

UNIVERSIDAD DE COSTA RICA
FACULTAD DE CIENCIAS AGROALIMENTARIAS
ESCUELA DE TECNOLOGÍA DE ALIMENTOS

Trabajo Final de Graduación bajo la modalidad de Tesis de Graduación,
presentado a la Escuela de Tecnología de Alimentos para optar por el grado de
Licenciatura en Ingeniería de Alimentos

**Comparación del efecto del uso de una muestra iniciadora sobre la
sensibilidad y la potencia de seis distintas pruebas de discriminación en una
bebida de naranja**

Elaborado por:

María del Mar González Vargas

Carné: B02786

Ciudad Universitaria Rodrigo Facio

Mayo, 2020

TRIBUNAL EXAMINADOR

Proyecto de graduación presentado a la Escuela de Tecnología de Alimentos como requisito parcial para optar por el grado de Licenciatura en Ingeniería de Alimentos

Elaborado por:

María del Mar González Vargas

Aprobado por

M.Sc. Ana I. Incer González

Presidenta del Tribunal

Ph.D. Elba Cubero Castillo

Directora del Proyecto

M.Sc. Yorleny Araya Quesada

Profesora asesora del Proyecto

M.Sc. Marcia Cordero García

Profesora asesora del Proyecto

Lic. Andrea Marín Fonseca

Profesora Designada

DEDICATORIA

Esta tesis la quiero dedicar a varias personas

En primer lugar, a mi madre porque ella con su ejemplo me ha enseñado a luchar y superarme a pesar de los obstáculos y adversidades que hemos atravesado en la vida. Me ha dado fuerzas y me ha sostenido cuando me siento rendida y me ha apoyado y escuchado desde siempre. Gracias mami por enseñarme a ser una mujer independiente, bondadosa y con gran corazón, este trabajo es un reflejo de todo lo que has inculcado en mí. Lo logramos juntas madre mía.

A mis hermanos Susan, Juan y Andrea por que han sido mi luz desde que nací, mis confidentes y mis mejores amigos, me han apoyado, regañado y aconsejado para que pueda crecer y ser una mejor mujer y hermana.

A mis sobrinos Asher, Cate, Astrid y Amber, mis chiquitines a ustedes les quiero transmitir todo mi conocimiento y quiero recordarles que el camino no es fácil pero que si yo lo logré ustedes también lo harán.

A mi abuelito que siempre creyó en mi porque sus palabras cada vez que hablábamos del tema eran las siguientes “chica ese título la va llevar muy alto, así es como se hace estoy muy orgulloso. Apenas se gradúe vamos a poner un negocio en uno de los localillos”.

A mi papá por estar presente durante estos años dándome apoyo en mi crecimiento y también por participar en este proyecto y a mi hermano Edson.

A toda mi familia porque todos aportaron un granito de arena a lo largo de mi carrera y en el proyecto, especialmente mi prima Cheryl que ha sido como una hermana y confidente toda la vida.

Finalmente, a Dios por brindarme salud y resiliencia para sacar la tarea

LOS AMO MUCHÍSIMO A TODOS

AGRADECIMIENTOS

Agradecimiento absoluto a los 54 panelistas que hicieron realidad este trabajo, gracias por su disponibilidad, paciencia, actitud y gran corazón durante cada una de las sesiones que Dios les multiplique y les llene de bendiciones: Ana Laura, Sharon, Jorge, Stephanie, Ronald, Xime, Irene, Cristiana, Andrea, Jose, Leo, Adri, Lorena, Ana Paula, Felipe, Caro, Esther, Cata, Fer, Odette, Gloriana, Mawi, Montse, Magda, Guis, Priscila, Vale, Cate, Cheryl, Sebas, Ricky, Carol, Griselle, Susen, Cris, Juane, Andre, Juan B, Ana Li, Daniela, Adrián, Eduardo, Nati, Leandro, Michelle, Camilo, Elena, Mar, Claudio, Susana, Pablo, Vanessa, Ricardo, Álvaro.

A mi directora de tesis Elba Cubero por su arduo trabajo, dedicación y apoyo para poder darle forma a este trabajo desde el inicio hasta el fin, por compartir su gran conocimiento e impulsarme a investigar y pensar más allá de lo que creía posible y ponerle pasión al tema. Y sobre todo por tener paciencia conmigo.

También a mis profesoras asesoras Yorleny Araya y Marcia Cordero quienes estuvieron pendientes durante todo el proceso ayudando a mejorar e interpretar la investigación, así como los consejos a la hora de realizar las sesiones y los avances que se dieron a lo largo del trabajo.

A mis amigos Leo Barrantes, Andrea Alfaro, Sharon Silesky, Montse Muñoz, Eduardo Smith, Camilo Toro, Ana Laura Mora, Andrea Carrillo, Adriana Ramírez, Katherine Fernández por los momentos vividos que los atesoro muchísimo, por ayudarme y acompañarme durante toda mi carrera universitaria, así como en este proyecto en los días de paneles, también gracias todas por las terapias y por estar presentes durante todos estos años.

A Luis David por estar disponible siempre para abrirme el laboratorio a las 6 a.m durante muchos meses

A Giovanni González por todos sus consejos y ayuda desde que fui asistente de laboratorio hasta que logré culminar mi tesis

GRACIAS A TODOS, cada uno de ustedes ha dejado una huella en mí

ÍNDICE GENERAL

ÍNDICE DE CUADROS	vii
ÍNDICE DE FIGURAS	viii
ÍNDICE DE ABREVIATURAS	ix
RESUMEN.....	x
1. JUSTIFICACIÓN.....	1
2. OBJETIVOS	4
2.1 Objetivo General.....	4
2.2 Objetivos Específicos	4
3. MARCO TEORICO	5
3.1 Análisis sensorial aplicado en alimentos.....	5
3.2 Pruebas de discriminación sensorial.....	6
3.2.1 Generalidades.....	6
3.3 Clasificación de las pruebas de discriminación.....	8
3.3.1 Pruebas direccionadas	8
3.3.2 Pruebas no direccionadas	10
3.4 La sensibilidad de las pruebas de discriminación.....	12
3.5 Inconvenientes asociados a las pruebas de discriminación	12
3.6 Protocolos implementados para mejorar la sensibilidad de las pruebas de discriminación.	16
3.7 Análisis de datos en pruebas de discriminación.....	18
4. MATERIALES Y MÉTODOS	21
4.1 Localización	21
4.2 Materiales y Equipos	21
4.3 Preparación de las bebidas de naranja: bebida patrón y bebidas con diferente concentración de azúcar.	21
4.4 Metodología	22
4.4.1 Pruebas preliminares.....	22
4.4.2 Determinación de las concentraciones adecuadas para realizar las pruebas de discriminación direccionadas y no direccionadas	22
4.4.3 Determinación de las condiciones e instrucciones adecuadas para realizar las pruebas definitivas.....	26
4.5 Pruebas definitivas	26

4.5.1 Pruebas direccionadas	26
4.5.2 Pruebas no direccionadas	29
4.6 Análisis de datos	31
4.6.1 Análisis al pasar las rondas	31
4.6.2 Análisis por prueba y por condición.....	33
5. RESULTADOS Y DISCUSIÓN	36
5.1 Pruebas preliminares.....	36
5.1.1 Prueba de umbral de diferencia con estímulo constante (JND)	36
5.1.2 Prueba 2-AFC con diferentes concentraciones versus la disolución patrón	37
5.1.3 Prueba mediante paneles informales para definir las concentraciones para los panelistas en pruebas direccionadas y no direccionadas	38
5.2 Pruebas definitivas	39
5.2.1 Análisis al pasar las rondas mediante el índice Thurstoniano del grado de diferencia (d') y análisis cualitativo por codificación	39
5.2.2 Análisis por prueba y por condición por ordenamiento de rangos Newell McFarland y por análisis de Varianza	50
6. CONCLUSIONES	56
7. RECOMENDACIONES.....	56
8. REFERENCIAS BIBLIOGRÁFICAS.....	57
9. ANEXOS	59

ÍNDICE DE CUADROS

Cuadro I. Prueba preliminares realizadas para definir las concentraciones adecuadas que permiten discriminar la muestra más dulce en una bebida de naranja. _____ 22

Cuadro II. Determinación de la concentración más adecuada para percibir la diferencia en dulzor o intensidad respecto a una bebida de naranja estándar de 60 g/L de Tang® utilizando una prueba 2-AFC. _____ 37

Cuadro III. Valores del parámetro d' obtenidos al pasar las rondas para las pruebas 2-AFC, 3-AFC y Tétrada direccionada utilizando cuatro condiciones distintas $n=54$. _____ 40

Cuadro IV. Valores del parámetro d' obtenidos al pasar las rondas para las pruebas Dúo Trío, Triángulo y Tétrada utilizando cuatro condiciones distintas. _____ 45

Cuadro V. Comparación de los aciertos totales entre las pruebas 2-AFC, 3-AFC y Tétrada direccionada bajo cada condición en estudio mediante el método de ordenamiento por rangos con las tablas de Newell y McFarland (1987) con un $\alpha=0,05$ para 27 panelistas. _____ 50

Cuadro VI. Comparación de los aciertos totales entre las pruebas 2-AFC, 3-AFC y Tétrada direccionada bajo cada condición en estudio mediante el método de análisis de varianza con un $\alpha=0,05$ para 27 panelistas. _____ 51

Cuadro VII. Comparación de los aciertos totales entre las pruebas Dúo Trío, Triángulo y Tétrada no direccionada bajo cada condición en estudio mediante el método de ordenamiento por rangos con las tablas de Newell y McFarland (1987) con un $\alpha=0,05$ para 27 panelistas. _____ 52

Cuadro VIII. Comparación de los aciertos totales entre las pruebas Dúo Trío, Triángulo y Tétrada bajo cada condición en estudio mediante el método de análisis de varianza con un $\alpha=0,05$ para 27 panelistas. _____ 53

ÍNDICE DE FIGURAS

- Figura 1. Ejemplo de acomodo de datos según el número de aciertos que obtuvo un panelista para cada prueba, condición y ronda.*_____ 31
- Figura 2. Ejemplo de acomodo de datos para realizar el análisis cuantitativo por medio de los valores d' para cada una de las rondas por condición y prueba en estudio.* _____ 32
- Figura 3. Codificación generada en el análisis cualitativo de las pruebas finales.*_____ 33
- Figura 4. Ejemplo de acomodo de datos para realizar el análisis cuantitativo de ordenamiento de rangos Newell McFarland y ANDEVA.*_____ 34
- Figura 5. Porcentaje de personas que escogieron una concentración más ácida que la concentración estándar (60 g/L) en una prueba 2-AFC.* _____ 36
- Figura 6. Número de personas que mejoran, empeoran o no presenta ninguna tendencia al realizar las pruebas 2-AFC, 3-AFC y Tétrada direccionada bajo cuatro condiciones distintas.* 41
- Figura 7. Número de personas que mejoran, empeoran o no presenta ninguna tendencia al realizar las pruebas Dúo Trío, Triángulo y Tétrada bajo cuatro condiciones distintas.* _____ 46

ÍNDICE DE ABREVIATURAS

2-AFC= 2 Escogencia forzada alternativa

3-AFC= 3 Escogencia forzada alternativa

d' = distancia medida entre desviaciones estándar entre las medias de dos distribuciones normales que corresponden a las sensaciones de dos estímulos que se comparan en pruebas de discriminación

n= número de personas

p= probabilidad de escoger al azar la muestra correcta en una prueba de discriminación

α = probabilidad de cometer error tipo I, se concluye que hay diferencias cuando realmente no las hay

β = probabilidad de cometer error tipo II, se concluye que no hay diferencias cuando realmente las hay

RESUMEN

González Vargas, María del Mar

Comparación del efecto del uso de una muestra iniciadora sobre la sensibilidad y la potencia de seis distintas pruebas de discriminación en una bebida de naranja.

Licenciatura en Ingeniería de Alimentos- San José
González, M.; 2020
63 páginas, 7 il- 26 refs.

En este trabajo se comparó el efecto del uso de una muestra iniciadora sobre la sensibilidad y la potencia debido al tipo de prueba (2AFC, 3AFC, Tétrada especificada y Dúo Trío, Triángulo y Tétrada no especificada), la fatiga y las implicaciones de memoria en una bebida de naranja. Para ello se comparó el efecto de cuatro condiciones específicas para cada una de las pruebas en estudio: sin enjuague ni primer al inicio de cada set, solo con enjuague al inicio de cada set, con enjuague y “primer” al inicio de cada set y con enjuague entre cada una de las muestras de cada set.

Las muestras a comparar fueron dos bebidas de naranja de la marca Tang® con diferente grado de dulzor, la disolución estándar fue de 13 g/L de azúcar y la segunda disolución se definió mediante una prueba antes de la primera sesión de acuerdo con la sensibilidad de cada panelista, para las pruebas direccionadas se utilizaron 17, 19, 21 y 23 g/L de azúcar y para las no direccionadas 21, 23, 25 y 27 g/L de azúcar.

Las pruebas se presentaron siguiendo el orden descrito a continuación, el cual equivale a una repetición (ronda): 2-AFC, 3-AFC, Tétrada direccionada, Tétrada direccionada, 3-AFC, 2-AFC para las pruebas direccionadas y Dúo Trío, Triángulo, Tétrada, Tétrada, Triángulo, Dúo Trío para las pruebas no direccionadas. En total se realizaron 4 rondas o repeticiones para cada panelista con el fin de lograr un efecto de fatiga.

Las pruebas se realizaron con 54 participantes en total (27 para las direccionadas y 27 para las no direccionadas), cada uno de ellos realizó 4 sesiones (un panel

sensorial por cada condición que se desea estudiar). Para ello se agendaron citas, la duración de las sesiones estuvo sujeta a la condición evaluada y agilidad del panelista (20- 40) minutos, el tiempo promedio estimado resulto ser de 30 minutos.

Los datos se analizaron de dos maneras: una cualitativa y otra cuantitativa. El primero consistió en ver el comportamiento al pasar las rondas, para ello se analizó el número de aciertos que tuvo cada panelista en cada ronda cualitativamente mediante una agrupación en tres categorías (mejora o se mantiene al pasar las rondas, empeora o no siguen ninguna tendencia). También cuantitativamente utilizando un promedio para cada condición y prueba.

El segundo análisis se llevó a cabo por prueba y por condición utilizando una sumatoria del número de aciertos obtenidos en toda la sesión y se analizó utilizando estadística no paramétrica ordenamiento por rangos de Newell McFarland y estadística paramétrica ANDEVA.

En las pruebas direccionadas la 2-AFC fue la que presentó mejores resultados independientemente de la condición y las pruebas no direccionadas no se pudo ver una diferencia clara entre éstas, pero se esperaba que la tétrada fuera la que presentara mejores resultados ya que se ha comprobado que es más sensible que las otras dos, aunque se pudo afectar por el efecto de memoria. Respecto a las condiciones estudiadas, la que presentó mejores resultados fue la que no tenía enjuague ni “primer” para la mayoría de las pruebas. De acuerdo con lo anterior, se puede ver que no existe una condición universal que ayude a mantener o mejorar el desempeño de los jueces tanto en las pruebas direccionadas como las no direccionadas. Sino, que esto va depender de la potencia de la prueba y del número de estímulos que ésta presente.

Universidad de Costa Rica- Directora de tesis: Elba Cubero

PRUEBAS DE DISCRIMINACIÓN, PRIMER, PANEL, REPETICIONES, RONDAS

1. JUSTIFICACIÓN

El análisis sensorial ha tenido un gran desarrollo a través de los años, lo cual ha permitido la utilización de diferentes metodologías aplicables a la industria de alimentos. Varios científicos señalan que, al realizar investigación en dicho campo, la mejor manera de vincular adecuadamente sus resultados con el análisis sensorial de los alimentos es a través del uso de una matriz alimentaria lo más parecida a una solución estándar para que sea fácil de modificar, pero que cuente con características o atributos propios de los alimentos los cuales le confieren cierto grado de complejidad como lo son el color, sabor, textura, aroma, acidez, amargor, astringencia entre otros. Por esta razón, la mayoría de artículos científicos en esta área utilizan como productos bebidas en polvo de distintos sabores dentro de los cuales resaltan naranja y manzana (Rousseau *et al.*, 2002).

Las pruebas de discriminación se han vuelto una herramienta imprescindible para la industria alimentaria, ya que, al establecer diferencias entre productos muy similares entre sí, permiten asegurar estándares de calidad, desarrollar nuevos productos, realizar estudios de vida útil, reformulaciones, modificar procesos de producción entre otros. Estas pruebas permiten determinar el grado de diferencia entre dos estímulos confundibles o un único estímulo con diferente grado de intensidad (O'Mahony y Rousseau, 2003), de manera tal, que resulta difícil considerarlos iguales o diferentes entre sí (Angulo y O'Mahony, 2009).

