

APLICACIÓN DE LA TEORÍA DE DETECCIÓN DE SEÑALES AL ESTUDIO DE LA CATA DE ACEITE DE OLIVA VIRGEN

*María M. Moreno-Fernández, Manuel M. Ramos-Álvarez,
Concepción Paredes-Olay y Juan M. Rosas*

En la sociedad del siglo XXI la importancia de la calidad de los alimentos y de sus efectos beneficiosos para la salud se ha convertido en un objetivo para productores, empresas manipuladoras y comercializadoras, e incluso para los científicos. Esto lleva a que sea el consumidor el que marque claramente la orientación del mercado para tales alimentos. Esta orientación hacia la calidad, tradicionalmente reservada a un grupo más o menos pequeño de alimentos como el vino o las bebidas espirituosas, se está extendiendo de manera cada vez más generalizada a distintos productos como los lácteos, el agua y, en el caso que nos ocupa, el aceite de oliva. Las propiedades organolépticas y sus efectos beneficiosos sobre la salud de los consumidores se convierten en un valor añadido al producto comercializado que, si bien supone una exigencia añadida para la empresa, queda bien recompensada por la elección que realiza el consumidor frente a otros productos similares pero que carecen de ese marchamo (marca) de calidad.

En el caso del aceite de oliva virgen, los atributos físico-químicos (ej., acidez) y las características organolépticas (ej., olor, sabor, textura) determinarán su clasificación en diferentes categorías. De hecho, la prueba de calidad más importante a la que se somete el aceite de oliva es el análisis sensorial que nos dirá cómo es dicho aceite respecto a diferentes atributos como amargo o frutado y en qué categoría comercial se ha de catalogar: virgen extra, virgen, corriente o lampante. Tales atributos sensoriales se evalúan, pues, a través de un proceso de cata que ha de desarrollarse de manera formal y rigurosa para garantizar al consumidor que está recibiendo la calidad por la que realiza el desembolso.

El Consejo Oleícola Internacional (COI) establece claramente el procedimiento de cata para la valoración sensorial del aceite de oliva virgen. En la decisión DEC-21/95-V/07 el COI establece los documentos que regulan, entre otros aspectos, el método para la valoración organoléptica del aceite de oliva virgen (COI/T.20/Doc. No 15/Rev. 2), para la preselección de catadores a través de entrevistas, el procedimiento para determinar sus

umbrales básicos de sensibilidad, así como el entrenamiento específico que deben seguir para familiarizarse con las características olfativas, gustativas y táctiles del aceite de oliva virgen (COI/T.20/Doc. No 14/Rev. 2) (véase también Guerrero, 1996; Gutiérrez, 1997). En la práctica los catadores deben degustar y comparar diversas muestras de aceite de oliva y hacer estimaciones muy finas respecto a los distintos atributos sensoriales asociados a tales muestras, atributos que pueden ser positivos (frutado, amargo y picante) o negativos (atrojado, moho, avinado, borras, metálico, rancio u otros). Estas estimaciones reflejarán aquellos factores de calidad referentes a la maduración, medio agrológico, método de recolección y extracción e incluso a las características de cada variedad (Romero & Tous, 2003; Tous & Romero, 2001).

Estos procedimientos empleados en la selección, entrenamiento, medición y control de panelistas y de la propia situación de cata se basan en la psicofísica clásica de umbrales. La psicofísica clásica considera que las sensaciones que experimentamos ante un estímulo determinado son un reflejo de la intensidad de sus atributos, de tal manera que la percepción tiene lugar únicamente cuando la intensidad supera un determinado nivel o umbral absoluto (Blanco, 1996). Los paneles de cata parten de la idea clásica que entiende a las sensaciones como elementos discretos, pero esto supone importantes limitaciones metodológicas, básicamente porque las respuestas de un receptor cualquiera serán reflejo de su precisión sensorial pero inevitablemente también de su valoración subjetiva. Los catadores evalúan la intensidad de los atributos deseables e indeseables en situaciones que no permiten tener en cuenta los posibles factores subjetivos del proceso.

La limitación metodológica de esta aproximación se pone claramente de manifiesto con un sencillo ejemplo. Imaginemos que un catador recibe 10 muestras de aceite con el encargo de determinar si contienen aceite de oliva o no lo contienen. En nueve de estas muestras se diluye una mínima cantidad de aceite de oliva (ej., 0,8%) en aceite de girasol; en la décima sólo se presenta el aceite de girasol. Cuando el catador se expone aleatoriamente a las 10 muestras, imaginemos que responde que contienen aceite de oliva todas ellas. La tasa de aciertos del catador sería muy elevada, el 90%, y por ende tenderíamos a concluir que éste presenta una capacidad sensorial muy fina. Sin embargo, el mismo resultado podría obtenerse si el catador tuviera una capacidad sensorial más burda, pero un sesgo o

tendencia a responder que la muestra contiene aceite de oliva de manera generalizada. Aunque este ejemplo lo hemos exagerado a propósito con el fin de realzar su poder ilustrativo, lo cierto es que la psicofísica clásica no permite distinguir el proceso sensorial de los sesgos de respuesta que pueda tener el perceptor, siendo más conveniente un modelo que asuma, añadido al proceso sensorial, al menos un segundo proceso que tiene que ver con la estrategia decisoria del perceptor, relacionado con sus intereses, su motivación o su estado de ánimo (véase Ramos, Catena, & Trujillo, 2004). Estos problemas son directamente imputables al procedimiento regulado de la cata de aceite de oliva virgen, dado que como señalábamos más arriba, ésta se fundamenta en los principios de la psicofísica clásica de umbrales. Dentro de la psicofísica contemporánea, estas dificultades se solventan a partir de los principios de la Teoría de la Detección de Señales, cuyos elementos básicos pasamos a exponer.

