

*Modulation of bistable perception: a study based on multimodal stimulation and registries of oculomotor activity**

Fernando Marroquín-Ciendúa**
Guillermo Rodríguez-Martínez***
Harold Germán Rodríguez-Celis****

* Este estudio fue financiado por la Universidad Jorge Tadeo Lozano y cofinanciado por la empresa NEUROMIND S.A.S. Con el código de aprobación 918-17-18 de convocatoria interna convocada por la Dirección de Investigación, Creación y Extensión de la Universidad Jorge Tadeo Lozano.

Los autores expresan sus agradecimientos a la Dirección de Investigación, Creación y Extensión (DICE) de la Universidad Jorge Tadeo Lozano por haber financiado el presente estudio. Así mismo, los apoyos brindados por la Escuela de Publicidad y al equipo científico de la empresa Neuromind S.A.S., en cabeza de su gerente general Henry Castillo, y al Doctor Pedro Joel Rosa por sus aportes.

** Doctor en Psicología. Profesor Asociado II, Escuela de Publicidad, Universidad Jorge Tadeo Lozano. Orcid: 0000-0002-2213-4566 | Correspondencia: fernando.marroquinc@utadeo.edu.co

*** Doctor en Psicología. Profesor Titular, Escuela de Publicidad, Universidad Jorge Tadeo Lozano. Orcid: 0000-0003-4329-5745 | Correspondencia: guillermo.rodri-guez@utadeo.edu.co

**** Protoplast Desarrollos Tecnológicos S.A.S. Doctorando en Estudios Sociales. Centro de Pensamiento Crítico y Científico para el Desarrollo Consciente. Orcid: 0000-0001-6232-3540 | Correspondencia: info@talamoo.com

Modulación de la percepción biestable: un estudio basado en estimulación multimodal y registros de actividad oculomotora*

Cómo citar este artículo: Marroquín-Ciendúa, F., Rodríguez-Martínez, G., & Rodríguez-Celis, H.G. (2020). Modulación de la percepción biestable: un estudio basado en estimulación multimodal y registros de actividad oculomotora. *Tesis Psicológica*, 15(1), 106-124. <https://doi.org/10.37511/tesis.v15n1a6>

Recibido: septiembre 19 de 2019

Revisado: diciembre 18 de 2019

Aprobado: abril 15 de 2020

ABSTRACT

Background: Bistable images can be interpreted in two different ways, the one the observer does with both the area of the visual stimulus he is looking at and the exogenous information provided simultaneously with the exposure of the ambiguous image. When a determined perceptual configuration of a visual stimulus is made, some modulative factors associated with the observer's ocular fixations play a role, as well as the modulation that the semantic load from another stimulus can exert. **Objectives:** Consequently, the problem arises as to whether in the presence of a multimodal audiovisual stimulation in which the audio has a semantic charge associated with the percepts of a bistable image, a semantic modulation is actually produced. Also, additionally, if it is possible to recognize an association between the ocular fixations made over the bistable image and the percept recognized from that image. **Methodology:** An experimental task was designed to solve this question. In it, thirty-five (35) participants observed the bistable image *My girlfriend or my mother-in-law* in a fixed device of oculomotor activity registry of 120 Hz., while they were hearing different voice tones. **Results:** Significant differences between the duration of the congruent and non-congruent interpretations in the audio were found. Evidence of the relation between the ocular fixations displayed in the areas of analysis and the reported visual percepts with the acoustic modulator was also found. **Conclusions:** The ocular fixations, such as the auditory semantic load, are related to the perception of the bistable image.

Keywords: visual perception, visual attention, perceptual restructuring, semantic memory, ocular movements, ocular fixation.

RESUMEN

Antecedentes: Las imágenes biestables tienen la posibilidad de ser interpretadas de dos maneras diferentes. La que hace el observador tanto con el área del estímulo visual que está mirando, como con información exógena provista de manera simultánea a la exposición de la imagen ambigua. Cuando se hace una determinada configuración perceptual de un estímulo visual, factores moduladores asociados a las fijaciones oculares del observador juegan un rol, lo mismo que la modulación que puede ejercer la carga semántica proveniente de otro estímulo. **Objetivos:** Emerge, en consecuencia, el problema de, si ante la presencia de una estimulación multimodal audiovisual en la que el audio tiene carga semántica asociada a los percepts de una imagen biestable, se produce efectivamente una modulación semántica y si, adicionalmente, es posible reconocer una asociación entre las fijaciones oculares hechas sobre la imagen biestable y el percepto que de ella se reconoce. **Metodología:** Para resolver esta cuestión, se diseñó una tarea experimental en la que treinta y cinco (35) participantes observaron en un dispositivo fijo de registro de actividad oculomotora de 120 Hz. la imagen biestable *My girlfriend or my mother-in-law* mientras escuchaban diferentes tonos de voz. **Resultados:** se encontraron diferencias significativas entre las duraciones de las interpretaciones congruentes y no congruentes con el audio. También se encontró evidencia de la relación entre las fijaciones oculares manifestadas en las áreas de análisis y los percepts visuales reportados. El entendimiento de los audios moduladores no se constituyó en un factor que repercutiera en las percepciones visuales congruentes con el modulador acústico. **Conclusiones:** se concluye que las fijaciones oculares como la carga semántica auditiva guardan relación con la percepción de la imagen biestable.

Palabras clave: percepción visual, atención visual, restructuración perceptual, memoria semántica, movimientos oculares, fijación ocular.