Las pruebas de discriminación pueden dividirse en direccionadas y no direccionadas. Las primeras son aquellas en las que se utiliza una característica, ya que en las instrucciones se le indica al juez cuál es el atributo del alimento que debe encontrar en la prueba (muestra más salada, dulce, etc.). Entre las más conocidas resaltan 2AFC (2-escogencia forzada alternativa), 3AFC (3-escogencia forzada alternativa) y tétrada especificada. Son muy utilizadas en la industria alimentaria debido a que poseen una mayor potencia estadística en comparación con las pruebas no direccionadas, pero solo puede aplicarse cuando la naturaleza de la diferencia entre muestras es conocida (Xia *et al.*, 2015).

Por otro lado, en las pruebas no direccionadas el atributo del alimento no se le especifica al juez en las instrucciones, sino que éste debe identificar el atributo que genera diferencia. Dichas pruebas son utilizadas en la industria cuando se realizan cambios en productos o formulaciones y se modifican varios atributos a la vez. Algunas de las pruebas más utilizadas son dúo trio, triángulo y tétrada. Estas pruebas se consideran menos potentes; por lo tanto, es necesario utilizar un mayor número de jueces para alcanzar el mismo grado de significancia que las pruebas direccionadas (Xia *et al.*, 2015).

Las pruebas de discriminación pueden verse afectadas por diversos factores, especialmente si los panelistas realizan varias repeticiones. La cantidad de estímulos a degustar, dependiendo de la prueba, puede tener implicaciones de memoria afectando el rendimiento de éstas a través del tiempo. También, el orden secuencial de degustación explica que, en una secuencia de degustación tras probar el primer estímulo, este podría tener un efecto anulativo sobre el estímulo siguiente; entre otros factores importantes destacan sensibilidad de la prueba, efectos de fatiga, acarreamiento y adaptación (Lucak, 2008; Ishii *et al.*, 2014).

Ante las deficiencias que pueden obtenerse en las pruebas de discriminación, se han intentado implementar diferentes protocolos que ayuden a mitigar los errores generados en el análisis sensorial de alimentos. Uno de ellos es el uso de limpiadores de paladar ya sea antes de probar la primera muestra o entre cada una de las muestras de diferentes pruebas. Se han utilizado una amplia gama de matrices alimentarias como limpiadores de paladar, entre estos destacan: agua, agua carbonatada, agua mineral, soluciones con pectina, soluciones con etanol, galletas, pan, leche, cerveza, queso, zanahoria, pepino, entre otros (Lucak, 2008). Sin embargo, varias investigaciones concluyen que, para diferentes grupos de alimentos, el limpiador de paladar más viable es el enjuague con agua. Es importante esclarecer que hay discrepancias entre autores, ya que algunos señalan que el enjuague solo actúa como otro estímulo el cual podría interferir en la prueba y que la saliva es el mejor medio para limpiar el paladar (Lucak, 2008).

Otro protocolo que resalta en la literatura, es el uso de un iniciador o “primer”, el cual constituye una muestra del alimento que se degusta antes de iniciar la prueba de discriminación, pero ésta no es evaluada (Rousseau *et al.*, 2002; Xia *et al.*, 2014). Artículos científicos recientes han utilizado el iniciador antes de degustar la primera muestra con la intención de prevenir el efecto distorsionante que puede causar el enjuague de agua sobre la degustación de la primera muestra, con lo cual se podría inferir que el uso del iniciador ayuda a mejorar el desempeño en dichas pruebas (Rousseau *et al.*, 2002; Xia *et al.*, 2014). Sin embargo, actualmente no existe evidencia científica que permita afirmar si el uso de un iniciador causa un efecto o no en la sensibilidad y potencia de las pruebas de discriminación.

Basado en lo expuesto con anterioridad, se identifica la necesidad de comparar el efecto del uso de una muestra iniciadora sobre la sensibilidad y la potencia de pruebas de discriminación direccionadas (2AFC, 3AFC, Tétrada especificada) y no direccionadas (Dúo Trío, Triángulo, Tétrada).

2. OBJETIVOS

2.1 Objetivo General

Comparar el efecto del uso de una muestra iniciadora sobre la sensibilidad y la potencia debido al tipo de prueba (2AFC, 3AFC, Tétrada especificada y Dúo Trío, Triángulo y Tétrada), la fatiga y las implicaciones de memoria en una bebida de naranja.

2.2 Objetivos Específicos

- Determinar el efecto de una muestra iniciadora después de un enjuague con agua entre cada set sobre la potencia y la sensibilidad de las pruebas direccionadas y no direccionadas de manera que se considere la fatiga a lo largo de las pruebas.
- Evaluar el efecto de un enjuague con agua entre los estímulos de los sets y entre cada set sobre la potencia y la sensibilidad de las pruebas direccionadas y no direccionadas de manera que se considere el efecto sobre la memoria y la fatiga.
- Comparar el efecto de probar 2, 3 o 4 muestras por set sobre la memoria cuando no se usa enjuague entre sets, cuando se usa enjuague entre sets, cuando se usa un iniciador y enjuagues entre sets, y cuando se usa enjuagues entre los estímulos de cada set.

3. MARCO TEORICO

3.1 Análisis sensorial aplicado en alimentos

El análisis sensorial se define como una ciencia multidisciplinaria en la que a través de los sentidos que tienen los seres humanos se realizan mediciones de las características sensoriales y aceptación de los productos alimenticios. Es importante resaltar que no existe ningún otro instrumento que pueda sustituir o reemplazar la respuesta humana, por lo tanto, la aplicación de esta ciencia en alimentos es un factor esencial. Además, dicha aplicación tiene fundamento científico al ser respaldadas por diferentes áreas de estudio como psicología y estadística (Watts *et al.*, 1989; Olivas- Gastélum, 2008).

La evaluación sensorial de alimentos se realiza por medio de diferentes pruebas, existen tres tipos principales: pruebas de discriminación, descriptivas y afectivas. En las pruebas de discriminación se desea conocer que tan diferentes deben ser dos muestras entre sí para ser catalogadas como distintas, las pruebas descriptivas buscan definir las propiedades de un alimento y medirlas de la manera más objetiva posible, y las pruebas afectivas son las que buscan establecer el grado de aceptación de un producto a partir de la reacción del panelista (Olivas-Gastélum, 2008).

A la hora de realizar pruebas sensoriales es necesario conocer la información que se desea obtener, ya que cada tipo de prueba recopila información distinta sobre el alimento. La evaluación sensorial I, utiliza los sentidos humanos como una herramienta analítica para estudiar las propiedades de los alimentos, en este caso se busca utilizar la prueba más poderosa por sensibilidad y estadísticamente. Por otro lado, la evaluación sensorial II tiene como objetivo estudiar la sensibilidad de los consumidores para diferenciar entre dos productos, es decir, se evalúan los panelistas y no el producto, la sensibilidad de la prueba a utilizar debe de nivelarse a las condiciones típicas de consumo y, por ende, no necesariamente se utiliza la prueba que sea más sensible (Olivas - Gastélum, 2008; O'Mahony y Rousseau, 2003).

Dependiendo del objetivo de estudio sensorial, se debe hacer una selección rigurosa de los panelistas. En caso de que se desee obtener información sobre las características sensoriales de un alimento en particular, se requieren pruebas orientadas al producto y para ello es necesario utilizar panelistas entrenados. Por otro lado, si se desea obtener información sobre el agrado, desagrado y preferencia, se necesitan aplicar pruebas orientadas a los consumidores y, por tanto, se utilizan panelistas no entrenados, siempre y cuando éstos estén familiarizados con el producto (O'Mahony y Rousseau, 2003).

Es importante resaltar que para obtener resultados válidos y confiables en los análisis sensoriales se debe controlar la forma en que opera el panel, el diseño experimental utilizado, las pruebas sensoriales seleccionadas y un análisis estadístico adecuado que le aseguren al analista la obtención de resultados consistentes y reproducibles (Watts *et al.*, 1989).

3.2 Pruebas de discriminación sensorial

3.2.1 Generalidades

Las pruebas de discriminación sensorial se utilizan cuando el investigador desea conocer si dos muestras son perceptiblemente diferentes entre sí. No obstante, la diferencia entre las muestras debe ser confundible (Olivas - Gastélum, 2008).

En algunas ocasiones, la formulación química de dos muestras es diferente, sin embargo, los consumidores no tienen la capacidad para percibirlo. En el área de desarrollo de productos es deseable que, cuando se necesita realizar reformulaciones, o bien se quiera copiar un producto líder en el mercado, el consumidor no detecte diferencia alguna entre el producto actual y el producto reformulado o líder. Por otra parte, se desea que el consumidor perciba la diferencia cuando se reformula para crear un producto nuevo o mejorado (Lawless y Heymann, 2018).

Las pruebas de discriminación son catalogadas como pruebas simples en comparación con las técnicas de perfiles clásicos como el análisis descriptivo

cuantitativo, esto porque los métodos utilizados son más fáciles y rápidos de ejecutar; además requieren poco o ningún entrenamiento ya que en la práctica se realizan con consumidores e inclusive se pueden aplicar en niños. Sin embargo, la teoría asociada a éstas es compleja. Dichas pruebas son ampliamente aplicadas a alimentos tanto en la academia como en la industria, ya que se puede aplicar en diversas áreas como control de calidad, reformulaciones, cambios en proceso o para determinar la habilidad de los consumidores de discriminar entre dos productos (Ramírez, 2016).

La mayor utilidad de las pruebas de discriminación se reduce al uso de solamente dos productos, ya que cuando se utilizan más de dos productos resultan poco eficientes y carecen de fondo estadístico. Aunado a esto, una limitante de su uso es que dichas pruebas no ofrecen información sobre el grado de diferencia que existe entre las muestras que se desean analizar (Lawless y Heymann, 2018; Ramírez, 2016).

Por otra parte, es importante mencionar que cuando se realizan estudios con consumidores, existe la posibilidad que estos no cuenten con la habilidad para identificar atributos específicos o bien tengan una noción errónea de lo que un atributo significa provocando errores parciales o totales en los resultados (O'Mahony y Rousseau, 2003).

Las pruebas de discriminación tienen un problema cognoscitivo conocido como respuesta sesgada, la cual puede definirse como la tendencia de los jueces a responder en una forma particular de elección propia que no concuerda con la información sensorial dada para desarrollar una estrategia que le permita encontrar la respuesta correcta (O'Mahony, 1995). Dicho problema depende del tipo de prueba utilizada (O'Mahony y Rousseau, 2003), por lo tanto para minimizar el efecto de la respuesta sesgada suele utilizarse procedimientos de respuesta forzada (O'Mahony, 1995).

A la hora de seleccionar cuál prueba de discriminación se debe aplicar, un aspecto a tomar en consideración es la estrategia cognitiva que utiliza el cerebro

para encontrar una respuesta, ya que ésta varía tanto para pruebas direccionadas como no direccionadas. Por lo tanto, la prueba que se esté utilizando puede presentar mejores o peores resultados para encontrar diferencias entre muestras (O'Mahony y Rousseau, 2002).

3.3 Clasificación de las pruebas de discriminación

Las pruebas de discriminación se clasifican en dos grandes grupos, las denominadas direccionadas que se utilizan cuando la naturaleza de la diferencia entre muestras es conocida, en esta se identifica un atributo como la fuente de posible diferencia; mientras que las pruebas no direccionadas buscan establecer diferencia independientemente del atributo o razón por la cual ésta se podría generar (McClure y Lawless, 2010). Si el investigador conoce la diferencia que existe entre las dos muestras se pueden utilizar ambos tipos de prueba, pero si la diferencia no se conoce o se debe a más de un atributo solo se pueden utilizar las pruebas no direccionadas (Olivas-Gastélum, 2008).

A pesar de que los protocolos son diferentes, todas estas pruebas comparten el mismo principio, en el cual los panelistas deben agrupar o separar los productos según su similitud o diferencia. En caso de que se les pida encontrar el producto diferente del conjunto de productos que se les presenta, las respuestas obtenidas serán correctas si logra encontrar el producto diferente o en caso contrario incorrectas. Los datos obtenidos son binarios: cuando el evaluador acierta tiene valor de 1, cuando no, tiene valor de 0, valores "dummy" (Ramírez, 2016).

3.3.1 Pruebas direccionadas

Son aquellas que se utilizan para probar diferencias especificadas, ya que en las instrucciones se le indica al panelista buscar un atributo o característica definida, por ejemplo, la muestra más dulce, salada, entre otras. Estadísticamente estas pruebas son más potentes que las no direccionadas (Olivas- Gastélum, 2008; O'Mahony y Rousseau, 2003). Seguidamente se presenta una breve descripción de cada prueba.

2AFC

En esta prueba se le presenta al panelista dos muestras distintas, y se le solicita escoger aquella que presente un atributo específico en mayor o menor intensidad. Los dos órdenes de presentación son AB y BA (donde A representa la muestra más intensa y B la menos intensa) los cuales se deben balancear durante el experimento y la probabilidad de acertar al azar es $p=1/2$ (Ennis *et al.*, 2014).

3AFC

En dicha prueba se presenta a los panelistas tres muestras, dos iguales y una diferente, y se le solicita escoger la muestra que presenta un atributo en específico con mayor o menor intensidad. Los posibles órdenes de presentación son seis: AAB, BBA, ABA, BAB, ABB, BAA (donde A representa la muestra más intensa y B la menos intensa), y se deben balancear durante el experimento. La probabilidad de escoger la respuesta correcta al azar (adivinar) es $p=1/3$ (Ennis *et al.*, 2014).

4AFC

En esta prueba se presentan cuatro muestras, tres iguales y una diferente, y se le solicita al panelista escoger la muestra que presenta un atributo específico en mayor o menor intensidad. Los posibles órdenes de presentación son ocho: BABB, AB BB, BBAB, BBBA, ABAA, AABA, AAAB (donde A representa la muestra más intensa y B la menos intensa), y se deben balancear durante el experimento. La probabilidad de escoger la respuesta correcta al azar (adivinar) es $p=1/4$. Cabe destacar que esta prueba proporciona un aumento trivial en la potencia de la prueba lo cual no resulta práctico y por eso no es de uso común (Ennis *et al.*, 2014).

Tétrada direccionada o especificada

En esta prueba se presenta a los panelistas cuatro muestras, dos de un grupo y dos de otro grupo, y se les solicita identificar las dos muestras que presentan el atributo específico en mayor o menor intensidad. Los posibles órdenes de presentación son seis: AABB, BBAA, ABBA, BAAB, ABAB, BABA (donde A

representa la muestra más intensa y B la menos intensa) y se deben balancear durante el experimento. La probabilidad de escoger la respuesta correcta al azar (adivinar) es $p=1/6$ (Ennis *et al.*, 2014).

3.3.2 Pruebas no direccionadas

En estas pruebas el atributo del alimento que podría causar la diferencia no se le especifica al juez en las instrucciones, sino que éste debe encontrarlo por sí solo. En las condiciones ordinarias de consumo, generalmente las diferencias entre atributos en los alimentos no se suelen especificar, ya que, al realizar cambios en productos alimenticios, se modifican varios atributos a la vez; por lo tanto, estas pruebas se suelen aplicar en control de calidad para determinar si las diferencias entre productos se deben a cambios de ingredientes, procesamiento y almacenamiento. También se utilizan para juzgar o establecer diferencias entre productos entre los cuales no se conoce bien cuál es la diferencia, entre otros (O'Mahony y Rousseau, 2003; O'Mahony y Lee, 2004). Seguidamente se dará una breve descripción de cada una de estas pruebas.

Dúo Trío

En dicha prueba se presentan tres muestras, una de ellas se cataloga como referencia y los dos restantes se identifican con códigos, se le solicita al panelista escoger la muestra codificada que es igual a la referencia. Los dos posibles ordenes de presentación son AB y BA y se balancean a lo largo del experimento. La muestra de referencia puede ser cualquiera de estos dos estímulos. La probabilidad de escoger la respuesta correcta al azar es $p=1/2$ (Ennis *et al.*, 2014).

Triángulo

En esta prueba se presentan tres muestras, dos iguales y una diferente, y se le solicita al panelista escoger la muestra diferente. Los seis posibles órdenes de presentación son AAB, BBA, ABA, BAB, BAA, ABB y deben balancearse a lo largo del experimento. La probabilidad de escoger la muestra diferente al azar (adivinar) es $p=1/3$ (Ennis *et al.*, 2014).

Tétrada no especificada

En esta prueba se presentan cuatro muestras, dos de un grupo de productos y los dos restantes de un grupo distinto. Se le solicita al panelista agrupar en parejas según su similitud de sabor. Los seis posibles órdenes de presentación son AABB, BBAA, ABBA, BAAB, BABA, ABAB y deben balancearse durante todo el experimento. La probabilidad de acertar al azar es de $p=1/3$ (Ennis y Jesionka, 2011).

Recientemente el uso de la prueba tétrada ha tomado auge en pruebas de discriminación debido a que en diferentes investigaciones se ha comprobado que presenta un mayor desempeño o potencia (capacidad de encontrar la diferencia) en el panel sensorial, es decir se suele obtener un mayor número de aciertos en las pruebas sensoriales y detecta diferencias sensoriales muy pequeñas en comparación con otras pruebas no direccionadas como triángulo y dúo trío (Ishii *et al.*, 2014; Xia *et al.*, 2015).