Principios básicos de la Teoría de Detección de Señales (TDS)

La TDS surge a mediados del siglo XX en el campo de las telecomunicaciones en un intento de mejorar el rendimiento de los radares y sonares que se encontraban entonces en pleno desarrollo. Estos aparatos proporcionan información sobre las señales de interés (ej., un punto que simboliza la presencia de un avión en el espacio aéreo) mezcladas con estímulos irrelevantes que denominamos ruido (ej., otros símbolos y puntos que no representan un avión), lo que dificulta la detección de las señales de interés. Por ejemplo, imaginemos a un técnico de sonar en un submarino que detecta una señal demasiado débil para tener la certeza total de que se trata de un submarino enemigo. En estas circunstancias, la decisión de activar la alarma dependerá de muchos factores, y no sólo de la intensidad de la señal percibida. Por ejemplo, si el submarino se encuentra inmerso en una situación de guerra, aumentará la probabilidad de que el técnico considere la señal como un submarino enemigo, ya que en medio del conflicto, puestos a equivocarse, las consecuencias asociadas a la comisión de un error de falsa alarma son preferibles a las que derivan de dejar de activar la alarma ante la posibilidad de un ataque enemigo. En cambio, en tiempos de paz, es más probable que el técnico considere esta señal débil como parte del ruido. Este ejemplo pone de manifiesto que la respuesta en una situación de incertidumbre está determinada por las repercusiones que se derivan de aquélla y no únicamente por la sensación experimentada

(Wickens, 2002).

Siempre que la situación defina de manera completa los dos tipos de error, la *TDS* permite obtener medidas independientes de la sensibilidad y del criterio de decisión a partir de los datos de las pruebas psicofísicas, despertando gran interés en los psicólogos de la época (véase por ejemplo Green & Swets, 1966). Desde entonces, la *TDS* se ha aplicado con profusión al estudio de casi todos los procesos psicológicos tales como el aprendizaje, la percepción, la memoria o la toma de decisiones, y a tareas tan variadas como el diagnóstico médico (Marshall, 2001; Swets, 1996), la percepción visual (Zalevski, Henning, & Hill, 2007), la predicción meteorológica (Harvey, Hammond, Lusk, & Mross, 1992), o el análisis sensorial de bebidas (Lawless, 1985; Parr, Heatherbell, & White, 2002) y alimentos (Moskowitz, 1988; O'Mahony & Hautus, 2008), entre otras (véase por ej., Hutchinson, 1981 & Swets, 1996).

En un experimento prototípico de detección, los participantes se enfrentan a distintos ensayos en los que deben decidir si se ha presentado o no se ha presentado un estímulo objetivo o, en términos de la *TDS*, la señal. Por ejemplo, supongamos que el objetivo de la tarea es detectar la aparición de un determinado sonido. La situación puede caracterizarse de modo muy sencillo si entendemos que existen dos posibles situaciones estimulares, con el sonido presente (ensayos señal+ruido) o ausente (ensayos ruido), y dos posibles respuestas por parte del participante, “sí” apareció la señal o “no” apareció la señal. Así, toda la información referente a esta situación puede resumirse en una tabla de doble entrada que se conoce como matriz de confusión.

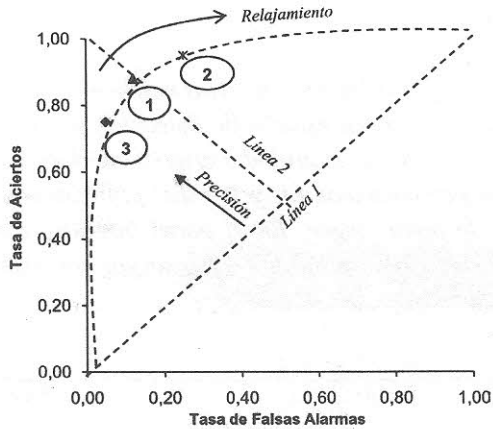
En el panel superior de la Figura 1 se presentan las matrices de confusión de tres participantes hipotéticos, expuestos a un total de 200 ensayos, en 100 de los cuales se presentó el sonido. Así, cada participante recibió 100 ensayos tipo señal+ruido entremezclados aleatoriamente con 100 ensayos tipo ruido. En esta situación se pueden dar cuatro posibles combinaciones. El participante puede responder que la señal está presente en los ensayos “señal+ruido”, lo que categorizamos como acierto; puede responder que la señal no está presente ante este mismo tipo de ensayos, lo que categorizamos como fallo. Por otra parte, el participante puede responder que la señal está presente en los ensayos “ruido”, lo que denominamos falsa alarma. Finalmente, puede responder que la señal no está presente en el ensayo

ruido, manifestando un rechazo correcto.

Figura 1. En el panel superior se presentan las matrices de confusión de tres participantes hipotéticos en un estudio de detección. En el panel central se presentan los índices de sensibilidad y criterio de decisión clásicos (d' y $\ln\beta$) y los no paramétricos (A' y B') que aplicamos a la situación de cata de aceite de oliva virgen. En el panel inferior se representa la curva COR (curvas de iso-sensibilidad e iso-sesgo) correspondiente a la actuación de esos mismos participantes.