Introducción

Una imagen biestable es una composición visual diseñada para que un observador la pueda percibir de maneras diferentes, de manera tal que sus posibles interpretaciones no puedan darse en una misma unidad de tiempo (Rodríguez & Castillo, 2018a). Así, la percepción biestable es el fenómeno perceptual por el cual un observador interpreta de dos maneras diferentes un mismo estímulo (Borisjuk, Chik & Kazanovich, 2009; Grossmann & Dobbins, 2006; Leopold & Logothetis, 1999; Long & Toppino, 2004; Pressnitzer & Hupé, 2006; van Loon et al., 2013). Manteniéndose invariable el estímulo (en el caso de una imagen de tipo estático), o encontrando movimientos u oscilaciones (para el caso de estímulos visuales biestables o multiestables dinámicos), el observador cambia de una interpretación a otra, debido a que el estímulo ofrece varias posibilidades de interpretación (Brascamp, Sterzer, Blake & Knapen, 2018; Moreno-Bote, Rinzel & Rubin, 2007; Schauer, Kanai & Brascamp, 2016), sin que estas puedan ser percibidas simultáneamente (Weilnhammer, Stuke, Hesselmann, Sterzer & Schmack, 2017; Xiaogang et al., 2017; Yamamoto & Yamamoto, 2006), puesto que los estímulos visuales biestables (ver figura 1) admiten dos posibles perceptos (Sterzer, Russ, Preibisch & Kleinschmidt, 2002), o más de dos, para el caso de imágenes multiestables, se les puede denominar también figuras ambiguas (Brouwer & van Ee, 2006; Okazaki, Kaneko, Yumoto & Arima, 2008; Runnova et al., 2016), o estímulos reversibles (Davidson & Pitts, 2014). Así mismo, el fenómeno de bistabilidad perceptual visual puede ser denominado también biestabilidad visual (Intaité, Kovisto & Castelo-Branco, 2014); el salto o cambio de una interpretación a otra se denomina reversibilidad perceptual (Clément & Demel, 2012; Rodríguez & Castillo, 2018a). Estos cambios entre uno y otro percepto tienden a ser de naturaleza estocástica

(Denham et al., 2012), en razón a que factores de tipo sensorial y de tipo cognitivo pueden implicarse durante la decodificación de los estímulos visuales biestables (Rodríguez & Castillo, 2018a).

Figura 1. Imágenes biestables



Nota: a la izquierda, la figura *The Rat-Man*. En el centro, una imagen biestable de tipo perspectivo. A la derecha, La escalera de Schroder.

Fuente: (Chastain & Burnham, 1975; Clément & Eckardt, 2005).

Las características físicas de las imágenes biestables propician la reversibilidad perceptual. Son imágenes diseñadas para que el observador tenga la posibilidad de reconocer dos diferentes perceptos, nunca los dos al mismo tiempo (Brouwer & van Ee, 2006). La interpretación depende de las áreas de la imagen donde el observador hace las fijaciones oculares (Gale & Findlay, 1983; Hsiao, Chen, Spence & Yeh, 2012), lo mismo que del recorrido visual que hace el receptor durante la correspondiente observación (García-Pérez, 1989; García-Pérez, 1992). Cuando los aspectos físicos de la imagen biestable inciden en su interpretación, se hace referencia a una modulación de la percepción de abajo hacia arriba o ascendente (Kornmeier & Bach, 2005; Rodríguez & Castillo, 2018a). De otra parte, también se reconoce que la interpretación de una imagen biestable puede estar supeditada a un procesamiento de información exógena al estímulo visual (Meng & Tong, 2004). Esta información corresponde a conceptos que se involucran en el proceso perceptual y que están almacenados en memoria o pueden ser provistos desde otras modalidades sensoriales en simultánea con la observación del estímulo biestable, condicionando el resultado final de lo percibido visualmente. Es lo que en

términos de procesos psicológicos básicos se denomina procesamiento de arriba hacia abajo o descendente (Rodríguez & Castillo, 2018a; Sterzer, Kleinschmidt & Rees, 2009). Los mecanismos descendentes se hacen manifiestos por información o conocimientos previamente almacenados en la memoria (Kornmeier, Hein & Bach, 2009), o por información que de manera adicional incursiona en el sistema perceptual, generando un eco interpretativo en el momento de definir el significado de la imagen (Intaité, Noreika, Šoliūnas & Falter, 2013). Un ejemplo de este mecanismo de modulación de la percepción biestable de arriba a abajo se da cuando se utiliza estimulación auditiva con una carga semántica que influencia la percepción de la imagen biestable que tiene coherencia con la carga semántica del audio (Hsiao et al., 2012; Smith, Grabowecky & Suzuki, 2007).

Existen, por tanto, según lo ya explicado, dos mecanismos moduladores de la percepción visual biestable: uno ascendente (*bottom-up*) y otro descendente (*top-down*). Es por esto que la interpretación de las imágenes biestables se asume como un mecanismo dinámico que se modula de las dos maneras aquí mencionadas (Sterzer et al., 2009). La alternancia entre una y otra interpretación (reversibilidad perceptual) muchas veces es involuntaria (cuando se presentan modulaciones ascendentes), pero también es voluntaria, especialmente cuando el observador puede ejercer controles atencionales que dirigen la percepción (Brouwer & van Ee, 2006; Intaité, Koivisto, Rukšėnas & Revonsuo, 2010).

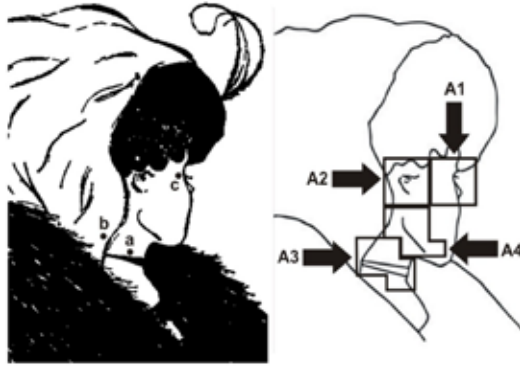
El modelo ascendente operante en la percepción biestable consiste en que una reversibilidad perceptual ocurre por mediación de mecanismos sensoriales, de tal suerte que cierto procesamiento perceptual da sustento a una específica configuración perceptual hasta que,

por efecto de la fatiga, el percepto competidor emerge, encontrando sustento en otro mecanismo (Kogo, Hermans, Stuer, van Ee & Wage-mans, 2015). En cuanto al modelo basado en procesamientos descendentes, la configuración de un percepto (entre los dos posibles), se da por una correspondencia con un procesamiento de información de tipo cognitivo que implica un reconocimiento de conceptos mediado por la memoria semántica (Barrera & Calderón, 2013; Rodríguez, 2016; Sterzer & Rees, 2009).