ABX

Se presentan tres muestras, dos de referencia o estándar, una corresponde al control y otra es la matriz modificada que por lo general tiene un cambio en su formulación o procesamiento, la tercera muestra no rotulada es igual a una de las muestras de referencia mencionadas anteriormente. Se le indica al panelista que especifique cuál de las dos muestras de referencia es igual a la tercera. La probabilidad de acertar al azar es de $p=1/3$ (Lawless y Heymann, 2018).

Igual – Diferente

Se presentan tres muestras iguales y una diferente agrupadas en parejas, se le solicita al panelista que indique si percibe los pares de muestra dados como iguales o diferentes. Los posibles ocho ordenes de presentación son los siguientes: AA BA, AA AB, AB AA, BA AA, BB BA y BA BB los cuales deben presentarse aleatoriamente durante el experimento. Cabe resaltar que esta prueba no usa el análisis binomial debido a que produce una respuesta sesgada, por lo tanto, no se

puede calcular la probabilidad de acertar al azar porque hay muchas categorías (Lawless y Heymann, 2018).

3.4 La sensibilidad de las pruebas de discriminación

En ciertas pruebas de discriminación, la información se procesa de manera más eficiente que en otras, por lo tanto, el desempeño de los jueces varía en función de la prueba utilizada. Esta diferencia se explica por la forma en que el cerebro procesa la información que recibe de los sentidos, particularmente del gusto (Ramírez, 2016).

Cuando se aplican diferentes pruebas de discriminación para evaluar el mismo par de estímulos, algunas permiten que los jueces logren discriminar mientras que con otras los jueces no reportan diferencias entre los estímulos. Esto se conoce como la paradoja de las pruebas de discriminación, la cual se relaciona con la estrategia cognitiva usada por el cerebro durante la ejecución de las pruebas (Angulo y O'Mahony, 2009).

Las pruebas especificadas como 2-AFC y 3-AFC utilizan la estrategia denominada descremado. En este caso el juez escoge la muestra con mayor intensidad de percepción como la diferente (Angulo y O'Mahony, 2009). Por otro lado, pruebas no direccionadas como triángulo utilizan la estrategia de comparación de distancias, la cual señala que cada muestra tiene dos distancias, cada una de ellas corresponde a la separación existente con las otras dos muestras; la muestra que tenga ambas distancias con los valores más altos es aquella que se ubica lejos de las otras dos muestras, por lo tanto, será la diferente (O'Mahony, 1995). Se ha determinado que la técnica del descremado resulta más eficiente que la comparación de distancias, y requieren un menor número de panelistas para detectar el mismo grado de diferencia (Angulo y O'Mahony, 2009). Es decir, son más potentes.

3.5 Inconvenientes asociados a las pruebas de discriminación

Debido a que las pruebas de discriminación no presentan la misma sensibilidad, se han desarrollado diferentes teorías y modelos. El modelo Thrustoniano es uno

de los más conocidos y aunque originalmente se aplicó a la prueba triangular, dúo trío y prueba de comparación pareada (2-AFC selección forzada de dos alternativas), hoy en día se han desarrollado modelos similares univariados y multivariados para otras pruebas de diferencia, entre ellas la tétrada (Angulo y O'Mahony, 2009).

Estos modelos permiten realizar el cálculo del parámetro d' (d prima) que se conoce como el índice Thurstoniano de grado de diferencia; dicho cálculo se realiza a partir de la proporción de respuestas correctas obtenidas en la prueba de diferencia en estudio. Asimismo, estos modelos hicieron posible el desarrollo de tablas que relacionan la proporción de respuestas correctas y el parámetro d' (O'Mahony y Rousseau, 2003).

El modelo Thurstoniano está basado en la variabilidad de percepción sensorial, que toma lugar al evaluar sensorialmente un estímulo, ya que la degustación del mismo es enviada al cerebro a través del sistema nervioso, el cual presenta fluctuaciones en el número de nervios estimulados por la sensación y la intensidad de la degustación debido a las diferentes interacciones que se dan en la boca, lo que genera que algunas degustaciones sean más intensas que otras, pero la intensidad promedio se presentará con mayor frecuencia. Gráficamente este comportamiento se representa mediante una distribución normal (Angulo y O'Mahony, 2009).

Al comparar entre dos o más estímulos similares, las distribuciones de frecuencia de cada estímulo se sobreponen entre si y por lo tanto son confundibles, mientras que en caso contrario no hay sobre posición y la distancia entre dichas distribuciones aumenta infiriendo así, que son estímulos fáciles de distinguir. Debido a esto el nivel de separación entre las medias de las distribuciones de los estímulos se utiliza como una medida para verificar que tan bien discriminan los jueces los estímulos en estudio, dicha distancia es el parámetro d' ; entre mayor sea este valor, mayor será el grado de diferencia entre los estímulos (Angulo y O'Mahony, 2009).

El valor d' disminuye significativamente al aumentar la desviación estándar o varianza de las distribuciones de los estímulos. Esto se debe a la presencia de diferentes variables experimentales como el grado de reacción del sistema gustativo de cada juez, interacciones en la cavidad oral como fluidos orales, diferente velocidad de secreción de la saliva, residuos de degustaciones previas. También, el orden secuencial en el que se presentan las degustaciones que, dependiendo de la intensidad de la primera degustación, hace que el segundo alimento se pueda percibir más o menos intenso de lo que debería (Angulo y O'Mahony, 2009).

Otra variable que influye significativamente es el efecto de la memoria. En varias investigaciones se mostró que las pruebas que utilizan menos estímulos pueden tener una ventaja práctica sobre aquellas pruebas que teóricamente presentan mayor potencia. Por esta razón, es que a pesar de que la prueba tétrada posee mayor potencia que la prueba triángulo (Ishii *et al.*, 2014), el desempeño de esta puede verse negativamente afectado debido a la adición de un cuarto estímulo comparado con los tres estímulos presentes en la prueba triángulo, ya que cuatro productos implican una carga mayor en la memoria del juez lo que puede ocasionar una reducción en el desempeño de la prueba debido a pérdida de memoria entre degustaciones (Xia *et al.*, 2014).

Las tablas que relacionan el valor d' con la proporción de respuestas correctas para diferentes pruebas de discriminación permite realizar comparaciones entre ellas, aunque las probabilidades de respuesta aleatoria sean diferentes. Por lo tanto, el parámetro d' es una medición imprescindible en las pruebas de discriminación ya que es independiente de la metodología empleada (Angulo y O'Mahony, 2009).

El otro modelo de gran importancia en pruebas de discriminación es el Modelo Secuencial Sensitivo, que considera la secuencia de degustación de los estímulos y cómo, después de degustar un estímulo, puede darse un efecto anulativo sobre el siguiente (O'Mahony, 1995).

Considerando dos estímulos a discriminar uno débil (D) y uno fuerte (F), tanto para la prueba tétrada como la prueba triángulo existen seis posibles órdenes de

presentación respectivamente DDFF, FFDD, DFFD, FDDF, DFDF y FDFD y para la segunda prueba DDF, FDD, DFD, FDF, FFD, DFF (O'Mahony, 1995).

Aquellos órdenes de presentación en los que un estímulo fuerte sea probado inmediatamente después de otro fuerte (FF) no se degustará de manera tan fuerte como un estímulo fuerte precedido de uno débil (DF); esto se debe a que el primer estímulo será el que adapte el sistema sensorial y la pérdida de sensibilidad resultante hará que el segundo estímulo sea percibido como menos fuerte. Entre más fuerte sea el primer estímulo mayor será el efecto de adaptación y menor será la sensación del segundo estímulo (O'Mahony, 1995).

De acuerdo con lo anterior, en ambas pruebas será más fácil acertar con los órdenes de presentación que inicien con el estímulo débil seguido del fuerte y no al revés; por lo tanto, se podría inferir que para aquellos órdenes de presentación que inicien con el estímulo fuerte existe una mayor probabilidad de contestar incorrectamente debido a fenómenos como supra y subadaptación del sentido del gusto, dilución del estímulo por la saliva y respuesta sesgada. Por esta razón es de gran importancia considerar las implicaciones de dicho modelo, ya que las pruebas tétrada y triángulo tienden a dar variaciones en el valor d' dependiendo del orden secuencial de degustación (O'Mahony, 1995). El uso de una muestra iniciadora podría afectar positivamente esta secuencia de pruebas, lo mismo que un enjuague con agua al inicio de cada set, ya que cambia el ambiente de los receptores logrando colocarlos en nivel base de adaptación antes de que prueben la primera muestra a degustar (Lucak, 2008).

Para cualquier prueba con una hipótesis nula H_0 (producto X= producto Y), existen dos tipos de errores que se pueden cometer. El primero es el error tipo I (α) que ocurre cuando se rechaza la hipótesis nula y esta es correcta, es decir, se percibe diferencia entre los productos cuando no la hay. El otro error es el tipo II (β) que se refiere al riesgo de no encontrar una diferencia cuando en realidad existe (Ramírez, 2016).

La potencia de la prueba ($1 - \beta$) se define como la capacidad que tiene la prueba para detectar de forma confiable diferencias entre los productos, estadísticamente permite expresar la probabilidad de encontrar una diferencia sensorial cuando ésta realmente existe. La potencia depende de la magnitud de la diferencia (d'), error tipo I (α) y el número de jueces. Este parámetro es indispensable a la hora de realizar el diseño de la prueba, ya que los investigadores determinan a priori el tamaño de la diferencia que sea significativo para ellos y realizan un análisis de potencia para escoger el tamaño de muestra adecuado (Angulo y O'Mahony, 2009; Ramírez, 2016).

3.6 Protocolos implementados para mejorar la sensibilidad de las pruebas de discriminación.

En las pruebas de discriminación usualmente se desean evaluar dos productos. Generalmente cuando el panelista prueba el primer producto inmediatamente se empieza a adaptar a éste, incluso si la muestra es tragada o expectorada, los residuos del producto permanecen en la boca provocando este efecto. Asimismo, dependiendo del orden en que se presenten las muestras y la cantidad de estímulos en la prueba se puede ver afectada la percepción de intensidad entre los productos, impactando negativamente el desempeño de los panelistas en encontrar la respuesta correcta (Lucak, 2008).

Por esta razón, es deseable encontrar un procedimiento que le permita al panelista regresar a su nivel base de sensibilidad donde se obtienen los niveles adecuados de intensidad después de probar la primera muestra. Para ello se ha estudiado el uso de diferentes limpiadores de paladar. El efecto residual de una muestra en el paladar se ha investigado utilizando una variedad de medidas psicofísicas entre las cuales destacan umbrales, clasificación de intensidad y estimación de magnitud (Lucak, 2008).

En un estudio se examinó el efecto residual de una solución con cloruro de sodio y cloruro de litio y se concluyó que los enjuagues requeridos para eliminar la sal residual y regresar el paladar al estado base exceden el número comúnmente usado en evaluaciones sensoriales. Otra investigación en la cual se midió la pérdida

de sensibilidad de una solución de cloruro de sodio por el método de umbrales determinó que el uso de dos enjuagues de agua seguidos de un minuto de descanso fue efectivo para contrarrestar la pérdida de sensibilidad y además este procedimiento mostró un menor efecto de adaptación en los panelistas (O'Mahony, 1972; O'Mahony y Godman, 1974).

Aunado a esto, también se ha utilizado el enjuague con agua no solo antes de probar la primera muestra, sino entre cada uno de los estímulos que se presentan en una determinada prueba. Para ello se utilizó el método de umbrales en una solución de cloruro de sodio a diferentes concentraciones bajo tres distintos procedimientos entre cada uno de los estímulos (15 segundos de descanso, 2 minutos de descanso y 15 segundos de descanso con un enjuague con agua posteriormente expectorado), los menores valores de umbral se obtuvieron al aplicar el procedimiento con enjuague (O'Mahony, 1972). Sin embargo, es necesario esclarecer que este protocolo podría no ser efectivo para todas las pruebas de discriminación, ya que aquellas que presenten un mayor número de estímulos como la tétrada o triángulo podrían verse afectadas por implicaciones de memoria (O'Mahony, 2018).

Varios investigadores han estudiado el uso de ciertos limpiadores de paladar, los cuales han sido utilizados para grupos específicos de alimentos. A pesar de que los alimentos se pueden clasificar de muchas maneras, para efectos de los limpiadores de paladar se han dividido en siete categorías: dulce, amargo, grasos, astringentes, picantes, fríos y alimentos no prolongados (Lucak, 2008)

Lucak (2008) seleccionó el mejor limpiador de paladar para cada categoría de alimentos y probó la efectividad de cada uno de ellos combinándolos con todas las categorías descritas en párrafos anteriores, esto con el fin de definir un posible limpiador de paladar universal para utilizar con varios alimentos. Los limpiadores utilizados fueron solución de pectina, agua, leche entera, chocolate, agua tibia y galletas tipo crackers seguidas de enjuague con agua. Éste último fue el más efectivo para las siete categorías de alimentos. Sin embargo, la autora señala que en términos generales la efectividad del limpiador depende de varios factores como

atributo particular y tipo de alimento que se probará, la cantidad de saliva de cada panelista y la cantidad de limpiador que utilizará a la hora de realizar la prueba.

Recientemente en diferentes estudios se ha puesto en práctica el uso de un “primer” o iniciador, éste corresponde a un estímulo que no forma parte de la prueba pero que se debe probar y luego expectorar antes de probar el primer estímulo de la prueba. Se utiliza con el fin preparar al paladar, adaptarlo al producto que se desea probar y eliminar las diferencias de percepción que se pueden dar entre el primer estímulo y los siguientes (Lee *et al.*, 2009). Aunado a esto, en caso de que se utilice un enjuague inicial, el primer se utiliza con la intención de prevenir el efecto distorsionante (por dilución) que puede causar el enjuague de agua sobre la degustación de la primera muestra (Rousseau *et al.*, 2002; Xia *et al.*, 2014). Sin embargo, es necesario realizar más investigaciones que sustenten la validez de esta práctica, ya que, su efectividad aún no ha sido comprobada.

Por otro lado, algunos autores señalan que el uso de enjuagues, tanto al inicio como entre cada uno de los estímulos de cada prueba, no son efectivos para limpiar el paladar y que más bien pueden afectar la memoria. Éstos argumentan que la saliva es el mejor medio para limpiar el paladar, por lo que infieren que las pruebas se deben aplicar utilizando únicamente los productos que se desean discriminar (Lucak, 2008).

3.7 Análisis de datos en pruebas de discriminación

Los datos de los estudios sensoriales pueden clasificarse según las siguientes escalas: nominal, ordinal, de escala y de razón. El tipo de análisis estadístico de los datos está sujeto al tipo de escala elegido, la cual se selecciona después de analizar y determinar los objetivos del estudio (Villa, 2014).

Los datos provenientes de las pruebas de discriminación se clasifican en número nominales, es decir, se acomodan en dos categorías: las personas que acertaron la prueba y las que no lo hicieron. Estos datos se analizan con estadística no paramétrica con prueba binomial. También se pueden analizar utilizando

estadística paramétrica si se calcula el d' , que es un número continuo (Paiz y Bustos, 2009).

Una prueba estadística paramétrica se identifica por tener un modelo en el cual se especifican ciertas condiciones (normalidad, medias iguales y homogeneidad de varianzas) que deben tener los parámetros de la población de la que se obtuvo la muestra investigada, además se requiere que los valores o puntajes analizados sean producto de una medición fuerte, y por ello se utilizan datos de escalas de intervalo y razón (Paiz y Bustos, 2009).

El análisis de varianza (ANDEVA) es la prueba paramétrica que frecuentemente se utiliza para los datos sensoriales expresados en escala de intervalos o racionales. Este método estadístico busca determinar si diversos conjuntos de muestras aleatorias de una determinada variable forman parte de la misma población o poblaciones. Cada conjunto de muestras se ve afectado por un tratamiento específico, que puede influir en los valores que tome la variable objeto de estudio (Fernandez, s. f.).

Por otro lado, una prueba estadística no paramétrica es aquella en la que su modelo no especifica condiciones de los parámetros de la población de la que se sacó la muestra, el tipo de mediciones en este caso no deben ser tan fuertes y generalmente se utilizan datos de una escala ordinal y en algunos casos de una escala nominal. Cabe resaltar que los métodos no paramétricos permiten un grado de discriminación menor, pero tienen la ventaja de que no requieren que los datos tengan una distribución determinada (Paiz y Bustos, 2009; Villa, 2014).

Cuando se trata de pruebas no paramétricas, los datos obtenidos suelen analizarse con la prueba de Friedman o Kramer que se basa en el análisis mediante la suma de los rangos de las muestras. Sin embargo, se ha observado que esta prueba no es recomendable, ya que las sumas de los rangos de las muestras diferentes no son independientes entre sí y el nivel de probabilidad de la prueba es incorrecta (Newell y MacFarlane, 1987). Ante esto se suele aplicar otro procedimiento llamado ordenamiento por rangos y suele ser de dos tipos: el primero

es la comparación entre todas las muestras entre sí que busca discernir aquellas muestras que son superiores o inferiores a otras muestras y el segundo es una comparación entre la referencia y varias muestras, donde se desea probar si una referencia es superior o inferior dentro de un grupo de muestras o si es diferente. Para dicho análisis se utilizan las tablas obtenidas de Newell y MacFarlane (Severiano *et al.*, s. f.).