Estímulo	Señal+Ruido		Ruido	
	Sí	No	Sí	No
Respuesta	<i>Aciertos</i>	<i>Fallos</i>	<i>Falsas Alarmas</i>	<i>Rechazos Correctos</i>
Participante 1:	88	12	12	88
Participante 2:	95	5	25	75
Participante 3	75	25	5	95

Índices Psicofísicos	Sensorial		Criterio Decisorio	
	<i>Paramétrico (d')</i>	<i>No Paramétrico (A')</i>	<i>Paramétrico ($\ln\beta$)</i>	<i>No Paramétrico (B')</i>
Participante 1:	2.35	.93	.00	.00
Participante 2:	2.32	.92	-1.13	-.60
Participante 3:	2.32	.92	+1.13	+.60



Con objeto de que los distintos tipos de ejecución sean comparativos, las frecuencias se transforman en probabilidades o tasas relativas a la densidad de ensayos “señal+ruido” por un lado y de ensayos “ruido” por otro. Esto es así, además, porque el investigador fija de antemano ambos tipos estímulares, lo que implica que el cálculo de tales probabilidades se condicione a cada tipo de ensayo. Finalmente, esta fijación lleva a que las tasas sean redundantes dos a dos. Bastaría conocer por ejemplo la de Aciertos y la de Falsas Alarmas para caracterizar de manera completa la matriz (las otras dos quedan determinadas automáticamente por el hecho de conocer el total de ensayos “señal+ruido” y de ensayos “ruido”).

Esta tabla permite calcular la *tasa de aciertos* (*TA*) definida como la probabilidad de responder que “sí” se ha presentado la señal en los ensayos tipo señal+ruido (88/100, 95/100 y 75/100, respectivamente para los tres participantes de la Figura 1). Asimismo se calcula la tasa de falsas alarmas (*TFA*) definida como la probabilidad de responder que “sí” hay señal en los ensayos de tipo ruido (12/100, 25/100 y 5/100 respectivamente para los tres participantes de la Figura 1). Obsérvese que ninguna de las dos tasas es suficiente si se toma de manera aislada sino que aportan información complementaria. Tomando sólo en cuenta la tasa de aciertos la actuación del participante 2 es mejor que la del participante 1, y la de éste mejor que la del participante 3. Sin embargo, tomando la tasa de falsas alarmas lo que observamos es lo contrario. Estas dos medidas de ejecución se han empleado tradicionalmente para la realización de las curvas de características operativas del receptor o curvas de iso-sensibilidad e iso-

sesgo (curvas COR) que relacionan gráficamente las probabilidades de falsas alarmas y de aciertos permitiendo un rápido análisis general de los resultados de la tarea. Asimismo, a partir del par que determina la tasa de aciertos y la de falsas alarmas se puede estimar una medida de la capacidad sensorial del individuo para detectar la señal y otra medida del criterio de decisión o sesgo de respuesta. Este último informaría sobre la inclinación del individuo a responder con una u otra respuesta, es decir el grado de exigencia sobre su experiencia sensorial para llegar a afirmar que la señal está presente.

Para el planteamiento original de la *TDS* las distribuciones de probabilidad asociadas a la señal y al ruido podían ajustarse al modelo Gaussiano de la normal, con varianzas equivalentes (Tanner & Swets, 1954). Así, la sensibilidad puede entenderse como una medida de la distancia entre las medias de ambas distribuciones que, en uno de los índices más utilizados, d' , se calcula como la distancia entre las puntuaciones típicas asociadas a la *TA* y *TFA* según la distribución normal:

$$d' = Z_{TFA} - Z_{TA}$$

En esta fórmula la falta de sensibilidad (la ausencia de discriminación entre señal y ruido) tiene un valor de cero (la distancia entre dos curvas normales superpuestas) y la sensibilidad creciente se traduce en un incremento en valores positivos de la misma.

Asimismo, con los datos procedentes de la *TA* y la *TFA*, Tanner y Swets (1954) proponían el índice Beta que representa el criterio de decisión y que se calcula como el cociente de las dos distribuciones de probabilidad en los puntos correspondientes a los valores de ambas tasas (el cociente entre el valor de la ordenada correspondiente a la puntuación típica asociada a la *TA* y el asociado a la de *FA*). No obstante, este índice tiene importantes desventajas desde el punto de vista interpretativo, lo que ha llevado a proponer dos alternativas, el criterio y el logaritmo neperiano de Beta, de los cuales optaremos por el segundo por ser de los más utilizados (Wickens, 2002):

$$\ln \beta = \frac{1}{2} (Z_{TA}^2 - Z_{TFA}^2)$$

En este caso, el criterio neutral, sin sesgos, lo refleja el valor 0 en la fórmula. Valores *positivos* indican un criterio *estricto*; esto es, el participante exige más a sus sentidos para afirmar que la señal está presente, lo que le

lleva a inclinarse hacia el “no” más que al “sí”, protegiéndose así contra las Falsas Alarmas (estas disminuyen respecto al participante que adopta un criterio neutral) pero inevitablemente disminuyendo también la de Aciertos. En contraposición, valores *negativos* de este índice revelan un criterio *relajado*, es decir una tendencia a contestar Sí con más ligereza que en el caso neutral, lo que aumentará los dos tipos de tasas, tanto la de Aciertos como la de Falsas Alarmas: una tendencia a potenciar el número de aciertos que arrastra un incremento de las Falsas Alarmas.