Áreas de fijación ocular como moduladores de la percepción biestable

La percepción de cada uno de los posibles perceptos de la imagen biestable *My girlfriend or my mother-in-law* (ver en figura 1), puede darse como consecuencia de fijaciones oculares hechas sobre áreas o partes específicas del estímulo visual (Hsiao et al., 2012). Se implica acá una influencia endógena sustentada en características físicas de los estímulos biestables (Brouwer & van Ee, 2006). Gale y Findlay (1983) demostraron que existen áreas críticas dentro de esa imagen biestable que favorecen la percepción de cada uno de los posibles perceptos. Habiendo realizado un análisis pormenorizado de líneas críticas constitutivas de la imagen biestable *My girlfriend or my mother-in-law*, se sugirió que ciertos trazos de la imagen posibilitaban que el observador percibiera mayormente una imagen (*mujer joven o mujer anciana*) sobre la otra. Haciendo un ejercicio de síntesis gráfica sobre la imagen original de Boring (1930), el estudio refirió cuatro áreas en específico, cada una de ellas con información visual útil para el reconocimiento de uno y otro percepto. Sobre la base de este estudio, se infirieron puntos de fijación atencional que en específico (ver figura 2) favorecen más a un percepto que al otro (Hsiao et al., 2012).

Figura 2. Áreas de modulación de la imagen biestable de Boring



Fuente: (Gale & Findlay, 1983).

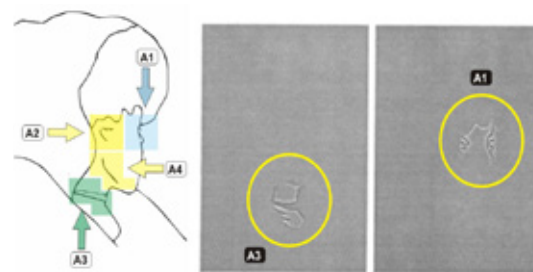
Sobre el modelo simplificado en línea utilizado por Gale y Findlay (1983), se aprecian las cuatro áreas de interés estimadas por ellos como críticas para la percepción de la imagen biestable *My girlfriend or my mother-in-law*. El área A1, moduladora del percepto mujer joven, contiene líneas definitorias del ojo y de la nariz de la mujer joven; la A2 define preponderantemente la oreja de la mujer joven y el ojo de la mujer anciana; la A3, moduladora del percepto mujer anciana, hace referencia a la boca de la mujer anciana; la A4 contiene una línea que define la nariz de la mujer anciana y, a su vez, un contorno de la quijada de la joven.

Por otra parte, García-Pérez (1989) realizó un estudio en el cual puso en evidencia una relativa no homogeneidad en los patrones de actividad oculomotora durante la observación de imágenes biestables. La homogeneidad visual es la percepción sostenida y enfocada de un determinado estímulo dentro del contexto de todo el flujo de información visual presente en el campo visual (García-Pérez, 1989). Con respecto a la imagen de Boring (1930) *My girlfriend or my mother-in-law*, el estudio refiere que áreas específicas de la imagen repercuten en el reconocimiento de características específicas de alguno de los perceptos posibles, de conformidad con lo que Gale y Findlay (1983) habían sugerido.

Los estudios de Gale y Findlay (1983), así como los de García-Pérez (1989), hacen referencia a qué áreas en particular inducen la percepción de cada uno de los posibles perceptos de la imagen biestable de Boring (1930). En efecto, el área que se corresponde con la quijada y la boca de la anciana no implica un favorecimiento hacia el percepto *mujer joven*. De hecho, supone un favorecimiento del percepto *mujer anciana*, pero no de la manera en que el área de los ojos de la mujer joven favorece al percepto *mujer joven* (García-Pérez, 1989). Esto puede explicar por qué el porcentaje de reportes congruentes con mujer anciana cuando se observa su área de modulación (un poco por encima del 60%) no es tan alto con relación a las congruencias manifestadas entre las fijaciones en el área moduladora de la mujer joven y el reporte del percepto *mujer joven*, expresadas en casi un 90% de los casos (ver Gale & Findlay, 1983).

En la figura 3, se establece un comparativo entre las áreas de modulación dispuestas en los estudios de Gale y Findlay (1983) y García-Pérez (1989). Para efectos del comparativo, se estimó la designación de áreas referidas en la figura 2, de manera tal que las convenciones usadas (A3 y A1) hacen referencia a las áreas moduladoras, la primera para el percepto *mujer anciana*, la segunda para el percepto *mujer joven*.

Figura 3. Áreas de modulación para la imagen biestable de Boring (1930)



Fuente: Gale & Findlay,(1983); García-Pérez,(1989).

Considerando los estudios de Gale y Findlay (1983) y de García-Pérez (1989) respecto de la imagen *My girlfriend or my mother-in-law*, se muestra, a la izquierda, la imagen basada del primer estudio; a la derecha, la referente al segundo estudio. El área A3 designa la zona de modulación para el percepto mujer anciana; A1 designa el área moduladora para el percepto mujer joven.

Una de las dificultades para realizar estudios sobre la relación que puede existir entre áreas observadas y perceptos configurados, es la falta de homogeneidad durante el proceso de observación del estímulo biestable (García-Pérez, 1989). Esto se debe a que son demasiados los factores que hacen que una persona mueva sus ojos, de manera tal, que los recorridos visuales no necesariamente tienen un mismo patrón entre persona y persona, más allá de condicionantes que puedan existir, derivados de las instrucciones impartidas en función de reportar uno u otro percepto. Para que se dé la homogeneidad, se hace necesario que el área de enfoque esté más bien constante durante el proceso de observación de una imagen, lo que también supone mayor permanencia de las fijaciones oculares (García-Pérez, 1989). Se encuentran otros estudios realizados sobre la injerencia que tienen las fijaciones oculares sobre la interpretación de imágenes biestables a manera de moduladores de abajo-arriba en los que se reseña la importancia de la atención sobre áreas particulares de los estímulos visuales biestables (Hsiao, Chen, Spence & Yeh, 2010) o sobre rasgos diferenciados de este tipo de imágenes (Chastain & Burnham, 1975). Otro estudio (Laubrock, Engbert & Kliegl, 2008) concluye que los movimientos oculares microsacádicos proporcionan una señal que permite predecir el modo en que se va a percibir la dirección de un estímulo visual ambiguo en movimiento. También se ha encontrado evidencia de la relevancia que tiene

en la interpretación, la primera fijación ocular en la observación de imágenes biestables (Chastain & Burnham, 1975; Gale & Findlay, 1983).