Dado que el análisis con estadística no paramétrica es más limitado y menos potente y que, por otro lado, el análisis estadístico paramétrico es más potente y robusto, en el análisis sensorial de datos se suelen aplicar los dos en conjunto, con el fin de contrarrestar los efectos negativos que ambos métodos presentan, concluyendo con ambos resultados.

4. MATERIALES Y MÉTODOS

4.1 Localización

Las pruebas sensoriales que se realizaron en este trabajo final de graduación se llevaron a cabo en su mayoría en el Laboratorio de Análisis Sensorial de la Escuela de Tecnología de Alimentos de la Universidad de Costa Rica, sede Rodrigo Facio y también se realizaron a domicilio en diferentes áreas de la provincia de Heredia.

4.2 Materiales y Equipos

La bebida de naranja se preparó a partir de una bebida en polvo de naranja de marca Tang®. El dulzor fue el atributo definido para discriminar entre las muestras, esto se realizó mediante la adición de azúcar de caña refinada de marca Doña María®. Para las mediciones de masa se emplearon balanzas granatarias y para la preparación de las disoluciones se utilizaron balones aforados de (500-1000) mililitros.

Se hizo uso de un software de análisis sensorial y pruebas de consumidores marca FIZZ®, mediante el cual se generaron los códigos con números al azar y balanceo de órdenes de presentación para cada tipo de prueba. Posteriormente se elaboró una hoja de respuestas de manera manual para cada condición y cada panelista.

4.3 Preparación de las bebidas de naranja: bebida patrón y bebidas con diferente concentración de azúcar.

Cabe mencionar que tanto para la bebida de naranja patrón como para las bebidas de naranja con diferente concentración de azúcar fue necesario realizar varias pruebas preliminares para definir la formulación, dichas pruebas se describen a continuación.

4.4 Metodología

4.4.1 Pruebas preliminares

Inicialmente se deseaba trabajar con disoluciones de bebida de naranja sin adición de azúcar, para ello se tomó como referencia la concentración de 60 g/L establecida por Ramírez (2016), ya que la concentración recomendada en las instrucciones de uso de la bebida en polvo de naranja marca Tang® resultó ser muy concentrada y ácida (90 g/L), como consecuencia generó problemas de acarreamiento, adaptación y fatiga.

4.4.2 Determinación de las concentraciones adecuadas para realizar las pruebas de discriminación direccionadas y no direccionadas

En las pruebas de discriminación sensorial es necesario que las diferencias entre los estímulos de ambas muestras sean confundibles. Por lo tanto, las concentraciones para comparar que se deseaban definir, deben de estar cercanas a la concentración estándar 60 g/L de Tang, ya sea que estén más o menos concentradas (Angulo y O'Mahony, 2009). Para ello se realizaron varias pruebas preliminares, las cuales se presentan en el siguiente cuadro

Cuadro I. Prueba preliminares realizadas para definir las concentraciones adecuadas que permiten discriminar la muestra más dulce en una bebida de naranja.

Prueba	Objetivo	Concentraciones	Resultado
Umbral de diferencia con estímulo constante (JND) en una prueba 2- AFC con el atributo de muestra más ácida.	Determinar la concentración mínima detectable para producir una diferencia de percepción	Concentración estándar de Tang: 60 g/L Concentraciones de prueba de Tang: 28, 37,8; 51,03; 68,89; 93,02 y 125,55 g/L	JND= 20 g/L de Tang. Los panelistas no lograron distinguir entre el sabor ácido y dulce en la muestra (generó confusión).
Prueba 2- AFC 3 Sets donde se preguntó por la muestra más dulce 3 Sets donde se preguntó por la muestra más intensa	Determinar el atributo que resultara más fácil para que los jueces logaran discriminar entre muestras	Concentración estándar de Tang: 60 g/L Concentraciones de prueba de Tang: 40, 52 y 80 g/L	La concentración más apropiada fue la de 52 g/L de Tang. No hubo grandes diferencias entre los atributos escogidos.

Continuación Cuadro I. Prueba preliminares realizadas para definir las concentraciones adecuadas que permiten discriminar la muestra más dulce en una bebida de naranja.

Prueba	Objetivo	Concentraciones	Resultado
Panel informal 2-AFC con el comité asesor utilizando la de muestra más dulce	Determinar las concentraciones adecuadas para panelistas poco, medio y altamente sensibles	Concentración estándar de Tang: 60 g/L Concentraciones de prueba de Tang: 50, 52, 54 y 56 g/L.	Inconsistencia en resultados Se discriminó en todos los casos menos en la muestra 60 g/L vs. 54 g/L.
Panel informal 2-AFC con adición de azúcar basado en recomendaciones de Ramírez (2016).	Determinar las concentraciones adecuadas para panelistas poco, medio y altamente sensibles para pruebas de discriminación direccionadas	Concentración estándar: 45 g/L de Tang + 13 g/L de azúcar. Concentraciones de prueba: 45 g/L de Tang como base y 17, 19, 21 y 23 g/L de azúcar.	Tanto la concentración estándar, así como las de prueba resultaron eficaces para discriminar la muestra más dulce.
Panel informal Dúo Trío con adición de azúcar basado en recomendaciones de Ramírez (2016).	Determinar las concentraciones adecuadas para panelistas poco, medio y altamente sensibles para pruebas de discriminación no direccionadas	Concentración estándar: 45 g/L de Tang + 13 g/L de azúcar. Concentraciones de prueba: 45 g/L de Tang como base y 21, 23, 25 y 27 g/L de azúcar.	Tanto la concentración estándar, así como las de prueba resultaron eficaces para discriminar entre las muestras.

Como se puede ver en el Cuadro I, en la prueba de umbral, se definieron seis concentraciones por arriba y debajo de la concentración estándar 60 g/L de Tang, las cuales se compararon una a una contra esta. Las muestras se balancearon de acuerdo a los órdenes de presentación existentes (AB y BA) y se presentaron de manera aleatoria para cada juez.

El panel se realizó con un total de 40 personas. Cada panelista realizó la prueba de manera individual en un cubículo con luz blanca para evitar alteraciones en el color de las disoluciones, se le presentó una bandeja con seis pares de muestras, cada par contenía la concentración estándar y una de las concentraciones de prueba.

En las instrucciones se le indicó al panelista lo siguiente: A continuación, se le presentan seis pares de muestras de una bebida de naranja, pruebe cada uno de los sets de arriba hacia abajo y de izquierda a derecha y marque la muestra que presenta el sabor más ácido. Entre cada uno de los sets enjuáguese 3 veces con agua. Los datos se tomaron de manera electrónica con el programa FIZZ®.

Para la segunda prueba preliminar, se definieron tres nuevas concentraciones basadas en el resultado obtenido del JND (ver Cuadro I). Las muestras se balancearon de acuerdo a los órdenes de presentación existentes (AB y BA) y se presentaron de manera aleatoria para cada juez.

El panel se realizó con un total de 70 personas. Cada panelista realizó la prueba de manera individual en un cubículo con luz blanca para evitar alteraciones en el color de las disoluciones. A cada juez se le presentaron dos bandejas, cada una de ellas con tres pares de muestras donde cada par estaba compuesto por la concentración estándar y una de las concentraciones de prueba.

En las instrucciones se le indicó al panelista lo siguiente: A continuación, se le presentan seis pares de muestras de bebida de naranja, inicialmente pruebe los tres primeros sets de arriba abajo y de izquierda a derecha y marque la muestra que presente el sabor más intenso. Seguidamente evalúe los últimos tres sets de la misma manera, pero esta vez escoja la muestra que presente el sabor más dulce. Debe enjuagarse una vez entre cada uno de los sets.

En este panel el asistente estuvo dándole seguimiento a cada panelista para asegurarse que éste evaluara correctamente los sets con sabor más intenso y los sets con sabor más dulce. Los datos se tomaron de manera electrónica con el programa FIZZ®.

Como cada panelista tiene un nivel de sensibilidad distinto, es necesario encontrar varias concentraciones cercanas a la estándar que se puedan utilizar tanto con las personas más sensibles como las menos sensibles. Para ello se realizó la tercera prueba preliminar con el comité asesor mediante un panel informal para evaluar si se lograba discriminar entre la concentración estándar y las

concentraciones definidas con base en los resultados de la segunda prueba preliminar (ver Cuadro I).

Esta prueba presentó un comportamiento inusual, ya que cuando la diferencia entre la concentración de prueba y la concentración estándar fue de 6g/L (par 60g/L y 54 g/L) los jueces no lograron acertar la muestra más dulce pero cuando la diferencia entre las muestras se hizo más pequeña 4 g/L (par 60 g/L y 56 g/L) los jueces si lograron discriminar. Dicho comportamiento puede deberse a la naturaleza de la muestra, la cual resultó ser compleja, ya que enmascara sensaciones entre ácido y dulce.

Por esta razón, no se continuó buscando concentraciones de Tang que se diferencien de la concentración estándar, sino que se decidió acatar las recomendaciones de Ramírez (2016), quien plantea que es necesario la adición de azúcar extra a la concentración estándar para atenuar el sabor ácido y evitar efectos de adaptación, acarreamiento, fatiga entre otros. Dicha autora definió como concentración estándar 60 g/L de bebida en polvo y la adición de 15 g/L, para la prueba 2-AFC se definió una adición de azúcar de 17 g/L y para la prueba triángulo la adición fue de 25 g/L.

Basado en las recomendaciones de Ramírez (2016) se realizaron dos paneles informales. El primero para definir las concentraciones adecuadas para las pruebas de discriminación direccionadas y el segundo para definir las concentraciones de las pruebas de discriminación no direccionadas (ver Cuadro I).

Cabe destacar que en estas pruebas preliminares se utilizó 2-AFC por ser la más sensible y con menor número de estímulos, lo que permite reducir el tiempo y cantidad de panelistas. Por otro lado, para las pruebas no direccionadas se utilizó la prueba Dúo Trío debido a que es la menos sensible, si los panelistas lograban detectar la diferencia con dicha prueba, se aseguraba que también la detectarían con las más sensibles.

Es importante aclarar que el lote de producción de la bebida en polvo se mantuvo constante en todas las pruebas con el fin de evitar posibles alteraciones en los resultados por variabilidad de la muestra.

4.4.3 Determinación de las condiciones e instrucciones adecuadas para realizar las pruebas definitivas

En las pruebas finales se utilizó la concentración que resultó efectiva para cada panelista, es decir, se buscó la diferencia que éste lograba percibir tras probar pares de muestras con diferencia de concentración muy pequeña hasta pares de muestra con diferencias de concentraciones mayores.

Se realizaron pruebas preliminares para definir los protocolos de calentamiento, la secuencia de presentación de las pruebas de discriminación y el número de repeticiones en que se debía presentar esa secuencia para que el panelista sufriera deterioro en su capacidad para detectar la diferencia, es decir que estuviera expuesto a fatiga, adaptación y acarreamiento. Se determinó que se harían cuatro repeticiones de la secuencia (4 rondas).

4.5 Pruebas definitivas

4.5.1 Pruebas direccionadas

Utilizando las soluciones de la bebida Tang® en las concentraciones determinadas previamente en las pruebas preliminares, se realizaron cuatro paneles sensoriales (uno para cada condición que se desea estudiar) de la siguiente manera:

1. Panel sin muestra iniciadora y sin enjuague entre sets.
2. Panel con enjuague con agua entre sets.
3. Panel con muestra iniciadora y con enjuague entre sets.
4. Panel con enjuague con agua entre cada una de las muestras para cada set.

En cada uno de los paneles se presentaron las tres pruebas de discriminación direccionadas en estudio: 2-AFC, 3-AFC y Tétrada direccionada. De esta forma se pretende eliminar las variaciones que podrían presentarse entre pruebas debido a

efectos psicológicos como motivación, confianza, tiempo, estrés entre otros que generan diferencias en rendimiento y sensibilidad, de un día a otro. Al realizar las tres pruebas en una misma sesión, estas serán comparables entre si bajo el mismo grado de rendimiento del juez (O'Mahony, 2018).

Las pruebas se presentaron en una secuencia siguiendo el orden descrito a continuación, el cual equivale a una repetición (ronda): 2-AFC, 3-AFC, Tétrada direccionada, Tétrada direccionada, 3-AFC, 2-AFC. Este orden se presentará por cuadruplicado para cada juez, ya que, según los resultados de las pruebas preliminares, 4 rondas corresponden a la cantidad máxima de repeticiones para los jueces sin perder sensibilidad. Es necesario mencionar que todas las muestras, los enjuagues y "primer" o iniciador, se deben de expectorar para evitar fatiga y saciedad (O'Mahony, 2018).

Haciendo uso del programa FIZZ®[®], se generaron los códigos respectivos para cada set y para cada una de las pruebas direccionadas, tomando en cuenta todos los órdenes de presentación según corresponda a cada prueba y balanceándolos a lo largo del panel para evitar fenómenos de adaptación y aprendizaje entre los panelistas, y también, para minimizar el efecto por orden secuencial sensitivo en caso de que lo hubiese. Asimismo, es importante esclarecer que fue necesario balancear entre los jueces el orden en que se realizaban los 4 paneles sensoriales para evitar las variaciones producto de efectos psicológicos entre las condiciones en estudio.

Se elaboró una hoja de respuesta por cada condición para cada uno de los panelistas, esto se realizó de manera manual acomodando los códigos de cada prueba de discriminación que se le presentó al juez para cada una de las rondas (ver Anexo C).

Las pruebas se realizaron con 27 participantes, cada uno de ellos realizó 4 sesiones (un panel sensorial por cada condición que se desea estudiar). Para ello se agendaron citas con todos los panelistas en un horario acordado según la disponibilidad de los mismos. La duración de las sesiones estuvo sujeta a la

condición evaluada y agilidad del panelista (20- 40) minutos, el tiempo promedio estimado resultó ser de 30 minutos.

Es imprescindible señalar que, en cada una de las sesiones, los participantes estuvieron acompañados y guiados de principio a fin por el evaluador. Ya que, debido a la complejidad de la prueba, fue necesario asegurarse que el panelista entendiera el objetivo de la investigación, discriminara adecuadamente entre las disoluciones escogidas según su sensibilidad y acatará las instrucciones a lo largo de la sesión. Además, el evaluador (estudiante tesaria), se encargaba de anotar los resultados ya que el panelista solamente seleccionaba la muestra o muestras consideradas como las más dulces.

Para la selección de la concentración adecuada para cada panelista, en la primera sesión, se realizó una prueba de calentamiento que consistió en darle a los participantes dos vasos (30 mL de disolución), uno con la concentración estándar (45 g/L de bebida en polvo y 13 g/L de azúcar) y otro con la concentración más cercana a ésta (45 g/L de bebida en polvo y 17 g/L de azúcar). Se le solicitó al panelista probar de izquierda a derecha en forma continua (sin detenerse) varios tragos intercalando ambas disoluciones hasta que lograran seleccionar el vaso con el sabor más dulce (Mata-García *et al.*, 2007). Inmediatamente se iniciaba la secuencia de pruebas (primera ronda), y si la persona no acertaba las primeras 3 pruebas, se repetía el procedimiento de calentamiento con una disolución con mayor contenido de azúcar hasta que éste lograra acertar.

Ahora bien, para asegurarse de que en el resto de sesiones los panelistas tuvieran el mismo nivel de sensibilidad, se realizó el mismo calentamiento utilizando las concentraciones definidas en la primera sesión. Dicho procedimiento se realizó con la intención de mitigar los cambios asociados a diferencias de sensibilidad que podrían presentarse de un día a otro, esto permite que las sesiones (condiciones) sean comparadas de una mejor manera (O'Mahony, 2018).

4.5.2 Pruebas no direccionadas

Utilizando las soluciones de la bebida Tang® en las concentraciones determinadas previamente en las pruebas preliminares, se realizaron cuatro paneles sensoriales (uno para cada condición que se desea estudiar) de la siguiente manera:

1. Panel sin muestra iniciadora y sin enjuague entre sets.
2. Panel solamente con enjuague con agua entre sets.
3. Panel con muestra iniciadora y con enjuague entre sets.
4. Panel con enjuague con agua entre cada una de las muestras para cada set.

En cada uno de los paneles se presentaron las tres pruebas de discriminación no direccionadas en estudio: Dúo Trío, Triángulo y Tétrada no direccionada. De esta forma se pretende eliminar las variaciones que podrían presentarse entre pruebas debido a efectos psicológicos como motivación, confianza, tiempo, estrés entre otros que generan diferencias en rendimiento y sensibilidad. Al realizar las tres pruebas en una misma sesión, estas serán comparables entre si bajo el mismo grado de rendimiento del juez (O'Mahony, 2018).