En el panel central de la Figura 1 se ilustra el cálculo de los índices d' y logaritmo de beta ($\ln\beta$), para cada uno de nuestros participantes, de acuerdo con las fórmulas explicitadas. Si atendemos primero al índice sensorial podemos observar que todos los participantes muestran aproximadamente el mismo valor de d' (2.32), lo que indica que muestran la misma sensibilidad a la detección del sonido en nuestro ejemplo y además según un valor muy elevado. Sin embargo, el valor del Logaritmo de Beta varía, siendo 0 para el participante 1, negativo para el participante 2 (-1.13) y positivo y simétrico respecto al anterior para el participante 3 (1.13). Así, de acuerdo con lo dicho hasta ahora, aunque los tres participantes son igualmente buenos en la detección del sonido desde un punto de vista puramente sensorial, el criterio de decisión sólo es neutral en el participante 1, siendo relajado en el participante 2 y estricto en el participante 3. Respecto a los dos últimos, manifiestan un sesgo del criterio equivalente en magnitud pero en direcciones opuestas.

Los resultados de este experimento hipotético se representan gráficamente de manera sencilla e ilustrativa mediante curvas COR. En el panel inferior de la Figura 1 aparece recogida la curva COR a la que se ajustan las ejecuciones de los tres participantes de nuestro estudio, que recordemos tenían la misma sensibilidad. Obsérvese que en la figura aparece un segmento diagonal (línea 1) que abarca desde el punto de ejecución mínima (0.0) hasta el de ejecución más extrema (1.1), en el vértice superior derecho. Los puntos situados sobre la línea 1 representan actuaciones al azar, pues las tasas de aciertos y de falsas alarmas son idénticas y nos servirá para delimitar los valores de sensibilidad. Por otro lado, obsérvese la recta (línea 2) que va desde el punto intermedio (0.5; 0.5) hasta el de ejecución perfecta (0.0; 1.0); los puntos situados a lo largo de esta recta corresponderán a un criterio decisorio neutral y por tanto servirán para

delimitar los valores de sesgo en el criterio.

En teoría, las ejecuciones deberían quedar circunscritas al área triangular que queda por encima de la línea 1, puesto que lo contrario correspondería a valores de *TFA* superiores al de la *TA*, lo que en principio es un contrasentido. Presumiblemente esto sucedería porque el participante está respondiendo al contrario de lo que se le pide (Wickens, 2002).

Las ejecuciones ubicadas en las áreas por encima o por debajo de la línea 2 representan una actuación sesgada. Los participantes cuya ejecución se sitúe en el área que queda a la derecha de la línea 2 están utilizando un criterio flexible, siendo poco exigentes para determinar que el estímulo está presente, por lo que tendrán muchos aciertos pero también muchas falsas alarmas. Por el contrario, los participantes cuya ejecución se sitúe en el área a la izquierda de la línea 2 están utilizando un criterio estricto, siendo muy exigentes para determinar que el estímulo está presente, por lo que tendrán muy pocas falsas alarmas pero también menos aciertos. Así, distintos sujetos pueden mostrar la misma sensibilidad para detectar la señal y sin embargo mantener criterios diferentes, lo que aparecería reflejado en la curva COR.

En la figura puede observarse que los participantes se sitúan sobre la misma curva COR lo que indica que presentan un nivel de detección idéntico. Sin embargo, el participante 2 está utilizando un criterio flexible, lo que le lleva a tener un número de aciertos alto, pero también a incrementar el de falsas alarmas. Mientras, el participante 3 está imponiendo un criterio estricto, lo que le lleva a tener muy pocas falsas alarmas, pero también a disminuir la tasa de aciertos. Este ejemplo da una idea de la utilidad de las curvas COR para realizar un análisis exploratorio rápido de los resultados de los experimentos de detección.

Aplicación de los principios de la TDS a la cata de aceite de oliva

Como señalábamos más arriba, el catador de aceite de oliva, al igual que el catador de otro tipo de alimentos, debe ser capaz de realizar estimaciones muy finas de distintos atributos sensoriales del alimento que es evaluado. Tal y como señalan Hughson y Boakes (2001, 2002) en sus estudios sobre la cata de vino, el catador debe ser capaz de detectar y discriminar entre compuestos –como sería el caso de la discriminación entre dos aceites muy parecidos pero de calidad distinta y debe ser también capaz de detectar y

distinguir entre elementos de un mismo compuesto —como la detección de los atributos positivos y negativos de un determinado aceite. En ambos casos nos encontramos con tareas discriminativas en las que el catador tiene que detectar la presencia o ausencia de distintos atributos perceptivos en situación de incertidumbre, exactamente como las situaciones de operario de radar que inspiraron la *TDS*, por lo que contar con una técnica que nos permita separar los factores puramente sensoriales, como el grado de similitud entre dos muestras de aceite, de factores internos o subjetivos, como las expectativas del sujeto o las consecuencias que puedan derivarse de su juicio, puede ser de gran trascendencia.

Como hemos visto en el apartado previo, la *TDS* nos proporciona las herramientas necesarias para separar entre estos dos procesos, superando las limitaciones de la psicofísica clásica de umbrales que impera en las situaciones de cata y según un modelo teórico que optimizará la medida de ejecución de los paneles de cata.

Sin embargo, a la hora de aplicar la *TDS* a la cata de aceite nos enfrentamos a un problema que proviene de las propias características del proceso. La cata depende de las habilidades gustativas y olfativas del individuo y estas habilidades son fácilmente saturables por la exposición repetida a los estímulos, lo que impone un número muy reducido de muestras para una misma sesión, entre 4 y 6 cuando se trabaja con aceites de oliva puros. Asimismo, la muestra de participantes es habitualmente pequeña.