La voz humana como modulador semántico de la percepción biestable

Es posible usar una historia corta narrada oralmente (estimulación acústica) de modo tal que ella (su contenido) module la interpretación final de una imagen biestable (Hsiao et al., 2012). Como es de suponerse, la carga semántica de la historia tendrá que estar relacionada con el contenido semántico de una de las posibles interpretaciones de la imagen biestable (Balçetis & Dale, 2007). Este tipo de modulación implica una congruencia semántica multimodal, en que el fenómeno de modulación hace emergencia debido a la asociación existente entre la información provista desde las diferentes modalidades sensoriales, de modo tal que una unificación de la información visual y acústica es asumida por el observador (Hsiao et al., 2010). La interpretación que se hace en el sentido de que hay una unidad perceptual audiovisual es conocida como una unidad asumida o *unity assumption* (Vatakis & Spence, 2007). En ese orden de ideas, es posible observar la transmisión de la congruencia semántica a partir del uso de tonos de voz aislando el efecto que pueda producir la carga semántica de las palabras utilizadas, apelando a un idioma que sea desconocido para el perceptor (Hsiao et al., 2012). En efecto, solo los tonos de voz tienen la capacidad de modular la percepción visual (Chen & Spence, 2011), si el contenido semántico del tono (no de las palabras) guarda relación con el contenido del estímulo visual observado (Smith et al., 2007).

Lo anterior implica que cuando un observador está apreciando una figura biestable como *My girlfriend or my mother-in-law*, puede percibir a

la mujer joven si aparece en simultanea un modulador sonoro, que podría ser una voz juvenil femenina; en contraparte, si el modulador sonoro es una voz de anciana, por congruencia semántica, el observador podrá advertir en su plano perceptual la presencia de la mujer vieja (Hsiao et al., 2012). La memoria semántica, que se define como un conjunto de información registrada relativa a los conocimientos culturales y lingüísticos (Patiño, 2007), cumple un papel esencial en la manifestación del fenómeno de congruencia semántica, en los que se implican fenómenos de primado semántico (Goolkasian & Woodberry, 2010; Gulán & Valerjev, 2010). A su vez, la implicación del efecto de congruencia semántica reivindica el procesamiento visual descendente.

La modulación semántica supone una conexión contextual (o de contenido semántico), de manera tal que se hace manifiesto un efecto en la percepción producido por las relaciones en términos de contenido (o carga semántica) entre dos o más unidades procesadas perceptualmente (Rodríguez & Castillo, 2018b). En otras palabras, al estudiar los efectos de las modulaciones descendentes en términos de congruencia semántica cotejando los efectos moduladores ascendentes referidos a fijaciones oculares, se encuentra una prevalencia de este último efecto sobre el primero (ver Hsiao et al., 2010).

En consideración a lo anterior, este estudio se diseñó con el propósito de establecer si se hace manifiesto el efecto de congruencia semántica cuando se observa la imagen biestable *My girlfriend or my mother-in-law* mientras se escuchan audios moduladores (tonos de voz) con correspondencia semántica a cada uno de los posibles perceptos. Adicionalmente, se quiso establecer si las fijaciones oculares manifestadas en áreas críticas de modulación ascendente sugeridas en el estudio de Gale y Findlay (1983), efectivamente están asociadas a los reportes

perceptuales dados durante la observación de la imagen biestable mencionada. Así mismo, el estudio tuvo por propósito observar el efecto que podía tener el posible entendimiento del contenido textual del estímulo auditivo sobre los perceptos visuales reportados.

Metodología

Tipo y diseño

Este estudio, inscrito en la investigación básica de corte empírico-experimental, implicó un diseño intrasujeto mediante el cual se quiso observar el efecto modulador de la carga semántica de dos tonalidades de voces en la interpretación de una imagen biestable. Cada tonalidad de voz tenía una afinidad semántica con alguno de los posibles perceptos de la imagen biestable, siguiendo la hipótesis que será percibido mayormente el percepto de la imagen biestable relacionado con la tonalidad de voz con carga semántica congruente. Dado que se deseaba analizar el efecto modulador de los tonos de voz y no del contenido de las palabras emitidas, se tuvo en consideración el nivel de entendimiento de los monólogos como una covariable. También se quiso cotejar el nivel de asociación entre las áreas de la imagen observadas y los perceptos reportados. Se planteó la hipótesis de que las fijaciones oculares realizadas en las áreas de modulación de la imagen biestable están asociadas a los perceptos reportados.

Participantes

Treinta y cinco (35) voluntarios participaron en este estudio (57%=mujeres; 43%= hombres; rango de edad entre 18 y 37; promedio de la edad= 25.14; SD= 6,43). Seleccionados de manera no probabilística, los participantes debían cumplir con el criterio de no tener en su historia clínica reporte de antecedentes de daño cerebral.

Adicionalmente, se incluyeron únicamente personas que no reportaran problemas de visión ni necesidad del uso de aditamentos para corregir imperfecciones a nivel visual y auditivo. No se accedió a su historia clínica, sólo se tuvo en cuenta su auto-reporte. Las pruebas experimentales fueron llevadas a cabo en un laboratorio de psicología. Todos los participantes firmaron su correspondiente consentimiento informado.