Las pruebas se presentaron en una secuencia siguiendo el orden descrito a continuación, el cual equivale a una ronda: Dúo Trío, Triángulo, Tétrada, Tétrada, Triángulo, Dúo Trío. Este orden se presentará por cuadruplicado para cada juez, ya que, según los resultados de la tercera prueba preliminar, 4 rondas corresponden a la cantidad máxima de réplicas para los jueces sin perder sensibilidad. Es necesario mencionar que todas las muestras, los enjuagues y "primer" o iniciador, se deben de expectorar para evitar fatiga y saciedad (O'Mahony, 2018).

Haciendo uso del programa FIZZ, se generaron los códigos respectivos para cada set y para cada una de las pruebas no direccionadas, tomando en cuenta todos los órdenes de presentación según corresponda a cada prueba y balanceándolos a lo largo del panel para evitar fenómenos de adaptación y aprendizaje entre los panelistas y también para minimizar el efecto por orden secuencial sensitivo en caso

de que lo hubiese. Asimismo, es importante esclarecer que fue necesario balancear entre los jueces el orden en que se realizaban los 4 paneles sensoriales para evitar las variaciones producto de efectos psicológicos entre las condiciones en estudio.

Se elaboró una hoja de respuesta por cada condición para cada uno de los panelistas, esta se realizó de manera manual acomodando los códigos de cada prueba de discriminación que se le presentó al juez para cada una de las rondas (ver Anexo C).

Las pruebas se realizaron con 27 participantes, cada uno de ellos realizó 4 sesiones (un panel sensorial por cada condición que se desea estudiar). Para ello se agendaron citas con todos los panelistas en un horario acordado según la disponibilidad de los mismos. La duración de las sesiones estuvo sujeta a la condición evaluada y agilidad del panelista (20- 40) minutos, el tiempo promedio estimado resultó ser de 30 minutos.

Es imprescindible señalar que, en cada una de las sesiones, los participantes estuvieron acompañados y guiados de principio a fin por el evaluador. Ya que, debido a la complejidad de la prueba, fue necesario asegurarse que el panelista entendiera el objetivo de la investigación, discriminara adecuadamente entre las disoluciones escogidas según su sensibilidad y acatará las instrucciones a lo largo de la sesión. Además, el evaluador (estudiante tesiaría), se encargaba de anotar los resultados ya que el panelista solamente seleccionaba la muestra o muestras denominadas como diferentes o iguales.

Para la selección de la concentración adecuada para cada panelista, en la primera sesión se les brindó tres sets utilizando la prueba Dúo Trío bajo la primera condición (sin "primer" y sin enjuague) con la disolución estándar (45 g/L de bebida en polvo y 13 g/L de azúcar) y la disolución más cercana a ésta (45 g/L de bebida en polvo y 21 g/L de azúcar). Se le solicitó al panelista probar de izquierda a derecha todas las muestras y seleccionar aquella que fuera igual a la referencia, como mínimo el panelista debía acertar dos sets. En caso contrario se repetía el

procedimiento y se cambiaba la disolución por una con mayor contenido de azúcar hasta que acertara.

Por otro lado, es importante resaltar que en estas pruebas no fue posible realizar el calentamiento, ya que, al ser no direccionadas, el atributo en el cual difieren las muestras no se debe mencionar. Por lo tanto, en este caso es necesario contemplar que podría presentarse variaciones entre las sesiones (condiciones) producto de las diferencias de sensibilidad asociados a efectos psicológicos que varían de un día a otro

4.6 Análisis de datos

4.6.1 Análisis al pasar las rondas

Los datos obtenidos de las hojas de respuesta de los panelistas fueron acomodados de acuerdo con el número de aciertos máximo que éstos lograron en cada ronda, para cada prueba y para cada condición tal y como se muestra en la siguiente figura

CONDICION	N. RONDA	TIPO DE PRUEBA		
		2AFC	3AFC	TETRADA DIRECCIONADA
Sin enjuague ni primer	R1	1	0	0
	R2	2	1	1
	R3	2	1	0
	R4	1	1	1
Solo con enjuague	R1	0	1	0
	R2	2	1	1
	R3	2	0	1
	R4	2	1	0
Con enjuague y primer	R1	1	2	1
	R2	2	1	1
	R3	2	0	1
	R4	2	2	0
Con enjuague entre cada muestra	R1	1	0	2
	R2	1	1	1
	R3	2	2	0
	R4	2	1	0

Figura 1. Ejemplo de acomodo de datos según el número de aciertos que obtuvo un panelista para cada prueba, condición y ronda.

4.6.1.1 Análisis cuantitativo mediante el índice Thurstoniano del grado de diferencia (d')

Para dicho análisis, se ordenaron los datos de número de aciertos de cada panelista (ver Figura 1) para cada ronda y condición realizada. Se determinó el número de aciertos máximo que podían lograrse para cada prueba y condición con la totalidad de los panelistas (el máximo valor que un panelista puede obtener es 2 para cada ronda, por prueba y condición). Se hizo una sumatoria del número de aciertos de todos los participantes (el máximo valor que la sumatoria de los panelistas podría obtener es 54 por ronda, prueba y condición), se calculó la proporción de esos aciertos y con las tablas de valor d' se estableció un valor para cada ronda y condición. Esto se realizó para cada una de las pruebas de discriminación tanto direccionadas como no direccionadas tal y como se muestra en la siguiente figura

JUEZ	RONDA	CONDICION			
		A	B	C	D
1	R1	1	1	2	2
2	R1	1	2	1	1
3	R1	2	2	2	1
4	R1	1	2	1	2
5	R1	1	1	2	1
6	R1	1	2	0	0
7	R1	2	1	0	1
8	R1	2	2	2	2
9	R1	0	0	2	1
10	R1	2	2	1	2
11	R1	2	1	1	2

RONDAS	A= SIN "PRIMER" Y ENJUAGUE AL INICIO		
	ACIERTOS	PROPORCION	D PRIMA
R1	37	0,685	0,68
R2	37	0,685	0,68
R3	41	0,759	0,99
R4	39	0,722	0,83

Numero total de jueces es 27 Numero maximo de aciertos para los 27 jueces por cada ronda seria 54.

Figura 2. Ejemplo de acomodo de datos para realizar el análisis cuantitativo por medio de los valores d' para cada una de las rondas por condición y prueba en estudio.

4.6.1.2 Análisis cualitativo

Este análisis se realizó en varias etapas, las cuales se describen a continuación.

Inicialmente se analizó el comportamiento a través de las rondas para cada uno de los panelistas con el fin de identificar si los mismos compartían alguna tendencia en particular y formar así grupos o “clusters”. Se observaron varios comportamientos y ante esto se generó la siguiente codificación:

CODIFICACION " CONFORME PASAN LAS RONDAS "

0= mal (0/2)

1= empeora (1 a 0)

2= empeora levemente (2 a 1)

3= no mejora ni empeora (inconsistencia entre rondas, se identifica más azar)

4= mejora, pero en **4ta ronda** se satura (cansancio)

5= mejora levemente (0 a 1)

6= mejora al pasar las rondas (1 a 2)

7= muy bien (2/2) en todo o solo al final (por cansancio) pierde alguna

8= se mantiene constante (no mejora ni empeora, pero no se ve inconsistencia entre rondas, podría fallar una al final)

Figura 3. Codificación generada en el análisis cualitativo de las pruebas finales.

Luego, dicha codificación se asignó a cada panelista conforme pasaban las rondas, en cada condición para cada una de las pruebas de discriminación en estudio. Seguidamente, se realizó un conteo de la cantidad de personas que obtuvieron cada una de las categorías de la codificación establecida para cada condición y tipo de prueba con el objetivo de analizar a todos los panelistas en conjunto.

Las nueve categorías de la codificación se agruparon para formar tres nuevas categorías (comportamientos) de la siguiente manera: categorías en las que los panelistas mejoran o se mantienen a través de las rondas (códigos 5,6,7 y 8), categorías en las que empeoran o se deterioran a través de las rondas (códigos 1,2 y 4) y categorías donde no siguen ninguna tendencia a través de las rondas (códigos 0 y 3). Con estos datos se elaboraron gráficos para cada prueba en donde se logró evidenciar cual condición resultó más favorable para la mayoría de los panelistas.

4.6.2 Análisis por prueba y por condición

Para este análisis, se utilizaron los datos de número de aciertos de cada panelista (ver Figura 1) y se realizó una sumatoria del número de aciertos obtenidos en cada ronda para cada uno de los panelistas de la siguiente manera (ver Figura 4):

- I. Número de aciertos totales para cada prueba y condición (máximo 8 por panelista).
- II. Los datos anteriores que corresponden al número de aciertos totales para cada prueba y condición se ordenaron de menor a mayor para cada uno de los panelistas.
- III. Al ordenamiento anterior se le asignó una numeración del 1 al 12, donde 1 correspondía al número de aciertos más bajo y 12 el número de aciertos más alto (ranking) para cada panelista.

CONDICION	N. RONDA	TIPO DE PRUEBA				
		2AFC	3AFC	TETRADA DIRECCIONADA		
Sin enjuague ni primer	R1	2	2	1		
	R2	2	2	2		
	R3	2	2	2		
	R4	2	2	1		
Solo con enjuague	R1	2	1	1		
	R2	1	0	1		
	R3	1	0	1		
	R4	2	2	0		
Con enjuague y primer	R1	2	2	1		
	R2	1	2	2		
	R3	2	2	1		
	R4	2	1	2		
Enjuague entre cada muestra	R1	1	0	0		
	R2	1	1	0		
	R3	2	1	1		
	R4	2	0	1		
Sin enjuague ni primer	2-AFC	8	Con enjuague entre cada muestra	3-AFC	2	1
Solo con enjuague	2-AFC	6	Con enjuague entre cada muestra	Tetrada	2	2
Con enjuague y primer	2-AFC	7	Solo con enjuague	3-AFC	3	3
Con enjuague entre cada muestra	2-AFC	6	Solo con enjuague	Tetrada	3	4
Sin enjuague ni primer	3-AFC	8	Solo con enjuague	2-AFC	6	5
Solo con enjuague	3-AFC	3	Con enjuague entre cada muestra	2-AFC	6	6
Con enjuague y primer	3-AFC	7	Sin enjuague ni primer	Tetrada	6	7
Con enjuague entre cada muestra	3-AFC	2	Con enjuague y primer	Tetrada	6	8
Sin enjuague ni primer	Tetrada	6	Con enjuague y primer	2-AFC	7	9
Solo con enjuague	Tetrada	3	Con enjuague y primer	3-AFC	7	10
Con enjuague y primer	Tetrada	6	Sin enjuague ni primer	2-AFC	8	11
Con enjuague entre cada muestra	Tetrada	2	Sin enjuague ni primer	3-AFC	8	12

Figura 4. Ejemplo de acomodo de datos para realizar el análisis cuantitativo de ordenamiento de rangos Newell McFarland y ANDEVA.

4.6.2.1 Análisis por ordenamiento de rangos Newell McFarland

Se utilizaron los datos ordenados del 1 al 12 de cada panelista (ranking) y se sacó el promedio del ranking obtenido para todos los panelistas bajo cada condición y tipo de prueba. Se utilizaron las tablas de Newell McFarland para establecer el número mínimo de diferencia entre cada prueba y condición que debe haber para obtener diferencias significativas entre cada uno de los tratamientos.

4.6.2.2 Análisis de Varianza

Para este análisis se utilizaron los números de aciertos totales de cada panelista para cada prueba y condición, y mediante el uso del programa XLSTAT se realizó el ANDEVA por tipo de prueba y condición.

5. RESULTADOS Y DISCUSIÓN

5.1 Pruebas preliminares

5.1.1 Prueba de umbral de diferencia con estímulo constante (JND)

En la Figura 3 se presenta la curva psicofísica que muestra cómo se calculó la diferencia apenas perceptible del sabor ácido, en una bebida de naranja con concentración estándar 60 g/L que se obtuvo con 40 participantes

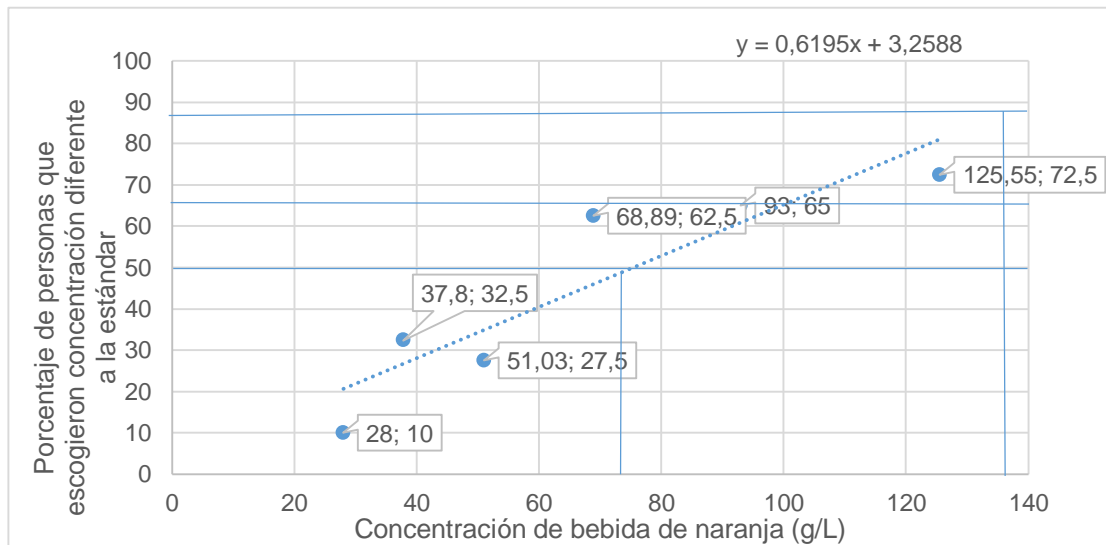


Figura 5. Porcentaje de personas que escogieron una concentración más ácida que la concentración estándar (60 g/L) en una prueba 2-AFC.

Utilizando la ecuación de la recta que presenta el gráfico de la Figura 5, se encontró que el umbral de diferencia apenas perceptible del sabor ácido para una bebida de naranja con concentración estándar de 60 g/L de Tang® corresponde a 20 g/L. Es decir, con un incremento o reducción de 20 g/L respecto a la concentración estándar, los panelistas son capaces de percibir una diferencia en el sabor ácido.

Cabe señalar que dicho valor representa una diferencia bastante grande en comparación con la concentración estándar y, en pruebas subsecuentes unos pocos panelistas mostraron que en pruebas de discriminación las diferencias entre los dos estímulos si se percibieron, por lo que se requirió usar diferencias menores a 20 g/L. Es decir, el JND no se pudo calcular correctamente.

5.1.2 Prueba 2-AFC con diferentes concentraciones versus la disolución patrón

Debido a lo anterior, se decidió realizar una segunda prueba preliminar en la cual se definieron tres concentraciones con valores cercanos al JND resultante, una concentración alta, una media y una baja (80, 52 y 40) g/L con respecto a la concentración estándar 60g/L, esto con el objetivo de encontrar aquella concentración que resulte confundible con respecto a la estándar.

En el Cuadro II se resumen los resultados obtenidos al aplicar seis pruebas consecutivas de 2-AFC a 70 panelistas, en las que se buscaba discriminar la muestra más intensa en los primeros tres sets y la muestra más dulce en los sets restantes respecto a la bebida estándar de 60 g/L.

Cuadro II. Determinación de la concentración más adecuada para percibir la diferencia en dulzor o intensidad respecto a una bebida de naranja estándar de 60 g/L de Tang® utilizando una prueba 2-AFC.

Atributo a encontrar	Concentración de la bebida (g/L)	Aciertos obtenidos	Proporción de aciertos	d' Ennis ¹	Mínimo de aciertos ²	Resultado
<i>Intensidad</i>	40	64	0,914	1,93	43	Hay diferencia significativa
	52	53	0,757	0,98	43	Hay diferencia significativa
	80	64	0,914	1,93	43	Hay diferencia significativa
<i>Dulzor</i>	40	66	0,943	2,23	43	Hay diferencia significativa
	52	57	0,814	1,26	43	Hay diferencia significativa
	80	64	0,914	1,93	43	Hay diferencia significativa

1.Ennis 1993 ($\alpha=0,05$), 2. Ennis 2013 ($\alpha=0,05$)

Como se puede observar en el cuadro anterior, en todos los casos se encontró diferencia significativa respecto a la muestra patrón, ya que el número de aciertos para cada prueba fue mayor que el número de aciertos mínimo requerido según las

tablas binomiales de Ennis (1993). También se puede notar que casi no hubo diferencia entre preguntar por la muestra más intensa o la más dulce.