La utilización de los índices clásicos de la *TDS*, d' y medidas derivadas de Beta, sólo está justificada en aquellas situaciones en las que los datos se ajustan a la curva normal o campana de Gauss, lo que impone una elevada cantidad de información (véanse los ejemplos representativos en Snodgrass & Corwin, 1998). En el caso de la cata de aceite de oliva el número de muestras es limitado para no saturar las vías sensoriales de los individuos con lo que resulta muy difícil comprobar la Normalidad de los datos. A fin de superar estas limitaciones Ramos-Álvarez, Callejas-Aguilera, Abad, Rosas y Paredes-Olay (2006, 2007) proponen la utilización de índices no paramétricos que no requieren del supuesto de la Normalidad de los datos. Específicamente, se seleccionaron los índices sistematizados por Grier (1971), A' y B' para el proceso sensorial y decisorio, respectivamente. La formalización se ha tomado del trabajo de Stanislaw y Todorov (1999), pues permite resolver algunas incoherencias matemáticas de la propuesta

original. Así, el índice A' se calcularía de acuerdo con la ecuación siguiente:

$$A' = 0.5 + \left[\text{signo} \{TA - TFA\} \frac{(TA - TFA)^2 + |TA - TFA|}{4 \cdot \text{Max} \{TA, TFA\} - 4 \cdot TA \cdot TFA} \right]$$

Los valores numéricos del índice A' oscilan entre 0 y 1, pero es el valor 0.5 el que representa la ejecución al azar (véase la parte izquierda de la ecuación) y el 1 representa la ejecución perfecta. Pensemos que este índice cuantifica el área bajo la curva COR, con lo cual los valores en el intervalo que abarca desde 0 hasta 0.5 corresponden a ejecuciones posibles desde el punto de vista empírico pero inadecuadas desde la perspectiva teórica (recuérdese el apartado precedente).

Por otra parte, el índice B' se calcula con la ecuación siguiente:

$$B' = \text{signo} \{TA - TFA\} \frac{TA \cdot (1 - TA) - TFA \cdot (1 - TFA)}{TA \cdot (1 - TA) + TFA \cdot (1 - TFA)}$$

Los valores del índice B' oscilan entre -1 y 1. El valor 0 representa el criterio neutral, mientras que valores positivos y negativos representan criterios estrictos y flexibles, respectivamente. Aplicando estos índices a los datos que presentábamos como ejemplo en la Figura 1 (véase la columna “No Paramétrico” en el panel central) podemos comprobar que, aunque la escala es distinta, no existen variaciones fundamentales con respecto a la actuación obtenida con los índices d' y Logaritmo de Beta. Así, el índice A' de los tres participantes se encuentra en torno a 0.92, reflejando un nivel sensorial similar. Asimismo, el índice B' presenta un valor neutro para el participante 1 (0.00), un valor negativo (-0.60) para el participante 2 y un valor positivo (+0.60) para el participante 3, reflejando los criterios flexible y estricto, respectivamente, que aparecían recogidos con los índices paramétricos. Los valores de estos índices son claramente convergentes con lo que apuntaban los índices paramétricos de la *TDS* (compárense las columnas Paramétrica y No Paramétrica de la Figura 1).

Teniendo en cuenta estas consideraciones, en nuestro laboratorio de la Universidad de Jaén hemos realizado una serie de experimentos con la intención de poner a punto una metodología experimental derivada de esta propuesta psicofísica. Así, el objetivo del trabajo experimental que recogemos en este capítulo fue, fundamentalmente, establecer una

metodología adecuada para el estudio de laboratorio de la cata de aceite de oliva virgen y comprobar la adecuación de la *TDS* para el estudio de los factores propios de la cata.

Específicamente, se evaluó la detección de tres concentraciones distintas de aceite de oliva diluido en parafina (0.4%, 0.8% y 1.6%), que habían sido seleccionadas partiendo de experimentos previos de umbrales. Esperábamos encontrar un aumento en la detección de la señal (índice A') mientras que el criterio de decisión (B') debería mantenerse neutral. En otras palabras, nuestra predicción es que las curvas COR para los 3 valores de concentración manipulados se diferenciarían a través de la línea 2 de la gráfica que presentábamos en el panel inferior de la Figura 1, de manera ortogonal a las curvas de los participantes hipotéticos que vimos en los apartados precedentes.

Método

Participantes

Los participantes en este estudio fueron 30 estudiantes de la Universidad de Jaén con edades comprendidas entre los 18 y los 28 años y sin experiencia previa con la tarea. La participación en los estudios fue voluntaria, recibiendo una pequeña compensación académica o económica por su participación (tarjetas de fotocopidora). Aproximadamente el 70% de los participantes fueron mujeres. Se aplicaron los principios del código deontológico de investigación psicológica, comenzando con un modelo de consentimiento informado estándar que utilizamos en todas nuestras investigaciones y en el que se indica a los participantes que pueden abandonar el experimento en cualquier momento.

Los participantes fueron previamente instruidos para acudir al laboratorio en las condiciones físicas adecuadas, esto es, descansados, sin fumar, sin haber comido y sin haber tomado excitantes al menos 30 minutos antes de la hora de la cita, posponiendo la participación en aquellos casos en los que existiesen alteraciones del gusto o del olfato relacionadas, por ejemplo, con procesos catarrales o alérgicos.

Estímulos y aparatos

Como estimulación de tipo ruido (recuérdense los tipos de ensayos según la

psicofísica contemporánea, tal y como se planteó en el segundo apartado) se utilizó parafina y como estimulación de tipo señal+ruido usamos una de tres disoluciones de aceite de oliva virgen extra de la variedad picual al 0.4%, 0.8% ó 1.6% en parafina. La parafina empleada en nuestro laboratorio es una sustancia apta para consumo, con textura similar a la del aceite alimentario pero con características organolépticas neutras.