Procedimiento

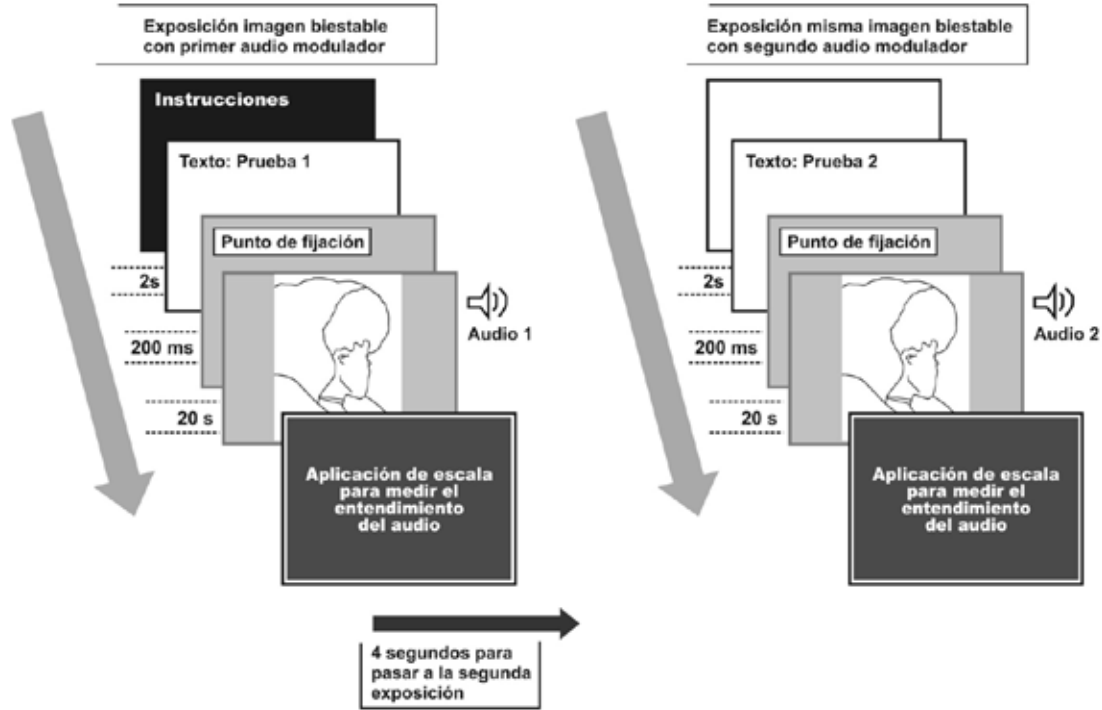
Para atender los objetivos de este estudio, se diseñó una tarea en la que cada participante tenía que observar en dos oportunidades, en un dispositivo fijo de registro de actividad oculomotora de 120 Hz. (referencia Tobii™ T120), la imagen biestable *My girlfriend or my mother-in-law* (a aproximadamente 60 cms. de distancia) mientras escuchaban la voz de una mujer a 54 dB SPL, mediante el uso de audífonos. La imagen se presentó en blanco y negro en un fondo gris (RGB= [127, 127, 127]). Un punto de fijación rojo colocado en un área de la imagen que no favoreciera ninguna posible interpretación de esta fue expuesto durante 200 ms. antes de la presentación de la imagen. La primera vez que cada participante observaba la imagen (durante 20 segundos), escuchaba el audio de una mujer anciana hablando.

En la segunda oportunidad, la imagen era la misma, pero el audio era el de una mujer joven hablando. Los audios fueron aleatoriamente

presentados (intercambiando su orden), para así tener un contrabalanceo y controlar la variable *orden de presentación*. Los dos monólogos estaban en idioma francés. Como se dijo, un audio se correspondía semánticamente con una posible interpretación de la imagen (mujer anciana), mientras que el otro lo hacía con la otra posible interpretación (mujer joven). El entendimiento de los monólogos fue tomado en cuenta. Para este propósito, todos los participantes fueron evaluados previamente en lo que refería a su nivel de dominio del idioma francés, mediante una escala de autovaloración de 0 a 9, donde cero implicaba ningún dominio del idioma y 9 total dominio. En la figura 4 se sintetiza el diseño de la tarea experimental.

La investigación, como se dijo en párrafos anteriores, buscaba establecer si los tonos de voz proveían contextos moduladores para las interpretaciones “mujer anciana” o “mujer joven”, tomando en consideración el entendimiento del contenido de los monólogos moduladores. Los participantes oprimían un botón de un ratón de computador, para reportar qué interpretación de la imagen estaban teniendo durante la experiencia. Para cada uno de los posibles perceptos correspondía un botón diferente. Adicionalmente, se hizo un análisis de las fijaciones oculares hechas a cada una de las áreas tipificadas por Gale y Findlay (1983), para establecer el nivel de asociación de las visitas oculares hechas a dichas áreas con el reporte dado en términos del percepto configurado.

Figura 4. Diseño de la tarea experimental



Nota: ms= milisegundos; s= segundos. La gráfica muestra la línea de tiempo de la tarea que cada participante tuvo que realizar.
Fuente: Autores

Análisis de datos

Con el ánimo de establecer si las fijaciones oculares manifestadas en áreas críticas de modulación estaban asociadas a los reportes perceptuales dados durante la observación de la imagen biestable, se utilizaron los datos procesados con el programa Tobii Studio (v.3.4.8), de manera tal que quedaran organizados el total de reportes perceptuales discriminados por cada uno de los posibles perceptos (mujer joven y mujer anciana), igualmente relacionados con las áreas de fijación ocular definidas para el estudio (áreas A1, A2, A3, A4; las zonas de la imagen que no correspondieran con estas áreas se codificaron como “fondo”).

Para efectos de establecer indicadores de modulación semántica, la duración de cada percepto visual congruente con el audio fue la medida

estimada para hacer los correspondientes análisis estadísticos. Dicho de otra manera, cada vez que un participante reportó una interpretación que fuera congruente semánticamente con el modulador auditivo, se tomó el tiempo desde el momento del reporte hasta que se manifestara un reporte de la interpretación incongruente. De este modo, se obtuvieron las duraciones de las interpretaciones congruentes e incongruentes con los audios moduladores. Posteriormente se hizo la correspondiente prueba de normalidad para establecer el estadístico a implementar. En este caso, fue necesario recurrir a una prueba estadística no paramétrica en virtud de que la distribución de los datos no fue normal.