Ahora bien, al observar los valores obtenidos del parámetro d' , se puede evidenciar que la concentración de la bebida de naranja que resultó ser más adecuada para discriminar la intensidad de sabor entre las muestras corresponde a 52 g/L, ya que ésta presentó un $d' = 0,98$, dicho valor es muy cercano a un $d' = 1$ el cual se recomienda aplicar en pruebas de discriminación porque se ha evidenciado que alrededor de ese valor la diferencia entre los estímulos es confundible (Worch y Delcher, 2013). Cabe destacar que fue posible discriminar entre muestras con una diferencia de 8 g/L respecto a la concentración estándar (60 g/L).

5.1.3 Prueba mediante paneles informales para definir las concentraciones para los panelistas en pruebas direccionadas y no direccionadas.

Tomando en cuenta los resultados obtenidos en la prueba preliminar anterior, se realizó un panel informal de 5 personas para discriminar la muestra con sabor más dulce, para ello se utilizó la concentración estándar 60 g/L y las siguientes concentraciones (50, 52, 54 y 56) g/L. En los resultados obtenidos, los panelistas fueron capaces de discriminar todas las concentraciones en estudio excepto la de 54 g/L, la cual era intermedia.

En esta prueba se pudo observar que la bebida de naranja a base de una solución en polvo marca Tang®, resultó ser una muestra muy compleja, ya que su comportamiento al disminuir o aumentar su concentración no fue lineal, es decir en algunas ocasiones al disminuir su concentración la diferencia fue perceptible y al aumentarla no se percibía, pero después de cierto punto se hacía perceptible de nuevo. Además, se logró ver que al realizar cambios en la concentración de la bebida también se modificaban varios de sus atributos sensoriales, ya que los jueces tendían a confundir el sabor ácido con el dulce.

Ante esto se decidió seguir las recomendaciones de Ramírez (2016) y tras varios paneles informales se definió la concentración estándar como 45 g/L de bebida en polvo y una adición de 13 g/L de azúcar. Para las pruebas direccionadas las

concentraciones para los panelistas fueron las siguientes: 45 g/L de bebida en polvo y (17, 19, 21 y 23) g/L de azúcar respectivamente. En el caso de las pruebas no direccionadas las concentraciones para los panelistas fueron: 45 g/L de bebida en polvo y (21, 23, 25 y 27) g/L de azúcar respectivamente.

5.2 Pruebas definitivas

5.2.1 Análisis de la capacidad de detectar la diferencia al pasar las rondas mediante el índice Thurstoniano del grado de diferencia (d') y análisis cualitativo por codificación

En el Cuadro III se muestran los valores del parámetro d' obtenidos al pasar las rondas para cada una de las pruebas direccionadas bajo las cuatro condiciones en estudio

Cuadro III. Valores del parámetro d' obtenidos al pasar las rondas para las pruebas 2-AFC, 3-AFC y Tétrada direccionada utilizando cuatro condiciones distintas n=54.

Prueba	Rondas	Condición											
		Sin enjuague ni "primer"			Enjuague al inicio			Enjuague y "primer" al inicio			Enjuague entre c/ muestra		
		Aciertos	Proporción	Valor d'	Aciertos	Proporción	Valor d'	Aciertos	Proporción	Valor d'	Aciertos	Proporción	Valor d'
2-AFC	R1	37	0,685	0,68	33	0,611	0,40	36	0,667	0,61	31	0,574	0,26
	R2	37	0,685	0,68	38	0,704	0,76	32	0,593	0,33	32	0,593	0,33
	R3	41	0,759	0,99	40	0,741	0,91	36	0,667	0,61	39	0,722	0,83
	R4	39	0,722	0,83	36	0,667	0,61	32	0,593	0,33	41	0,759	0,99
3-AFC	R1	38	0,704	1,26	36	0,667	1,12	37	0,685	1,18	29	0,537	0,68
	R2	35	0,648	1,05	29	0,537	0,68	25	0,463	0,44	32	0,593	0,86
	R3	39	0,722	1,32	28	0,519	0,62	31	0,574	0,80	25	0,463	0,44
	R4	40	0,741	1,40	31	0,574	0,81	35	0,648	1,05	30	0,556	0,74
Tétrada direccionada	R1	24	0,444	0,95	25	0,463	1,00	17	0,315	0,55	24	0,444	0,95
	R2	31	0,574	1,33	20	0,370	0,73	26	0,481	1,06	23	0,426	0,89
	R3	30	0,556	1,28	25	0,463	1,00	23	0,426	0,89	25	0,463	1,00
	R4	29	0,537	1,22	23	0,426	0,90	28	0,519	1,16	22	0,407	0,84

En la Figura 6 se presentan tres gráficos con los resultados obtenidos del análisis cualitativo al pasar las rondas para cada una de las pruebas direccionadas bajo las cuatro condiciones en estudio.

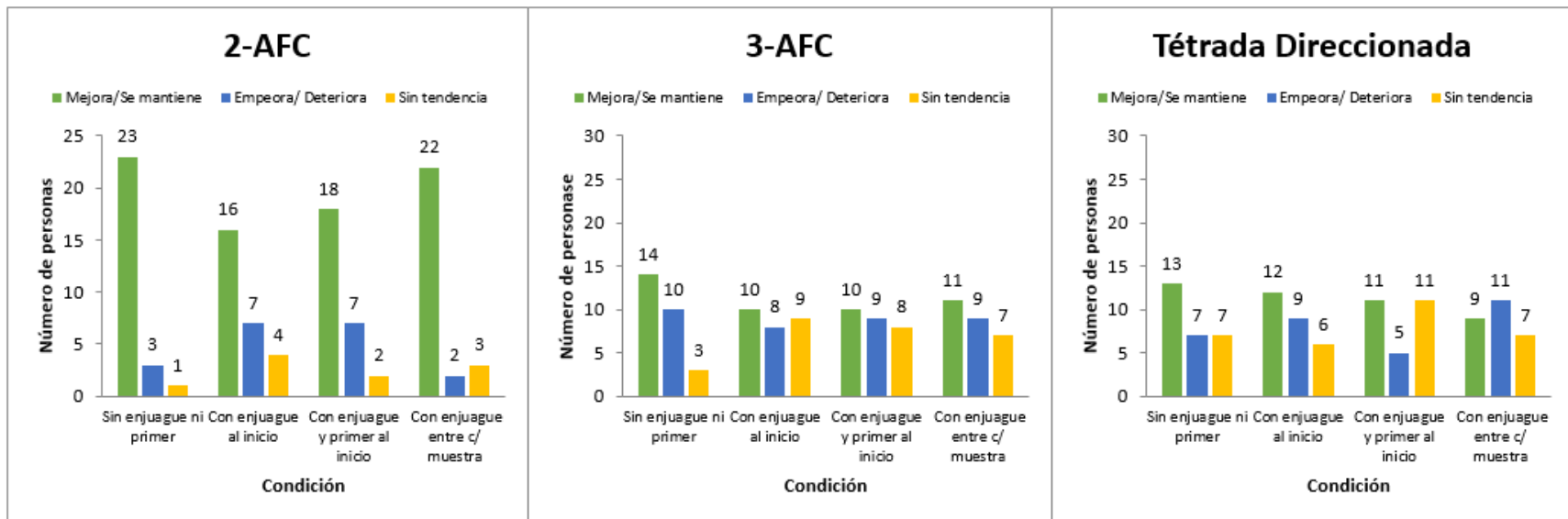


Figura 6. Número de personas que mejoran, empeoran o no presenta ninguna tendencia al realizar las pruebas 2-AFC, 3-AFC y Tétrada direccionada bajo cuatro condiciones distintas.

En el Cuadro III se puede ver que para todas las pruebas bajo cada una de las condiciones en estudio existe una tendencia a incrementar el valor d' al pasar las rondas, es decir, al aumentar el número de pruebas en la sesión también aumenta la capacidad de los panelistas para percibir la diferencia entre las muestras.

Para la prueba 2-AFC las condiciones sin enjuague ni “primer” y con enjuague al inicio tienden a aumentar su valor d' hasta la tercera ronda. Sin embargo, en la cuarta ronda disminuye, esto puede ser un indicador de que en ese punto los panelistas empiezan a perder la sensibilidad por efectos de cansancio, saturación y/o distracción.

En la condición que presenta un enjuague entre cada una de las muestras, se puede ver que en las primeras dos rondas los panelistas no logran discriminar, ya que el valor d' es muy pequeño, mientras que en la tercera y cuarta ronda el valor d' incrementa hasta casi llegar a una unidad. Con esto se puede evidenciar que el calentamiento realizado no fue suficiente y que en dicha condición existe una mayor dificultad tal y como era de esperarse, esto porque al agregar dos estímulos extra a la prueba (enjuagues), el panelista debe realizar un mayor esfuerzo tanto para recordar como para adquirir habilidad a la hora de hacer la prueba. Por otro lado, parece que al transcurrir la sesión se da un efecto de aprendizaje, de manera que logran encontrar la forma de recordar los estímulos en las rondas finales.

En la condición con enjuague y “primer”, se puede observar un comportamiento inconsistente, ya que al pasar las rondas no se presenta una tendencia a disminuir o aumentar el valor d' y además los valores reportados son bajos, por lo tanto, se puede concluir que esta condición no es adecuada para la prueba 2-AFC.

Al observar la figura 6 (que corresponde al histograma de mejoró o no en cada prueba al pasar las rondas), se confirma lo planteado con anterioridad para la prueba 2-AFC, ya que un mayor número de personas tiende a mejorar o mantener

su sensibilidad al pasar las rondas bajo la primera y última condición (sin enjuague ni “primer” y con enjuague entre cada una de las muestras).

Por otro lado, en el Cuadro III se puede ver que para la prueba 3-AFC la condición que resulta más adecuada para discriminar conforme pasan las rondas es la que no contiene ni enjuague ni primer. En esta condición los valores d' tienden a aumentar al pasar las rondas mientras que en las otras condiciones se presentan inconsistencias o el valor d' tiende a hacerse más pequeño, lo cual indica que dichas condiciones no ayudan a mejorar la respuesta de los panelistas ya que se les dificulta discriminar entre las muestras. Este comportamiento se puede evidenciar en el gráfico presente en la figura 6, donde se muestra que la mayoría de panelistas mejoran o se mantienen a través de las rondas bajo la primera condición.

Por otra parte, si se compara el valor d' de la primera ronda para cada condición, se observa que son similares e indican que los jueces discriminaron las muestras en todos los casos excepto para la condición con enjuague entre cada una de las muestras. Lo que indica que desde el inicio dicha condición afecta la memoria o bien la capacidad de reconocer tres estímulos con enjuagues en medio.

De igual manera, para la prueba Tétrada direccionada, la condición más favorable resultó ser la que no usa enjuague ni “primer” seguida de la condición enjuague y “primer”, esto se puede comprobar al observar los valores d' presentes en el Cuadro III, los cuales aumentan al pasar las rondas, mientras que las condiciones restantes presentan comportamientos variados, tanto la condición con enjuague al inicio como la de enjuague entre cada una de las muestras mantienen la sensibilidad, pero con valores d' más bajos.

Al comparar el d' de la primera ronda para cada condición se observa que el valor mayor se obtuvo con la condición con enjuague al inicio de cada set y el valor más bajo con enjuague y “primer”. Se esperaba que la condición con enjuague entre cada una de las muestras afectara la memoria, especialmente en la prueba tétrada debido a que presenta un mayor número de estímulos (Ishii *et al.*, 2014; Xia *et al.*,

2015). Sin embargo, no fue la condición que presentó el mayor deterioro de sensibilidad.

Al observar la Figura 6, se puede evidenciar que la mayoría de personas tienden a mejorar utilizando la primera condición mientras que en las condiciones restantes no se aprecia una gran diferencia entre usar una u otra.

Ahora bien, al comparar entre las pruebas direccionadas, se puede notar que la prueba 2-AFC, fue la que permitió a los panelistas mantener o mejorar su respuesta independientemente de la condición. Esto coincide con lo que dicta la literatura, ya que dicha prueba es más potente que la 3-AFC y Tétrada direccionada. Además, presenta un menor número de estímulos; por lo tanto, las posibles implicaciones de memoria siempre serán menores que para las pruebas que presentan 3 o 4 estímulos. Cabe destacar que la prueba 3-AFC a pesar de tener menor cantidad de estímulos, no fue mejor que la prueba Tétrada direccionada en cuanto a deterioro al pasar las rondas (O'Mahony y Rousseau, 2002).

Por otra parte, la condición que permitió mantener o mejorar el desempeño de los panelistas al realizar las distintas pruebas direccionadas fue la primera (sin enjuague ni "primer"), confirmando así lo expuesto por varios autores en la literatura quienes señalan que el mejor medio para limpiar el paladar es la saliva (Lucak, 2008). Sin embargo, es necesario mencionar que las demás condiciones podrían resultar favorables dependiendo del tipo de prueba y número de repeticiones que se desee realizar, ya que como se mencionó en párrafos anteriores las implicaciones de memoria parecen afectar directamente estas dos variables en estudio (Lucak, 2008; Ishii *et al.*, 2014; Xia *et al.*, 2015).

En el Cuadro IV se muestra los valores del parámetro d' obtenidos al pasar las rondas para cada una de las pruebas no direccionadas bajo las cuatro condiciones en estudio.

Cuadro IV. Valores del parámetro d' obtenidos al pasar las rondas para las pruebas Dúo Trío, Triángulo y Tétrada utilizando cuatro condiciones distintas.

Prueba	Rondas	Condición											
		Sin enjuague ni "primer"			Enjuague al inicio			Enjuague y "primer" al inicio			Enjuague entre c/ muestra		
		Aciertos	Proporción	Valor d'	Aciertos	Proporción	Valor d'	Aciertos	Proporción	Valor d'	Aciertos	Proporción	Valor d'
<i>Dúo Trío</i>	R1	42	0,778	2,20	43	0,796	2,33	41	0,759	2,08	33	0,611	1,19
	R2	44	0,815	2,47	41	0,759	2,08	42	0,778	2,20	40	0,741	1,96
	R3	42	0,778	2,20	38	0,704	1,74	44	0,815	2,47	39	0,722	1,85
	R4	38	0,704	1,74	34	0,630	1,30	44	0,815	2,47	34	0,630	1,30
<i>Triángulo</i>	R1	36	0,667	2,32	37	0,685	2,42	36	0,667	2,32	27	0,500	1,50
	R2	42	0,778	2,98	43	0,796	3,10	32	0,593	1,94	25	0,463	1,27
	R3	40	0,741	2,74	37	0,685	2,42	39	0,722	2,63	33	0,611	2,03
	R4	42	0,778	2,98	36	0,667	2,32	32	0,593	1,94	25	0,463	1,27
<i>Tétrada</i>	R1	38	0,704	1,72	37	0,685	1,65	32	0,593	1,34	36	0,667	1,59
	R2	34	0,630	1,46	46	0,852	2,35	35	0,648	1,53	34	0,630	1,46
	R3	42	0,778	2,00	43	0,796	2,08	40	0,741	1,86	38	0,704	1,72
	R4	41	0,759	1,93	30	0,556	1,22	37	0,685	1,65	29	0,537	1,15

En la Figura 7 se presentan tres gráficos con los resultados obtenidos del análisis cualitativo al pasar las rondas para cada una de las pruebas no direccionadas bajo las cuatro condiciones en estudio.

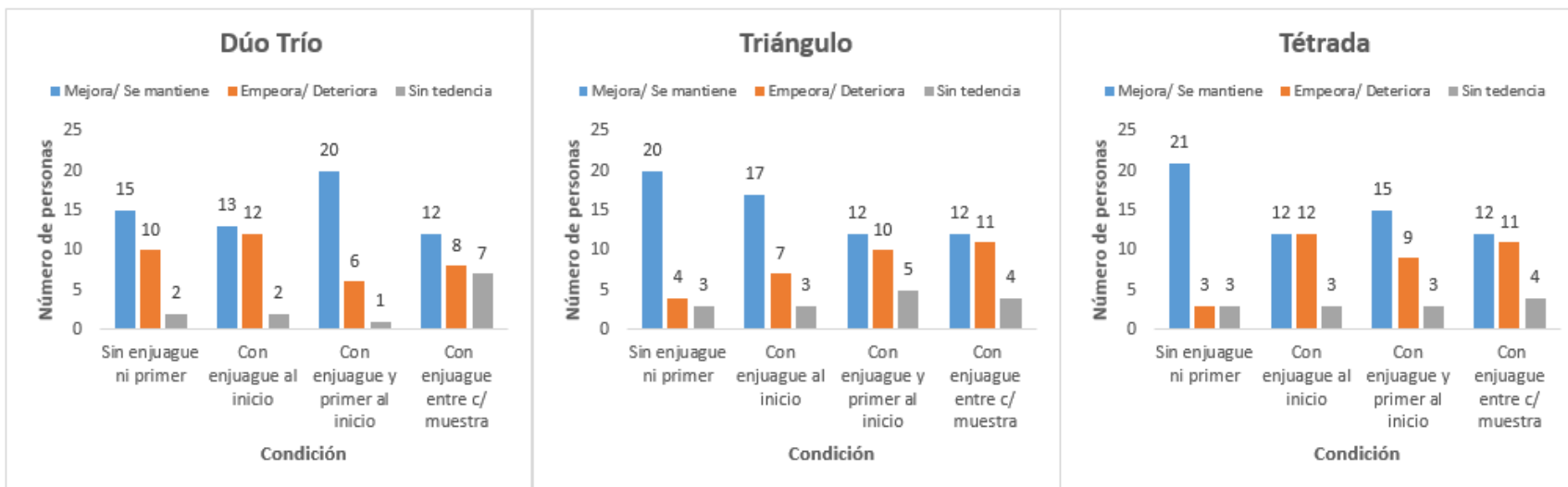


Figura 7. Número de personas que mejoran, empeoran o no presenta ninguna tendencia al realizar las pruebas Dúo Trío, Triángulo y Tétrada bajo cuatro condiciones distintas.