Cada muestra de cata estaba compuesta aproximadamente por 5 ml. de la disolución correspondiente. Las muestras se presentaron en vasos de cata estándar, cubiertos por vidrios de reloj, colocados sobre láminas numeradas lo que permitía al investigador y a los participantes identificar adecuadamente el vaso de cata correspondiente.

Procedimiento

A su llegada al laboratorio, cada participante fue asignado completamente al azar a uno de los tres grupos (grupos 0.4, 0.8 ó 1.6), de manera que cada uno de los grupos estuvo constituido por 10 personas. Una vez en el laboratorio, recibieron instrucciones por escrito en las que se describía el procedimiento adecuado para catar una muestra de aceite de oliva:

“Para proceder a la degustación toma la copa manteniéndola cubierta con su vidrio de reloj, inclínala ligeramente, y en esta posición da un giro total a fin de de mojar lo más posible la superficie interior. Hecha esta operación, retira el vidrio, huele la muestra y pruébala. Bebe la muestra intentando distribuir el aceite por toda la cavidad bucal desde la parte anterior de la boca y la lengua, pasando por los laterales y la parte posterior, hasta los pilares del paladar y la garganta. Después de consumir cada muestra debes masticar un pequeño trozo de manzana y beber un poco de agua para limpiar el paladar.”

Una vez leídas las instrucciones y para asegurar que las mismas se habían entendido, los participantes recibían dos muestras de práctica con parafina, bajo la supervisión del experimentador. Una vez resueltas las dudas, este último abandonaba el área de experimentación situándose fuera de la vista de los participantes pero en un lugar desde el que podía controlar el desarrollo de la sesión, en caso de que surgiera cualquier problema.

Los participantes recibieron 20 muestras al 50%, diez con parafina alimentaria (ruido) y diez con una de las tres disoluciones de aceite de oliva (señal+ruido) entremezcladas de manera aleatoria. Después de catar cada una de las muestras, el participante tenía que responder con un “sí” o

un “no” ante la siguiente pregunta “¿Crees que la muestra que has tomado contiene aceite de oliva?”, anotando en una hoja de registro presentada a tal efecto.

Los grupos difirieron, pues, en la concentración de aceite de oliva diluido en parafina, 0,4, 0,8 ó 1,6%, respectivamente. En terminología psicofísica, los tres grupos difirieron en la concentración de la señal.

Variable dependiente y análisis estadísticos

Se registró el número de aciertos, fallos, falsas alarmas y rechazos correctos para cada participante, como en la matriz de confusión, y se procedió al cálculo de la tasa de aciertos y de falsas alarmas para cada uno. Para evitar la indeterminación matemática en el cálculo de los índices psicofísicos en aquellas situaciones en las que la tasa de respuestas es extrema (con tasa de aciertos máxima o tasa de falsas alarmas nula) se optó por una transformación logarítmico-lineal consistente en sumar 0,5 a todas y cada una de las frecuencias (para una justificación de esta corrección se puede consultar Snodgrass & Corwin, 1998, o Stanislav & Todorov, 1999). Para evitar la sensibilidad a los valores extremos propia de las pruebas estadísticas estándar, se utilizaron estadísticos robustos vinculados a la perspectiva de Análisis Exploratorio de Datos, calculando las estimaciones de las medianas y de su correspondiente error típico a partir del estimador Harrell-Davis $-HD$ en adelante – (Wilcox 1996, 2005; véase Ramos-Álvarez, 2009, para los detalles técnicos).

Resultados y discusión

En la Figura 2 se presentan las curvas COR para la ejecución mediana de cada uno de los tres grupos de participantes en este experimento. Tal y como puede apreciarse, a medida que aumenta la concentración de aceite de oliva aumenta la capacidad de los participantes para identificar la presencia de aceite de oliva en la muestra, con los puntos gradualmente más próximos al vértice superior izquierdo, en los grupos 0,4, 0,8 y 1,6, aunque la diferencia es de mayor magnitud entre el primero de los grupos y cualquiera de los otros dos. Además, estos puntos se sitúan muy próximos a la diagonal de referencia para el criterio, indicando que los participantes en este estudio respondieron de manera neutra, sin sesgos de respuesta en una dirección u otra.

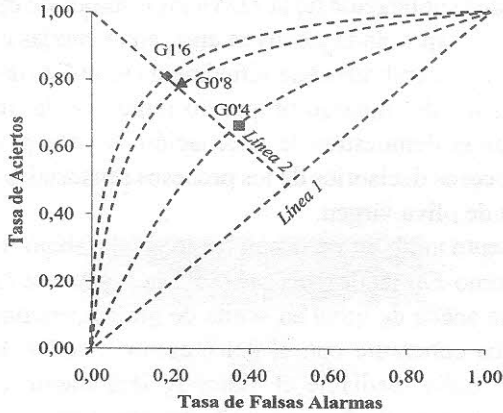


Figura 2. Curva COR obtenida para los participantes de los grupos 0'4, 0'8 y 1'6 en el estudio realizado.

Las apreciaciones que permite realizar la curva COR se confirman con el cálculo de los índices A' y B' para cada grupo de sujetos. Con respecto al índice A' , los estimadores HD de las Medianas \pm Error Típico asociado a los mismos fueron 0.732 ± 0.104 , 0.885 ± 0.039 y 0.895 ± 0.031 para los tres grupos de concentración creciente, respectivamente. Con respecto al índice B' , los estimadores $HD \pm$ Error Típico fueron -0.049 ± 0.035 , -0.095 ± 0.111 y -0.014 ± 0.071 para los tres grupos, respectivamente, es decir prácticamente cero en los tres casos.