Para analizar el entendimiento del monólogo modulador se aplicó una dicotomización de la variable, de modo tal que los valores menores a 3 (tres) incluido, implicaran el no entendimiento

del contenido del audio, en tanto que valores superiores refirieron una comprensión del contenido verbal de cada audio. Se analizó el nivel de asociación entre el entendimiento del audio y la percepción de los dos posibles perceptos de la imagen biestable cuando ellos fueron congruentes con el audio modulador. Posteriormente, se realizó un análisis de covarianza (MANCOVA) para observar la incidencia del entendimiento del audio en el efecto modulador de los audios sobre las interpretaciones dadas a la imagen.

Resultados

Consideraciones éticas

Esta investigación fue debidamente aprobada por el comité de ética de la Universidad de Bogotá Jorge Tadeo Lozano. Las pruebas experimentales ejecutadas en el marco de la ejecución de este estudio fueron coordinadas y supervisadas por psicólogos con tarjeta profesional vigente, debidamente expedida por el Colegio Colombiano de Psicología (Colpsic).

Análisis de modulación de las fijaciones oculares

Considerando los resultados, los perceptos que se configuran durante la observación de la imagen biestable presentada de manera multimodal junto a un estímulo acústico tienen una asociación estadísticamente significativa con las fijaciones oculares manifestadas sobre áreas críticas de modulación ($\chi^2(4, N=107)=19.451, p=0.003$).

En la tabla 1, se aprecia la relación entre los reportes de los dos posibles perceptos de la imagen biestable (MJ, mujer joven; y MA, mujer anciana) y las áreas de interés (AOIs) visitadas:

Tabla 1. Relación entre reportes de perceptos y áreas de interés moduladoras (AOIs)

AOIs	Reporte MJ	Reporte MJ (%)	Reporte MA	Reporte MA (%)
A1	23	79%	5	20%
A2	22	50%	23	50%
A3	3	32%	5	67%
A4	4	40%	6	60%
Fondo	10	65%	6	36%

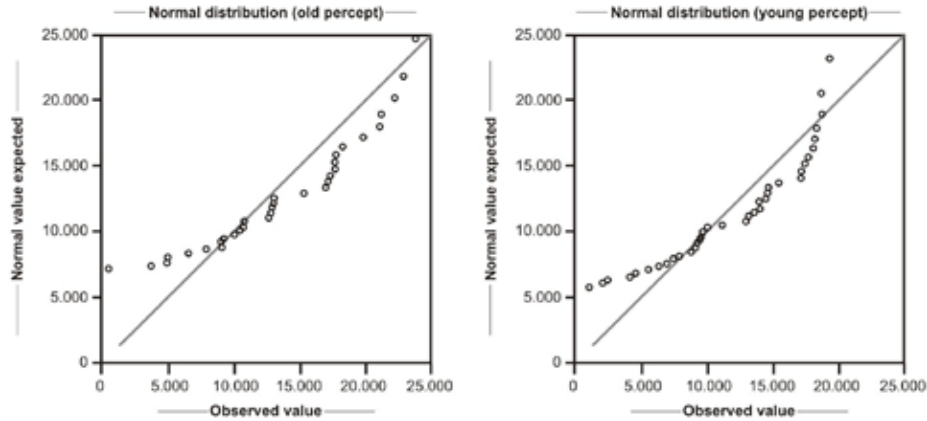
Nota: Prueba con estadístico Chi-Cuadrado ($\chi^2(4, N=107)=19.451$).
* $p<0.05$

Fuente: Autores

Análisis de modulación semántica

Se observó que la distribución de los datos (duraciones de las interpretaciones congruentes) no fue normal (ver figura 5). El estadístico aplicado fue la prueba de *Kolmogorov-Smirnov* ($p=0.000$, para la duración de la interpretación “mujer anciana” congruente con el audio modulador; $p=0.000$, para la duración de la interpretación “mujer joven” congruente con el audio modulador). El valor promedio de la duración de cada interpretación fue $M=6153.19$ ($SD=7875.74$) para “mujer anciana” y $M=7552.45$ ($SD=8353.43$) para “mujer joven” (ver comparación entre los promedios de los tiempos de duración de cada interpretación reportada en la figura 6).

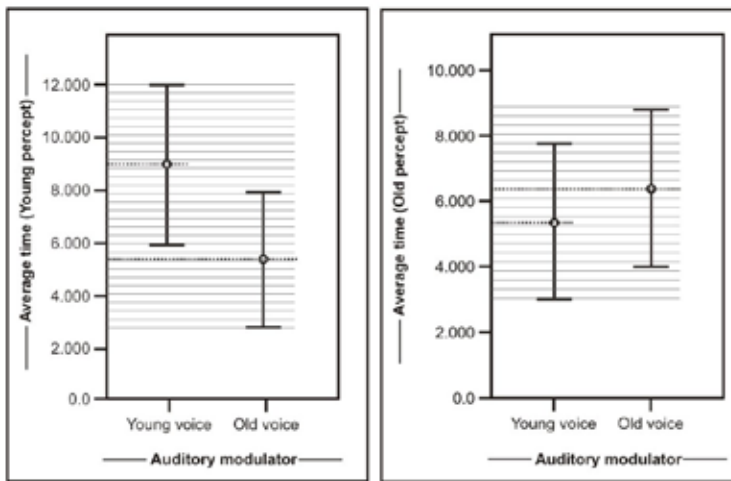
Figura 5. Distribución de los valores



Nota: duraciones en milisegundos de cada interpretación congruente reportada.

Fuente: Autores

Figura 7. Comparación entre los promedios de los tiempos de cada interpretación



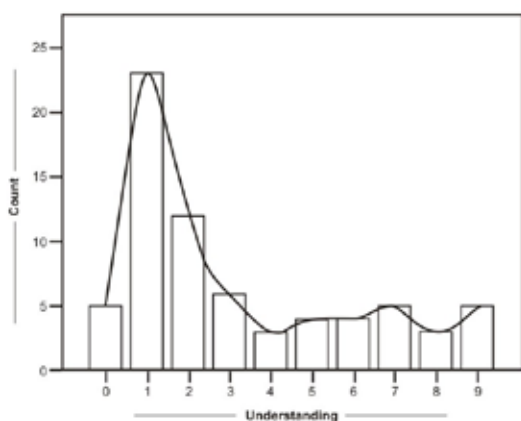
Nota: duración de cada interpretación reportada considerando los audios moduladores congruentes e incongruentes.