Al observar el Cuadro IV, se puede apreciar que para las pruebas no direccionadas bajo cada una de las condiciones en estudio existe una tendencia de incrementar el valor d' al pasar las rondas. Sin embargo, dicho comportamiento se da hasta la tercera ronda, ya que en la mayoría de los casos en la última ronda el valor d' tiende a caer. Ante esto, se puede inferir que la sensibilidad de los panelistas se ve mayormente afectada por el número de réplicas en comparación con las pruebas direccionadas. Es decir, en estas pruebas los panelistas presentaron más problemas de cansancio, concentración y/o distracción. A continuación, se analiza en detalle.

De acuerdo con el Cuadro IV, para la prueba Dúo Trío, se puede ver que la condición más favorable resultó ser la que contiene enjuague y primer, ya que con el pasar de las rondas los valores d' aumentan o se mantienen; es decir, esta condición les permite a los panelistas mejorar o mantener su sensibilidad a lo largo de la prueba, esto se puede evidenciar al observar el gráfico presente en la figura 5.

Con las condiciones sin enjuague ni “primer” y con enjuague entre cada una de las muestras, se puede ver un incremento aparente en los valores d' en las rondas uno y dos, pero en las rondas restantes los valores empiezan a disminuir, especialmente en la última condición. De manera similar en la condición con enjuague, al inicio la tendencia es disminuir los valores d' al pasar las rondas. Por lo tanto, se puede inferir que dichas condiciones no le permiten al panelista mejorar su desempeño en la prueba cuando se presentan 3 o 4 rondas, ya que al aumentar el número de réplicas se dificulta la capacidad para discriminar. Se recomendaría realizar solo dos rondas para no deteriorar la sensibilidad de los panelistas.

Al observar el Cuadro IV, se puede apreciar que para la prueba Triángulo la condición que resulta más favorable es la que no presenta ni enjuague ni “primer”, ya que los valores d' aumentan al pasar las rondas. La condición con solo enjuague presenta un aumento aparente hasta las primeras dos rondas y después los valores d' disminuyen. En las condiciones restantes se observa cierta inconsistencia al pasar las rondas; por lo tanto, se infiere que no ayudan a los panelistas a discriminar

entre las muestras durante la prueba en las últimas rondas. Este comportamiento se puede visualizar en el gráfico presente en la Figura 7, donde la mayoría de los jueces tienden a mejorar o mantenerse con la primera condición.

En la prueba Tétrada, según los valores d' presentes en el Cuadro IV, se puede ver que para las condiciones sin enjuague ni “primer” y con enjuague y “primer” al inicio, la tendencia es aumentar o mantener la capacidad de discriminar, pero en la cuarta ronda, este proceso se dificulta. Para la primera condición este comportamiento era de esperarse, ya que al no introducir un estímulo extra en la prueba se evita la pérdida de sensibilidad por efectos de memoria, aunque en la cuarta ronda se falla por fatiga y adaptación (Ishii *et al.*, 2014).

Por otra parte, en el caso de la condición con enjuague y “primer”, el comportamiento coincide con lo encontrado en artículos científicos recientes, los cuales plantean que el uso del primer ayuda a prevenir el efecto distorsionante que puede causar el enjuague de agua sobre la degustación de la primera muestra (Rousseau *et al.*, 2002; Xia *et al.*, 2014).

En las condiciones restantes se puede apreciar valores d' variados, lo cual evidencia un deterioro en la capacidad de discriminar, en la condición con enjuague al inicio de cada set por la dilución de la primera muestra y por efectos de memoria con la condición con enjuague entre cada una de las muestras (Lucak, 2008; Ishii *et al.*, 2014; Xia *et al.*, 2015).

Al comparar los resultados obtenidos entre las pruebas, no se puede ver una clara diferencia entre cuál de éstas presentó un mejor desempeño sin importar la condición utilizada. Sin embargo, se esperaba que la prueba Tétrada fuera la que presentara los mejores resultados, ya que se ha comprobado en diferentes estudios que resulta ser más sensible que la prueba Triángulo y Dúo Trío, pero al tener un mayor número de estímulos es posible que se viera afectada por problemas de memoria especialmente con las condiciones que presentan un mayor número de estímulos en la prueba como lo es la condición con enjuague entre cada una de las muestras (Ishii *et al.*, 2014; Xia *et al.*, 2015).

Por otro lado, la condición que permitió mantener o mejorar la sensibilidad de los panelistas al realizar las diferentes pruebas no direccionadas fue la que no presentó enjuague ni “primer”. Sin embargo, es necesario esclarecer que tanto para la prueba Dúo Trío como para la Tétrada no direccionada, la condición con enjuague y “primer” resultó también bastante favorable. Caso contrario ocurrió con la condición que presenta enjuagues entre cada una de las muestras, ya los valores d' en todos los casos fueron bastante menores a las otras condiciones, por lo que se puede inferir que esta condición dificultó a los panelistas su capacidad de discriminar entre muestras.

De acuerdo con lo anterior, se puede ver que no existe una condición universal que ayude a mantener o mejorar el desempeño de los jueces tanto en las pruebas direccionadas como las no direccionadas. Sino, que esto va depender de la sensibilidad de la prueba y del número de estímulos que ésta presente.

Además, es importante señalar que existen pruebas más sensibles y potentes estadísticamente que otras, por lo cual se requiere un número mayor de panelistas en las pruebas menos potentes para obtener el mismo grado de sensibilidad que las pruebas más potentes. Sin embargo, en esta investigación todas las pruebas se realizaron con el mismo número de panelistas con el fin de hacer comparaciones bajo igualdad de condiciones, dado esto, los resultados obtenidos pueden verse ligeramente afectados por este inconveniente.

Aun así, las pruebas 2-AFC son las más sensibles, seguidas por las 3-AFC y Tétrada (Ennis, Rousseau y Ennis., 2014). No obstante, en el presente estudio el d' muestra que el tamaño de la diferencia captado por 27 panelistas no fue menor para los métodos menos sensibles en comparación con el 2-AFC. Algo similar ocurrió con las pruebas no direccionadas, donde la diferencia captada por 27 panelistas no resultó menor para las pruebas menos sensibles.

Es importante tomar en cuenta que la diferencia entre las concentraciones que se compararon (estándar versus otra con mayor contenido de azúcar) fueron mayores para las pruebas no direccionadas en comparación con las direccionadas,

este aumento en concentración compensa el uso de un menor número de panelistas.

5.2.2 Análisis por prueba y por condición por ordenamiento de rangos Newell McFarland y por análisis de Varianza.

En los cuadros V y VI se muestra la comparación entre cada prueba direccionada y condición en estudio en muestras de bebida de naranja utilizando el método de ordenamiento por rangos con las tablas de Newell y McFarland (1987) y el análisis de varianza respectivamente.

Cuadro V. Comparación de los aciertos totales entre las pruebas 2-AFC, 3-AFC y Tétrada direccionada bajo cada condición en estudio mediante el método de ordenamiento por rangos con las tablas de Newell y McFarland (1987) con un $\alpha=0,05$ para 27 panelistas.

Condición	Prueba	Promedio
Con enjuague entre cada muestra	2-AFC	8,370 ^a
Sin enjuague ni "primer"	2-AFC	<u>8,000</u> ^{ab}
Solo con enjuague	2-AFC	7,963 ^{ab}
Con enjuague y "primer"	2-AFC	7,889 ^{ab}
Sin enjuague ni "primer"	3-AFC	<u>7,481</u> ^{abc}
Con enjuague y "primer"	3-AFC	6,037 ^{abc}
Con enjuague entre cada muestra	3-AFC	6,000 ^{abc}
Sin enjuague ni "primer"	Tétrada D	<u>5,926</u> ^{abc}
Solo con enjuague	3-AFC	5,889 ^{abc}
Con enjuague entre cada muestra	Tétrada D	5,037 ^{bc}
Con enjuague y "primer"	Tétrada D	4,815 ^{bc}
Solo con enjuague	Tétrada D	4,593 ^c

De acuerdo con los promedios obtenidos en el Cuadro V, se puede ver que los tratamientos que resultaron más efectivos para mantener o mejorar el desempeño de los jueces para discriminar entre las muestras de bebidas de naranja fueron las prueba 2-AFC y 3-AFC con las 4 condiciones en estudio. Sin embargo, dichos tratamientos no presentan diferencias significativas con respecto a la Tétrada direccionada bajo la condición sin enjuague ni primer. Por otro lado, no se aprecia diferencia significativa entre utilizar la prueba 3-AFC bajo todas las condiciones y

Tétrada direccionada bajo todas las condiciones (dicha prueba presentó el menor número de aciertos).

Se debe tomar en cuenta que los análisis estadísticos no paramétricos son poco sensibles para captar diferencias significativas. Por lo tanto, si se considera cada método de discriminación, la condición que permitió mejor desempeño para las tres pruebas fue sin enjuague ni “primer” (promedios subrayados, Cuadro V).

Cuadro VI. Comparación de los aciertos totales entre las pruebas 2-AFC, 3-AFC y Tétrada direccionada bajo cada condición en estudio mediante el método de análisis de varianza con un $\alpha=0,05$ para 27 panelistas.

Condición	Prueba	Promedio
Sin enjuague ni “primer”	2-AFC	<u>5,963</u> ^a
Solo con enjuague	2-AFC	5,778 ^a
Con enjuague entre cada muestra	2-AFC	5,704 ^a
Sin enjuague ni “primer”	3-AFC	<u>5,667</u> ^a
Con enjuague y “primer”	2-AFC	5,333 ^{ab}
Con enjuague y “primer”	3-AFC	4,778 ^b
Solo con enjuague	3-AFC	4,667 ^{bc}
Sin enjuague ni “primer”	Tétrada	<u>4,444</u> ^{bc}
Con enjuague entre cada muestra	3-AFC	4,296 ^{cd}
Con enjuague entre cada muestra	Tétrada	3,778 ^{cd}
Con enjuague y “primer”	Tétrada	3,778 ^d
Solo con enjuague	Tétrada	3,630 ^d

Al observar el Cuadro VI, se puede ver que el tratamiento que resultó más favorable para que los jueces mejoraran su desempeño o lo mantuvieran fue la prueba 2-AFC bajo las diferentes condiciones, ya que estos tratamientos no presentaron diferencias significativas entre ellos, y se incluye también la prueba 3-AFC bajo la condición sin enjuague ni primer. Tampoco se aprecian diferencias significativas entre la prueba 3-AFC con enjuague entre cada muestra ni la prueba Tétrada direccionada con las condiciones restantes, que representan los tratamientos que lograron un menor número de aciertos. Considerando cada método de discriminación, se puede observar que la mejor condición fue la de sin

enjuague ni “primer” para las tres pruebas de discriminación direccionadas (ver promedios subrayados, Cuadro VI).

Ambos métodos estadísticos coinciden en que la prueba más efectiva fue la 2-AFC especialmente con las condiciones sin enjuague ni “primer” y con enjuague entre cada una de las muestras. Además, dejan en evidencia que para la prueba Tétrada direccionada la condición más eficiente fue sin enjuague ni “primer”.

En los cuadros VII y VIII se muestra la comparación entre cada prueba no direccionada y condición en estudio en muestras de bebida de naranja utilizando el método de ordenamiento por rangos con las tablas de Newell y McFarland (1987) y el análisis de varianza respectivamente

Cuadro VII. Comparación de los aciertos totales entre las pruebas Dúo Trío, Triángulo y Tétrada no direccionada bajo cada condición en estudio mediante el método de ordenamiento por rangos con las tablas de Newell y McFarland (1987) con un $\alpha=0,05$ para 27 panelistas.

Condición	Prueba	Promedio
Solo con enjuague	Tétrada	<u>7,741</u> ^a
Con enjuague y “primer”	Dúo Trío	<u>7,630</u> ^a
Con enjuague y “primer”	Tétrada	7,556 ^a
Sin enjuague ni “primer”	Tétrada	7,259 ^a
Sin enjuague ni “primer”	Dúo Trío	6,963 ^a
Solo con enjuague	Triángulo	<u>6,704</u> ^{ab}
Con enjuague entre cada muestra	Tétrada	6,481 ^{ab}
Sin enjuague ni “primer”	Triángulo	6,481 ^{ab}
Solo con enjuague	Dúo Trío	6,000 ^{ab}
Con enjuague entre cada muestra	Dúo Trío	5,889 ^{ab}
Con enjuague y “primer”	Triángulo	5,667 ^{ab}
Con enjuague entre cada muestra	Triángulo	<u>3,630</u> ^b

Es importante resaltar que el análisis no paramétrico es poco sensible, como se aprecia en el Cuadro VII. De acuerdo con este cuadro, se puede ver que únicamente el tratamiento Triángulo con enjuague entre muestras fue significativamente diferente al tratamiento de Tétrada solo con enjuague, con

enjuague y “primer” al inicio de cada set y sin enjuague ni “primer”. También fue significativamente diferente respecto al tratamiento Dúo Trío con enjuague y primer al inicio de cada set y sin enjuague ni primer. De todos los tratamientos, el valor mayor fue para téttrada con enjuague al inicio de cada set.

Sin tomar en cuenta las pocas diferencias significativas, se considera cada método de discriminación para determinar cuáles condiciones mejoraron la sensibilidad promedio de los jueces. De ahí se obtuvo que la mejor condición para Téttrada no direccionada y Triángulo resultó ser solo con enjuague y para la prueba Dúo Trío la condición más favorable fue con enjuague y “primer” (ver los promedios subrayados, Cuadro VII).

Cuadro VIII. Comparación de los aciertos totales entre las pruebas Dúo Trío, Triángulo y Téttrada bajo cada condición en estudio mediante el método de análisis de varianza con un $\alpha=0,05$ para 27 panelistas.

Condición	Prueba	Promedio
Con enjuague y “primer”	Dúo Trío	<u>6,407</u> ^a
Sin enjuague ni “primer”	Dúo Trío	6,148 ^{ab}
Sin enjuague ni “primer”	Triángulo	<u>5,778</u> ^{abc}
Solo con enjuague	Dúo Trío	5,778 ^{abc}
Sin enjuague ni “primer”	Téttrada	<u>5,704</u> ^{abc}
Solo con enjuague	Téttrada	5,630 ^{bcd}
Solo con enjuague	Triángulo	5,630 ^{bcd}
Con enjuague entre cada muestra	Dúo Trío	5,407 ^{bcd}
Con enjuague y “primer”	Téttrada	5,407 ^{bcd}
Con enjuague y “primer”	Triángulo	5,185 ^{cd}
Con enjuague entre cada muestra	Téttrada	4,926 ^d
Con enjuague entre cada muestra	Triángulo	4,111 ^e

De acuerdo con el Cuadro VIII, se puede ver que los tratamientos que resultaron más favorables para los panelistas para mantener o mejorar su desempeño fue la prueba Dúo Trío con enjuague y “primer”, sin enjuague ni “primer” y solo con enjuague. También con la prueba Triángulo sin enjuague ni “primer” y la prueba Téttrada sin enjuague ni “primer”. Por otro lado, el tratamiento con la prueba

Triángulo y enjuague entre cada una de las muestras presentó diferencias significativas con el resto de los tratamientos.

Sin tomar en cuenta las pocas diferencias significativas, se considera cada método de discriminación para determinar cuáles condiciones mejoraron la sensibilidad promedio de los jueces. De ahí se obtuvo que la mejor condición para Tétrada no direccionada y Triángulo resultó ser sin enjuague ni “primer” y para la prueba Dúo Trío la condición más favorable fue con enjuague y “primer” (ver los promedios subrayados, Cuadro VIII).

Ambos métodos de análisis estadístico señalan que el tratamiento menos favorable fue para la prueba Triángulo con enjuague entre cada una de las muestras. Sin embargo, se presentan diferencias respecto al tratamiento que resultó ser más eficiente, ya que el método de ordenamiento por rangos (no paramétrico) señala que el tratamiento más eficaz fue la Tétrada solo con enjuague, mientras que el análisis de varianza expone que es la prueba Dúo Trío con enjuague y “primer”. Aunado a esto, el análisis por tablas Newell McFarland señala que entre dichos tratamientos no existe diferencia significativa, mientras que según el ANDEVA si la hay.