El experimento revela un patrón de resultados poco sorprendente con respecto al índice de detección o sensibilidad, con un incremento en la sensibilidad a medida que aumenta la concentración de aceite de oliva. El método utilizado permitió la detección de variaciones sensoriales prácticamente en toda la escala de valores del índice empleado, desde un nivel cercano al azar en el caso del grupo 0.4 hasta una ejecución prácticamente perfecta en los grupos de concentración superior. Sin embargo, lo interesante de estos resultados es que muestran que este patrón se obtuvo en ausencia de sesgos subjetivos de los participantes en el estudio, los cuales manifestaron un criterio de decisión neutral (todos cercanos a cero).

Conclusiones

El planteamiento teórico con el que iniciamos el desarrollo de este capítulo

hacía presuponer que la aplicación de la *TDS* a los estudios de cata de aceite de oliva podría resultar apropiada y conveniente, con ventajas con respecto a la psicofísica clásica de umbrales que actualmente se utiliza dentro de este ámbito. Los resultados del método propuesto junto con la investigación experimental realizada demuestran la adecuación de este procedimiento para separar los procesos decisorios de los procesos sensoriales en el marco de la cata de aceite de oliva virgen.

Este procedimiento también se ha empleado en el trabajo realizado por Paredes-Olay, Moreno-Fernández, Rosas y Ramos-Álvarez (2009), pero con disoluciones de aceite de oliva en aceite de girasol, encontrándose un patrón de resultados coherente con el del presente estudio: la ejecución sensorial según se midió mediante el índice A' incrementó conforme lo hizo la concentración de aceite de oliva, toda vez que el criterio se mantuvo en valores neutrales en torno a cero.

Asimismo, en un estudio posterior, Ramos-Álvarez, Moreno-Fernández, Paredes-Olay y Rosas (2009) encontraron que la manipulación de instrucciones que indicaban al sujeto que mantuviera un criterio de detección flexible o estricto afectaban al índice del criterio de decisión B' sin afectar al índice A' , que mejoraba con el aumento de la concentración de aceite de oliva.

Los resultados del experimento que presentamos aquí se unen a otros resultados de la literatura demostrando la adecuación de la *TDS* en general y de los índices propuestos en particular para el estudio de la cata de aceite de oliva virgen. El empleo conjunto de índices de sensibilidad y de medida del criterio de respuesta permite conocer si un catador es bueno detectando el atributo objetivo, por ejemplo un defecto sensorial, aportando además información sobre su tendencia en la forma de responder, lo que supone una mejora con respecto al método tradicional recomendado por el COI, ya que este último, aún siendo riguroso, no permite conocer el efecto de este segundo factor.

Las implicaciones prácticas del empleo de esta aproximación metodológica para el mundo de la cata serían inmediatas, ya que permitiría optimizar tanto la selección como al entrenamiento de los catadores profesionales y en última instancia la calidad y fiabilidad del proceso de evaluación organoléptica del aceite de oliva; y es que saber que un determinado tipo de entrenamiento está mejorando la habilidad de los futuros catadores para identificar defectos o atributos positivos del aceite sin afectar al criterio de decisión o que la ejecución de un candidato a

catador es buena (sensible) pero carente de sesgos son, a nuestro modo de ver, aspectos de gran utilidad para el diseño, desarrollo y evaluación de programas de selección y entrenamiento apropiados.

Supongamos por ejemplo un panel de cata que está siguiendo un programa de entrenamiento para conseguir que los panelistas aumenten la capacidad para detectar un atributo que por las condiciones climáticas durante la cosecha puede ser fundamental para establecer la calidad de los aceites durante ese año. Supongamos además que ese método consiste en presentar durante la sesión de entrenamiento muestras de aceite de oliva que en su mayoría presentan el atributo en cuestión. Este procedimiento (que supone en términos de la *TDS* un aumento de la densidad de la señal) puede sesgar el criterio de respuesta de los catadores y determinar su respuesta; de manera que el uso de una medida que no permita tener en cuenta este sesgo limitará cualquier conclusión respecto al resultado y efectividad del entrenamiento.

En la actualidad, en nuestro laboratorio estamos explorando factores como la densidad de la señal o los incentivos, pues la literatura sugiere que pueden afectar al criterio de decisión dejando intacto el índice de discriminación (Blanco, 1996; Ramos et al., 2004). Asimismo, estamos examinando cómo afectan distintos factores relacionados con el aprendizaje discriminativo y perceptivo (véase Gibson, 1963; Hall, 1991) y distintas variables de entrenamiento (véase por ej., Liu, Mercado III, Church & Orduña, 2008; Scahill & Mackintosh, 2004) a la discriminación en la cata de aceite de oliva virgen. Los resultados del estudio que informamos aquí nos llevan a ser optimistas en cuanto a la contribución de la *TDS* al estudio científico de la cata de aceite de oliva virgen y a sus aplicaciones prácticas al propio diseño del proceso de cata, abriendo un terreno hasta ahora inexplorado desde la Psicología.

■ Referencias

- Blanco, M. J. (1996). *Psicofísica*. Madrid: Editorial Universitas.
- Gibson, E. (1963). Perceptual learning. *Annual Review of Psychology*, 14, 29-56.
- Green D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.