Fuente: Autores

La prueba no paramétrica realizada con posterioridad fue la prueba *U de Mann-Whitney*. En términos estadísticos se encontró evidencia para rechazar la hipótesis de igualdad en la distribución de la duración de la interpretación “mujer anciana” congruente con su audio modulador ($Mdn= 4914.70$; $U=701.5$; $p=0.0425$) y la duración de la interpretación “mujer joven” congruente con su audio modulador ($Mdn= 10599.58$; $U=471.0$; $p=0.0481$).

Hágase ahora referencia a la covariable *entendimiento del audio modulador*. Como se aprecia en la figura 7, puesto que la variable entendimiento tuvo mayor frecuencia en los valores menores, es decir, la mayoría de las personas que participaron en el estudio presentaron un bajo nivel de entendimiento del audio, se tomó como corte (para dicotomizar la variable) el valor 3 (como se mencionó en un apartado anterior), en el que se agrupaba el 60% de la información (esto con el fin de tener datos suficientes para el análisis de asociación).

Figura 7. Distribución de los valores referidos al entendimiento del audio modulador



Nota: La escala usada implicó mediciones hasta el valor 9; entre más bajo el valor, menor el entendimiento del audio; entre más alto, mayor su entendimiento.

Fuente: Autores

Los resultados muestran que no existe una relación marcada entre el entendimiento del audio y la percepción de los dos posibles perceptos de la imagen biestable cuando ellos fueron congruentes con el audio modulador. Los resultados expuestos a continuación muestran la relación entre las interpretaciones que fueron congruentes con los audios moduladores asumidos como “comprendidos” ($\chi^2=3.84$): audio de mujer anciana: $\chi^2(1, N=70)=1.553$, $p=0.213$ ($1.553 < 3.84$; $0.213 > 0.05$); audio de mujer joven: $\chi^2(1, N=70)=1.544$, $p=0.214$ ($1.544 < 3.84$; $0.214 > 0.05$); todas las asociaciones (ambos audios): $\chi^2(1, N=70)=2.307$, $p=0.129$ ($2.307 < 3.84$; $0.129 > 0.05$).

Un análisis de covarianza (MANCOVA) fue implementado para ratificar que el entendimiento del audio no tuvo una incidencia en el efecto modulador de los tonos de voz sobre las interpretaciones dadas a la imagen biestable. Así, el efecto del audio de la mujer joven sobre la interpretación del percepto visual congruente luego de controlar el entendimiento del audio no resultó ser significativo ($F(1, 33)=0.12$, $p=0.914$). De igual manera

sucedió con el efecto del audio de la mujer anciana sobre la interpretación de su percepto visual congruente ($F(1, 33)=0.460$, $p=0.502$).

Discusión

Desde los análisis y resultados obtenidos, se observa una relación significativa entre las áreas de fijación ocular analizadas (áreas A1, A2, A3, y A4) y los reportes dados en relación con las configuraciones perceptuales manifestadas. Es importante mencionar que en el estudio de Gale y Findlay (1983) al remover los trazos clave de áreas como la A1 (moduladora del percepto MJ), se produjo un descenso significativo de las visitas hechas a dicha área con respecto a la observación de la imagen completa ($t= 2.94$, $df= 4$, $p < 0.05$). En otras palabras, las líneas definitorias de los ojos de la mujer joven repercuten en el direccionamiento de la mirada sobre el área A1, por ende, en el reconocimiento del percepto MJ ($t= 3.04$, $df= 4$, $p < 0.05$) (Gale & Findlay, 1983). Si bien en el presente estudio no se revisó en detalle la incidencia de cada trazo dentro de cada área moduladora de análisis, como sí se hizo en el estudio precedente señalado, se advierte que sí es posible referenciar efectos moduladores de tipo ascendente sobre la percepción del estímulo biestable, sin que se llegue a una pormenorización del aporte de cada trazo particular de cada área en específico para cada uno de los posibles perceptos configurados de la imagen utilizada. No obstante, la asociación encontrada entre las áreas observadas por los participantes y los reportes perceptuales, reivindican la relación entre las fijaciones oculares y la percepción biestable, esto en línea con los hallazgos en que la actividad oculomotora está implicada en la desambiguación que un observador hace de una imagen biestable (Chastain & Burnham, 1975; Gale & Findlay, 1983; García-Pérez, 1989; García-Pérez, 1992; Hsiao et al., 2010; Hsiao et al., 2012).

De otra parte, los audios moduladores tuvieron una incidencia significativa en la interpretación de la imagen biestable. Se evidencia una diferencia a favor de la interpretación congruente con la tonalidad de voz provista como modulador auditivo. Estudios preliminares sobre modulación semántica de la percepción visual que fundamentaron la tarea visual en el paradigma de imágenes biestables, indicaron que, si bien es posible usar un tono de voz como modulador (Hsiao et al., 2012, Smith et al., 2007), el efecto de congruencia semántica se puede manifestar, aunque factores moduladores de tipo ascendente pueden impactar la percepción final (Hsiao et al., 2010), suscitando un efecto en la lectura de la congruencia semántica.

En efecto, existe una interacción entre los dos mecanismos moduladores, de manera que se puede argüir que la percepción biestable supone una bicausalidad (Kornmeier et al., 2009). Desde los hallazgos obtenidos en el presente estudio no es posible establecer cuál tipo de modulación tiene mayor incidencia en la percepción (pues no era uno de sus objetivos). Sin embargo, estudios precedentes indican que las fijaciones oculares pueden producir un mayor efecto sobre la interpretación de una imagen biestable que una modulación de tipo semántico (ver Hsiao et al., 2010).