Ahora bien, tomando en cuenta los resultados del análisis al pasar las rondas con el parámetro d' presentes en el Cuadro IV, se puede apreciar que los valores d' aumentan al pasar las rondas para la prueba Dúo Trío con enjuague y “primer”, y además presenta valores más grandes con respecto al tratamiento con la prueba Tétrada solo con enjuague. Por lo que se podría inferir que el tratamiento más favorable corresponde a la prueba Dúo Trío con enjuague y “primer”.

Las discrepancias entre los métodos utilizados pueden deberse a que, a la hora de realizar los ordenamientos para cada uno de los tratamientos, muchos de estos tuvieron el mismo número de aciertos por parte de los panelistas. Esto provoca que por el método de ordenamiento de rangos algunos tratamientos se les asignaran los mismos valores, mientras que, en el análisis de varianza al ser más exacto, a cada tratamiento se le asigna un valor distinto (promedio).

De ahí que, para concluir tomando en cuenta los dos análisis de datos para las tres pruebas de discriminación se utilizaron los resultados provenientes del ANDEVA, donde se apreció que el Dúo trío con enjuague y “primer” y el Triángulo y Tétrada sin enjuague ni “primer” fueron las condiciones que mejoraron la sensibilidad promedio de los panelistas, resultados que coinciden con el análisis hecho con los cambios en los valores de d' a lo largo de las rondas.

6. CONCLUSIONES

Se logró observar que los panelistas empezaron a perder la sensibilidad una vez finalizada la tercera ronda por fatiga.

De las pruebas direccionadas, la prueba 2-AFC resultó ser la más efectiva para mantener o mejorar el desempeño de los panelistas utilizando cualquier condición.

La condición que resultó ser más efectiva para las pruebas direccionadas 2-AFC, 3-AFC y Tétrada direccionada fue la que no presentó enjuague ni primer.

En las pruebas no direccionadas hay diferencias en la precisión para identificar cuál fue la prueba que resultó ser más efectiva para mantener o mejorar el desempeño de los panelistas, ya que con el método no paramétrico fue más sensible la prueba Dúo Trío y con el método paramétrico (ANDEVA) fue la prueba Tétrada.

Los tratamientos que resultaron más efectivos para las pruebas no direccionadas son: prueba Dúo Trío con enjuague y “primer” y para Triángulo y Tétrada sin enjuague ni “primer”.

El “primer” resultó efectivo en mantener o mejorar el desempeño de los panelistas solo para las pruebas Dúo Trío y Tétrada no direccionada.

No existe una condición universal que mantenga o ayude a mejorar el desempeño de los panelistas en pruebas direccionadas o no direccionadas, sino que este está sujeto al tipo de prueba, cantidad de estímulos y réplicas.

7. RECOMENDACIONES

Se recomienda el uso de otra matriz para llevar a cabo la investigación ya que por medio de varias pruebas preliminares y durante la realización de las pruebas definitivas se pudo comprobar que la bebida de naranja a base de bebida en polvo marca Tang®, resultó ser muy compleja.

Se recomienda realizar el estudio tomando en cuenta las diferencias de sensibilidad y potencia estadística entre las pruebas direccionadas y no direccionadas.

8. REFERENCIAS BIBLIOGRÁFICAS

- Angulo, O; O'Mahony, M. 2009. Aplicación del modelo Thrustone a las pruebas sensoriales de diferencia. Archivos latinoamericanos de nutrición 59 (4): 349-357.
- Ennis, JM; Jesionka, V. 2011. The Power Of Sensory Discrimination Methods Revisited. Journal of Sensory Studies 26(5):371-382.
- Ennis, JM; Rousseau, B; Ennis, DM. 2014. Sensory difference tests as measurement instruments: A review of recent advances. Journal of Sensory Studies 29(2):89-102.
- Fernandez, S. (s. f.). Análisis de la Varianza : 1 Factor. Madrid, s.e.
- Ishii, R; O'Mahony, M; Rousseau, B. 2014. Triangle and tetrad protocols: Small sensory differences, resampling and consumer relevance (en línea). Food Quality and Preference 31(1):49-55.
- Lawless, H; Heymann, H. 2018. Sensory Evaluation of Food. 2th ed. Springer (ed.). NY, s.e. 211 p.
- Lee, YM; Chae, JE; Lee, HS. 2009. Effects of order of tasting in sensory difference tests using apple juice stimuli: Development of a new model. Journal of Food Science 74(6).
- Lucak, CL. 2008. Determination of various palate cleanser efficacies for representative food types. s.l., The Ohio State University. .
- Mata-Garcia, M; Angulo, O; O'Mahony, M. 2007. On warm-up. Journal of Sensory Studies 22(2):187-193.
- McClure, S; Lawless, HT. 2010. Comparison of the triangle and a self-defined two alternative forced choice test (en línea). Food Quality and Preference 21(5):547-552.
- Newell, GJ; MacFarlane, JD. 1987. Expanded Tables for Multiple Comparison Procedures in the Analysis of Ranked Data. Journal of Food Science 52(6):1721-1725.
- O'Mahony, M. 1972. Salt taste sensitivity: a signal detection approach. Perception 1(4):459-464.
- O'Mahony, M. y Godman, L. 1974. The effect of interstimulus procedures on salt taste thresholds. Perception and Psychophysics. 16(3):494-502.
- O'Mahony, M. 1995. ¿ Who told you the triangle test was simple? 6:227-238.
- O'Mahony, M; Rousseau, B. 2003. Discrimination testing: A few ideas, old and new.

- Food Quality and Preference 14(2):157-164.
- O'Mahony, M; Rousseau, B; Stroh, S. 2002. Investigating more powerful discrimination tests with consumers: Effects of memory and response bias. Food Quality and Preference 13(1):39-45.
- O'Mahony, M. y Lee, H.S. 2004. Sensory Difference testing: Thurstonian Models. Food Science and Biotechnology 13 (5): 1- 6.
- O'Mahony, M. 2018. Comunicación personal. 13 Agosto 2018.
- Olivas-Gastélum, R. 2008. Comparación y evaluación de las pruebas de diferencia Dúo-trío, triangular, ABX e igual diferente. s.l., s.e. p. 66-80.
- Paiz, M; Bustos, I. 2009. Facultad De Ciencias , Tecnología Y Ambiente (en línea). s.l., s.e. 195 p.
- Ramirez, M. 2016. Comparación entre el uso de las tablas binomiales sin repeticiones y las correcciones estadísticas con el uso de repeticiones, como métodos de análisis de pruebas de discriminación sensorial triángulo y 2-AFC. s.l., Universidad de Costa Rica. .
- Severiano, P; Gomez, D; Méndez, C; Pedrero, D; Gómez, C; Ríos, S; Escamilla, A; Utrera, M. (s. f.). Manual de Evaluación Sensorial. s.l., s.e.
- Villa, E. 2014. Métodos estadísticos para pruebas sensoriales en la industria (en línea). . Disponible en [http://7vpe.eventos.cimat.mx/sites/7vpe/files/Platica-VPyE2014_Parte lyII.pdf](http://7vpe.eventos.cimat.mx/sites/7vpe/files/Platica-VPyE2014_Parte%20II.pdf).
- Watts, BM; Ylimaki, GL; Jeffert, LE; Elías, LG. 1989. Metodos Sensoriales Basicos Para La Evaluación De Alimentos. International Development Research Centre, Ottawa.
- Worch, T y Delcher, R. 2013. A practical guideline for discrimination testing combining both the proportion of discriminations and thurstonian approaches. Journal of Sensory Studies 28: 396-40.
- Xia, Y; Zhang, J; Zhang, X; Ishii, R; Zhong, F; O'Mahony, M. 2015. Tetrads, triads and pairs: Experiments in self-specification (en línea). Food Quality and Preference 40(PA):97-105.

9. ANEXOS

ANEXO A. Resultados obtenidos en la prueba preliminar de umbral de diferencia apenas perceptible (JND) utilizando una concentración estándar de 60 g/L de Tang con 40 panelistas.

1: Escogieron la concentración estándar 2: Escogieron otra concentración

Fecha	Hora	Par 1	Par 2	Par 3	Par 4	Par 5	Par 6
23/08/2018	09:08	1	1	1	2	2	2
23/08/2018	09:14	1	1	1	1	2	1
23/08/2018	09:18	1	1	1	2	1	2
23/08/2018	09:21	1	2	1	1	1	1
23/08/2018	09:21	1	1	1	2	2	2
23/08/2018	09:26	1	1	1	2	2	2
23/08/2018	09:26	1	2	1	1	1	2
23/08/2018	09:30	1	1	1	2	1	1
23/08/2018	09:30	2	1	1	2	2	2
23/08/2018	09:31	1	1	1	2	2	2
23/08/2018	09:33	1	1	2	2	1	2
23/08/2018	09:34	2	2	1	2	1	2
23/08/2018	09:35	1	1	1	2	2	2
23/08/2018	09:37	1	1	1	2	2	2
23/08/2018	09:37	1	2	1	1	2	2
23/08/2018	09:39	1	1	2	2	2	2
23/08/2018	09:40	1	1	1	1	2	2
23/08/2018	09:40	1	1	2	2	1	2
23/08/2018	09:44	1	2	1	2	2	2
23/08/2018	09:44	1	1	1	1	2	2
23/08/2018	09:45	1	1	1	2	2	2
23/08/2018	09:47	1	2	1	2	2	2
23/08/2018	09:47	1	1	2	1	2	2
23/08/2018	09:48	1	2	2	1	1	1
23/08/2018	09:50	2	2	2	1	1	2
23/08/2018	09:52	1	2	1	2	1	2
23/08/2018	09:52	1	1	1	2	2	2
23/08/2018	09:52	1	1	1	2	2	2
23/08/2018	09:52	1	1	1	2	2	2
23/08/2018	09:54	1	1	1	2	2	2
23/08/2018	09:55	1	2	2	1	2	2
23/08/2018	09:58	1	1	1	1	2	2

23/08/2018	09:59	1	2	2	1	1	1
23/08/2018	10:00	1	1	1	1	2	1
23/08/2018	10:00	1	1	1	2	2	1
23/08/2018	10:01	1	1	2	2	2	1
23/08/2018	10:02	1	1	1	2	1	2
23/08/2018	10:03	2	2	2	2	1	1
23/08/2018	10:04	1	2	2	1	2	1
23/08/2018	10:05	1	1	1	1	1	1

ANEXO B. Resultados obtenidos en la prueba preliminar de 2-AFC con atributo intenso o dulce y escoger la concentración adecuadamente confundible con la concentración estándar 60 g/L de Tang con 70 panelistas.

1: Concentración estándar, 2: Otra concentración.

Fecha	Hora	Par 40/60 intenso	Par 52/60 intenso	Par 80/60 intenso	Par 40/60 dulce	Par 52/60 dulce	Par 80/60 dulce
30/08/2018	09:40	1	1	1	1	1	1
30/08/2018	09:40	1	1	1	1	1	1
30/08/2018	09:40	1	1	1	1	1	1
30/08/2018	09:40	1	1	1	1	1	1
30/08/2018	09:40	1	1	1	1	1	1
30/08/2018	09:40	1	1	1	1	1	1
30/08/2018	10:17	1	1	1	1	1	1
30/08/2018	10:18	1	1	1	1	1	1
30/08/2018	10:18	1	2	1	1	1	1
30/08/2018	10:18	1	1	1	1	1	1
30/08/2018	10:18	1	1	1	1	1	1
30/08/2018	10:18	1	1	1	1	1	1
30/08/2018	10:23	1	1	1	1	2	1
30/08/2018	10:24	1	1	1	1	2	1
30/08/2018	10:25	1	1	1	1	1	1
30/08/2018	10:25	1	1	1	1	1	1
30/08/2018	10:25	1	2	1	2	2	2
30/08/2018	10:27	1	1	2	1	1	2
30/08/2018	10:30	2	2	2	1	1	1
30/08/2018	10:30	1	1	1	1	1	1
30/08/2018	10:31	1	1	1	1	1	1
30/08/2018	10:34	1	1	1	1	1	1
30/08/2018	10:35	2	1	1	2	1	1
30/08/2018	10:39	1	2	1	1	1	1
30/08/2018	10:40	1	1	1	1	1	1
30/08/2018	10:40	1	1	1	1	1	2
30/08/2018	10:42	1	1	2	1	1	2
30/08/2018	10:45	1	1	1	1	1	1
30/08/2018	10:47	1	1	1	1	1	1
30/08/2018	10:49	1	1	1	1	2	1
30/08/2018	10:54	1	1	1	1	1	1
30/08/2018	10:56	1	2	1	1	1	2
30/08/2018	10:56	1	2	1	1	2	1
30/08/2018	10:59	2	2	1	1	1	1

30/08/2018	11:00	1	2	1	1	1	1
30/08/2018	11:01	1	1	1	1	1	1
30/08/2018	11:01	1	1	1	1	1	1
30/08/2018	11:02	1	1	1	1	1	1
30/08/2018	11:03	1	1	1	1	1	1
30/08/2018	11:07	1	1	1	1	1	1
30/08/2018	11:09	1	1	1	1	1	1
30/08/2018	11:10	1	1	1	1	2	1
30/08/2018	11:10	1	2	2	1	2	1
30/08/2018	11:10	1	1	1	1	1	1
30/08/2018	11:12	1	2	1	1	2	2
30/08/2018	11:13	1	1	1	1	1	1
30/08/2018	11:22	2	2	1	1	1	1
30/08/2018	11:24	1	1	1	1	1	1
30/08/2018	11:24	1	1	1	2	1	1
30/08/2018	11:24	1	2	1	1	1	1
30/08/2018	11:25	1	2	1	1	1	1
30/08/2018	11:27	1	1	1	1	1	1
30/08/2018	11:31	1	2	1	1	1	1
30/08/2018	11:31	1	1	1	1	1	1
30/08/2018	11:31	1	1	1	1	1	1
30/08/2018	11:31	2	2	1	1	1	1
30/08/2018	11:32	1	1	1	1	1	1
30/08/2018	11:52	1	1	1	1	1	1
30/08/2018	11:59	1	1	1	1	2	1
30/08/2018	12:00	1	1	1	1	1	1
30/08/2018	12:00	1	1	1	1	1	1
30/08/2018	12:00	1	1	1	1	2	1
30/08/2018	12:01	1	1	1	1	1	1
30/08/2018	12:01	1	2	1	2	2	1
30/08/2018	12:08	1	1	1	1	2	1
30/08/2018	12:08	2	1	2	1	1	1
30/08/2018	12:08	1	2	2	1	1	1
30/08/2018	12:08	1	1	1	1	2	1
30/08/2018	12:16	1	1	1	1	1	1
30/08/2018	12:16	1	1	1	1	1	1

ANEXO C. Ejemplo de hoja de respuestas de cada panelista para cada sesión que se elaboró manualmente a partir de los códigos generados en el programa FIZZ.

RONDA	C1	C2	C3	C4	SECUENCIA			
1	287	824	1	2	TANGDEBIL	TANGFUERTE		
1	854	265	676	2	TANGFUERTE	TANGDEBIL	TANGDEBIL	
1	768	124	357	590	TANGDEBIL	TANGFUERTE	TANGDEBIL	TANGFUERTE
1	595	104	743	382	TANGDEBIL	TANGFUERTE	TANGDEBIL	TANGFUERTE
1	757	870	983	1	TANGDEBIL	TANGFUERTE	TANGFUERTE	
1	902	143	1	2	TANGDEBIL	TANGFUERTE		
2	064	527	1	2	TANGDEBIL	TANGFUERTE		
2	23	536	339	1	TANGDEBIL	TANGDEBIL	TANGFUERTE	
2	341	108	985	752	TANGDEBIL	TANGFUERTE	TANGFUERTE	TANGDEBIL
2	791	194	307	420	TANGDEBIL	TANGDEBIL	TANGFUERTE	TANGFUERTE
2	463	537	284	2	TANGFUERTE	TANGFUERTE	TANGDEBIL	
2	186	905	2	1	TANGFUERTE	TANGDEBIL		
3	402	853	2	1	TANGFUERTE	TANGDEBIL		
3	534	197	860	1	TANGDEBIL	TANGFUERTE	TANGDEBIL	
3	504	391	278	875	TANGDEBIL	TANGDEBIL	TANGFUERTE	TANGFUERTE
3	243	130	727	614	TANGDEBIL	TANGFUERTE	TANGFUERTE	TANGDEBIL
3	121	524	637	2	TANGFUERTE	TANGDEBIL	TANGFUERTE	
3	952	489	1	2	TANGDEBIL	TANGFUERTE		
4	88	637	2	1	TANGFUERTE	TANGDEBIL		
4	347	884	421	1	TANGDEBIL	TANGFUERTE	TANGFUERTE	
4	610	61	512	963	TANGDEBIL	TANGDEBIL	TANGFUERTE	TANGFUERTE
4	199	480	931	382	TANGFUERTE	TANGDEBIL	TANGDEBIL	TANGFUERTE
4	599	262	925	2	TANGFUERTE	TANGDEBIL	TANGDEBIL	
4	894	347	2	1	TANGFUERTE	TANGDEBIL		