- Guerrero, L. (1996). Selección de catadores, por qué y cómo. *MERCACEI*, 25, 11-13.
- Gutiérrez, F. (1997). Análisis sensorial. En J. Alba, J. R. Izquierdo, & F. Gutiérrez, Aceite de oliva virgen. Análisis sensorial (pp. 49-81). Madrid: Editorial Agrícola Española.
- Hall, G. (1991). *Perceptual and associative learning*. New York: Clarendon Press/Oxford University Press.
- Harvey Jr., L. O., Hammond, K. R., Lusk, C. M., & Mross, E. F. (1992). The application of Signal Detection Theory to weather forecasting behavior. *Monthly Weather Review*, 120, 863-883.
- Hughson, A. L., & Boakes, R. A. (2001). Perceptual and cognitive aspects of wine expertise. *Australian Journal of Psychology*, 53(2), 103-108.
- Hughson, A. L., & Boakes R. (2002). The knowing nose: The role of knowledge in wine expertise. *Food Quality and Preference*, 13, 463-472.
- Hutchinson, T. P. (1981). A review of some unusual applications of Signal Detection Theory. *Quality and Quantity*, 15, 71-98.
- Lawless, H. T. (1985). Psychological perspectives of winetasting and recognition of volatile flavours. En G. Birch, & M. Lindley (Eds.), *Alcoholic beverages*. London: Elsevier Applied Science.
- Liu, E. H., Mercado III, E., Church, B. A., & Orduña, I. (2008). The easy-to-hard effect in human (*Homo sapiens*) and rat (*Rattus norvegicus*) auditory identification. *Journal of Comparative Psychology*, 122, 132-145.
- Marshall N. W. (2001). The practical application of Signal Detection Theory to image quality assessment in x-ray image intensifier-TV fluoroscopy. *Physics in Medicine and Biology*, 46, 1631-1649.
- Moskowitz, H. (1988). *Applied sensory analysis of foods*. Boca Ratón, FL: CRC Press.
- O'Mahony, M., & Hautus, M. J. (2008). The Signal Detection Theory ROC curve: Some applications in food sensory science. *Journal of Sensory Studies*, 2, 186-204.
- Paredes-Olay, C., Moreno-Fernández, M. M., Rosas, J. M., & Ramos-Álvarez, M. M. (2009). ROC curves in olive oil tasting: A Signal Detection Theory approach to tasting tasks. Manuscrito no publicado.
- Parr, W. V., Heatherbell, D., & White, K. G. (2002). Demystifying wine

- expertise: Olfactory threshold, perceptual skill and semantic memory in expert and novice wine judges. *Chemical Senses*, 27, 747-755.
- Ramos, M. M., Catena, A. y Trujillo, H. M. (2004). *Manual de métodos y técnicas de investigación en Ciencias del Comportamiento*. Madrid: Biblioteca Nueva.
- Ramos-Álvarez, M. M. (2009). Herramientas para la realización de estimaciones robustas en psicología. <http://www4.ujaen.es/~mramos/PsychoRobust.html>.
- Ramos-Álvarez, M. M., Callejas-Aguilera, J. E., Abad, M. J. F., Rosas, J. M., & Paredes-Olay, C. (2006). Aplicación de métodos psicofísicos al análisis de la cata del aceite de oliva virgen. XVIII Congreso de la Sociedad Española de Psicología Comparada. Málaga, España.
- Ramos-Álvarez, M. M., Callejas-Aguilera, J. E., Abad, M. J. F., Rosas, J. M., & Paredes-Olay, C. (2007). Análisis psicofísico del proceso de la cata de aceite de oliva virgen. X Congreso de Metodología de las Ciencias Sociales y de la Salud. Barcelona, España.
- Ramos-Álvarez, M. M., Moreno-Fernández, M. M., Paredes-Olay, C., & Rosas, J. M. (2009). Psychophysical evidence of dissociation between perception and decision processes within the context of organoleptic evaluation of olive oil. Manuscrito no publicado.
- Romero A., & Tous, J. (2003). Análisis sensorial de aceite de oliva virgen. Percepnet (recurso electrónico). Rubes.
- Scahill, V. L., & Mackintosh, N. J. (2004). The easy to hard effect and perceptual learning in flavor aversion conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 30, 96-103.
- Snodgrass, J. C., & Corwin, J. (1998). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology*, 117, 34-50.
- Stanislaw, H., & Todorov, N. (1999). Calculation of Signal Detection Theory measures. *Behavior Research Methods, Instruments and Computers*, 31, 137-149.
- Swets, J. A. (1996). *Signal detection theory and ROC analysis in psychology and diagnostics: Collected papers (Scientific Psychology Series)*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, 61, 401-409.

- Tous, J., & Romero, A. (2001). Evaluación sensorial de variedades de olivo. *Fruticultura profesional*, 120, 9-11.
- Wickens, T. D. (2002). *Elementary Signal Detection Theory*. Oxford: Oxford University Press.
- Wilcox, R. R. (1996). *Statistics for the Social Sciences*. London: Academic Press.
- Wilcox, R. R. (2005). *Introduction to robust estimation and hypothesis testing*. San Diego, CA: Academic Press.
- Zalevski, A. M., Henning, G. B., & Hill, N. J. (2007). Cue combination and the effect of horizontal disparity and perspective on stereoacuity. *Spatial Vision*, 20, 107-138.

Esta investigación ha sido financiada por los proyectos ACS/PP_2005 de la Universidad de Jaén y P06-HUM-01391 de la Consejería de Innovación, Ciencia y Empresa de la Junta de Andalucía.

La participación de M. M. Moreno-Fernández se financió a través de una beca del programa de FPU del Ministerio de Ciencia e Innovación del Gobierno de España (AP2007-01332).

Queremos agradecer los comentarios y sugerencias de M. J. F. Abad, J. E. Callejas-Aguilera y S. Parra a una versión previa de este manuscrito, así como su asistencia en la recolección de los datos.
