeab.1970.13-243
- Hodos, W. (1961). Progressive ratio as a measure of reward strength. *Science*, *134*(3483), 943-944. doi: 10.2307/1708217
- Holland, P. C. & Rescorla, R. A. (1975). The effect of two ways of devaluing the unconditioned stimulus after first- and second-order appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*, 355-363. doi: 10.1037/0097-7403.1.4.355
- Hursh, S. R. & Silberberg, A. (2008). Economic demand and essential value. *Psychological Review*, *115*, 186-198. doi: 10.1037/0033-295x.115.1.186
- Killeen, P. R. (1994). Mathematical principles of reinforcement. *Behavioral and Brain Sciences*, *17*(1), 105-135. doi: 10.1017/S0140525X00033628
- Lattal, K. A. (2010). Delayed reinforcement of operant behavior. *Journal of the Experimental Analysis of Behavior*, *93*, 129-139. doi: 10.1901/jeab.2010.93-129

- MacCorquodale, K. & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*(2), 95-107. doi: 10.1037/h0056029
- Madden, G. J. (2000). A behavioral economics primer. En W. K. Bickel & R. E. Vuchinich (Eds.), *Reframing health behavior change with behavioral economics* (pp. 3-26). Mahwah, NJ: Lawrence Erlbaum Associates.
- Madden, G. J. & Johnson, P. S. (2010). A delay-discounting primer. En G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 11-37). Washington, DC: American Psychological Association.
- Mankiw, G. (2004). *Principios de economía* (3.<sup>a</sup> ed.). Madrid: McGraw-Hill.
- Mazur, J. E. (1987). An adjusting procedure for studying delayed reinforcement. En M. L. Commons, J. E. Mazur, J. A. Nevin, & H. Rachlin (Eds.), *Quantitative analyses of behavior: The effect of delay and of intervening events on reinforcement value* (Vol. v, pp. 55-73). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Mazur, J. E. (2001). Hyperbolic value addition and general models of animal choice. *Psychological Review*, *108*, 96-112. doi: 10.1037/0033-295x.108.1.96
- Montague, P. R. & Berns, G. S. (2002). Neural Economics and the biological substrates of valuation. *Neuron*, *36*(2), 265-284. doi: 10.1016/S0896-6273(02)00974-1
- Myerson, J. & Green, L. (1995). Discounting of delayed rewards: Models of individual choice. *Journal of the Experimental Analysis of Behavior*, *64*, 263-276. doi: 10.1901/jeab.1995.64-263
- Nevin, J. A., Tota, M. E., Torquato, R. D., & Shull, R. L. (1990). Alternative reinforcement increases resistance to change: Pavlovian or operant contingencies? *Journal of the Experimental Analysis of Behavior*, *53*(3), 359-379. doi: 10.1901/jeab.1990.53-359
- Nordquist, R. E., Voorn, P., De Mooij-van Malsen, J. G., Joosten, R. N. J. M. A., Pennartz, C. M. A., & Vanderschuren, L. J. M. J. (2007). Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. *European Neuropsychopharmacology*, *17*, 532-540. doi: 10.1016/j.euroneuro.2006.12.005

- Odum, A. L. (2011a). Delay discounting: I'm a k you're a k. *Journal of the Experimental Analysis of Behavior*, *96*, 427-439. doi: 10.1901/jeab.2011.96-423
- Odum, A. L. (2011b). Delay discounting: Trait variable? *Behavioural Processes*, *87*(1), 1-9. doi: 10.1016/j.beproc.2011.02.007
- Odum, A. L. & Baumann, A. A. L. (2010). Delay discounting: State and trait variable. En G. J. Madden & W. K. Bickel (Eds.), *Impulsivity: The behavioral and neurological science of discounting* (pp. 39-65). Washington, DC: American Psychological Association.
- Ong, E. L. & White, K. G. (2004). Amount-dependent temporal discounting? *Behavioural Processes*, *66*(3), 201-212. doi: 10.1016/j.beproc.2004.03.005
- Orduña, V., Valencia-Torres, L., Cruz, G., & Bouzas, A. (2013). Sensitivity to delay is affected by magnitude of reinforcement in rats. *Behavioural Processes*, *98*, 18-24. doi: 10.1016/j.beproc.2013.04.011
- Padoa-Schioppa, C. & Assad, J. A. (2008). The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. *Nature Neuroscience*, *11*(1), 95-102. doi: 10.1038/nn2020
- Rachlin, H. (2006). Notes on discounting. *Journal of the Experimental Analysis of Behavior*, *85*, 425-435. doi: 10.1901/jeab.2006.85-05
- Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545-556. doi: 10.1038/nrn2357
- Reilly, M. P. (2003). Extending mathematical principles of reinforcement into the domain of behavioral pharmacology. *Behavioural Processes*, *62*, 75-88. doi: 10.1016/S0376-6357(03)00027-5
- Reynolds, G. S. (1975). *A primer of operant conditioning*. Glenview, IL: Scott Foresman.
- Schaal, D. W., Shahan, T. A., Kovera, C. A., & Reilly, M. P. (1998). Mechanisms underlying the effects of unsignaled delayed reinforcement on key pecking of

- pigeons under variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, 69(2), 103-122. doi: 10.1901/jeab.1998.69-103
- Sizemore, O. J. & Lattal, K. A. (1977). Dependency, temporal contiguity, and response-independent reinforcement. *Journal of the Experimental Analysis of Behavior*, 27, 119-125. doi: 10.1901/jeab.1977.27-119
- Sizemore, O. J. & Lattal, K. A. (1978). Unsignaled delay of reinforcement in variable-interval schedules. *Journal of the Experimental Analysis of Behavior*, 30, 169-175. doi: 10.1901/jeab.1978.30-169
- Skinner, B. F. (1932a). Drive and reflex strength. *Journal of General Psychology*, 6, 22-37.
- Skinner, B. F. (1932b). Drive and reflex strength: II. *Journal of General Psychology*, 6, 38-48.
- Skinner, B. F. (1950). Are theories of learning necessary? *The Psychological Review*, 57, 193-216.
- Stafford, D. & Branch, M. N. (1998). Effects of step size and break-point criterion on progressive-ratio performance. *Journal of the Experimental Analysis of Behavior*, 70(2), 123-138. doi: 10.1901/jeab.1998.70-123
- Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2004). Matching behavior and the representation of value in the parietal cortex. *Science*, 304(5678), 1782-1787. doi: 10.1126/science.1094765
- Verhave, T. (1963). Toward an empirical calculus of reinforcement value. *Journal of the Experimental Analysis of Behavior*, 6, 525-536. doi: 10.1901/jeab.1963.6-525
- Zuriff, G. E. (1985). *Behaviorism: A conceptual reconstruction*. New York: Columbia University Press.