Considérese adicionalmente, que en el dominio de la percepción biestable se ha asumido que condiciones propias del observador alusivas a su capacidad de aprendizaje y de rastreo del campo visual, entran a sumar en la manera en que se manifiestan las reversibilidades perceptuales, es decir, los cambios entre una y otra interpretación de la imagen biestable (Borisyuk et al., 2009). Esto sin considerar la naturaleza estocástica de las reversibilidades perceptuales (Denham et al., 2012), que dificulta el control sobre el fenómeno en laboratorio (Rodríguez & Castillo, 2018a), hecho que pudo repercutir en que la distribución

de los datos referidos a las duraciones de cada percepto no fuera normal. Adicionalmente, en el momento en que hace presencia un modulador acústico con carga semántica estimada como equivalente a los contenidos semánticos propios de las posibles interpretaciones de la imagen ambigua, el observador puede hacer una búsqueda de aquello que se asocie con lo percibido auditivamente, o retener dicha interpretación, todo mediante un control de sus movimientos oculares (Brouwer & van Ee, 2006).

Lo observado a partir de los resultados puede presuponer que es posible modular semánticamente la percepción de imágenes biestables, y que el tono de voz en sí mismo puede bastar para hacer la modulación, justo como se evidencia en otros estudios (Hsiao et al., 2012; Smith et al., 2007). Según lo indicaron los resultados de las pruebas estadísticas, para ambas interpretaciones congruentes con cada audio (mujer joven y mujer anciana), se encontró evidencia para aceptar la hipótesis de incidencia del audio modulador sobre los tiempos de configuración perceptual congruente con dicho modulador acústico.

Por su parte, el entendimiento del audio como covariable no fue un factor que impactara en el desempeño perceptual de los participantes, esto es, en sus reportes congruentes con el modulador acústico. No obstante, debe entenderse que el presente estudio presenta una limitante en el sentido de que la dicotomización aplicada a la variable se hizo de manera tal que, con las puntuaciones obtenidas, fuera posible hacer un análisis de asociación; como se refirió en párrafos anteriores, el punto de corte para asumir que había entendimiento fue el valor 3 (tres), el cual está por debajo del punto medio de la escala y que supone un resultado susceptible de discusión, el cual, a su vez, reclama de futuras investigaciones que permitan dilucidar el papel que juega el contenido del audio modulador diferenciadamente de la modulación mediada

solamente por tonalidades de voz y el reconocimiento que se haga de dichas tonalidades en términos de la identificación de alguna carga semántica relacionada con los posibles perceptos de la imagen biestable.

En consideración al análisis de covarianza realizado, se observó que el entendimiento del audio no tuvo una incidencia en el efecto modulador de los tonos de voz sobre las interpretaciones dadas a la imagen biestable. Como se dijo, estudios posteriores podrán ahondar en la modulación semántica que pueden ejercer los tonos de voz, proponiendo diseños experimentales en los que se coteje la congruencia semántica de conformidad con los niveles de entendimiento de los monólogos emitidos.

Conclusiones

La observación de una imagen biestable supone procesamientos de tipo ascendente y descendente. La estimulación auditiva, emitida de manera multimodal junto a un estímulo visual biestable, puede proveer contextos semánticos congruentes con cada posible interpretación de dicho estímulo. El uso de tonos de voces como

moduladores semánticos puede tener la posibilidad de ejercer un efecto en la decodificación de imágenes biestables, más allá del entendimiento de su contenido. Sin embargo, tendrán que sistematizarse y tipificarse los modos en que moduladores acústicos de este tipo operan en el contexto de la percepción biestable. Sobre la base de las ideas anteriormente referidas, puede presumirse que cuando existe una interferencia acústica (modulación) durante la observación de un estímulo biestable, si se está observando dicho estímulo por ciertas áreas en específico se podrá interpretar el percepto con carga semántica congruente, esto si las áreas observadas implican un favorecimiento del percepto en específico. Así, el posible vínculo multimodal proveniente de diversas señales sensoriales puede estar condicionado por factores de tipo físico y sensorial, como las áreas de observación, más los trazos o características que en específico orientan el procesamiento visual de la imagen. Investigación de tipo experimental tendrán que seguirse ejecutando para desentrañar el papel particular que cumplen tanto las fijaciones oculares como la información contextual que opera como modulador de tipo semántico durante la percepción de los estímulos visuales biestables.

Referencias

- Balcetis, E., & Dale, R. (2007). Conceptual set as top-down constraint on visual object identification. *Perception*, *36*(4), 581-595. <https://doi.org/10.1068/p5678>
- Barrera, M., & Calderón, L. (2013). Notes for supporting an epistemological neuropsychology: contributions from three perspectives. *International Journal of Psychological Research*, *6*(2), 107-118. Recuperado de http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S2011-20842013000200012
- Boring, E. (1930). A new ambiguous figure. *The American Journal of Psychology*, *42*(3), 444-445. <https://doi.org/10.2307/1415447>
- Borisyuk, R., Chik, D., & Kazanovich, Y. (2009). Visual perception of ambiguous figures: synchronization based neural models. *Biological Cybernetics*, *100*, 491-504. <https://doi.org/10.1007/s00422-009-0301-1>
- Brascamp, J., Sterzer, P., Blake, R., & Knapen, T. (2018). Multistable perception and the role of the frontoparietal cortex in perceptual inference. *Annual review of psychology*, *69*, 77-103. <https://doi.org/10.1146/annurev-psych-010417-085944>
- Brouwer, G. J., & van Ee, R. (2006). Endogenous influences on perceptual bistability depend on exogenous stimulus characteristics. *Vision Research*, *46*(20), 3393-3402. <https://doi.org/10.1016/j.visres.2006.03.016>
- Chastain, G., & Burnham, C. A. (1975). The first glimpse determines the perception of an ambiguous figure. *Perception & Psychophysics*, *17*(3), 221-224. <https://doi.org/10.3758/BF03203203>
- Chen, Y. C., & Spence, C. (2011). Crossmodal semantic priming by naturalistic sounds and spoken words enhances visual sensitivity. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1554-1568. <https://doi.org/10.1037/a0024329>
- Clément, G., & Demel, M. (2012). Perceptual reversal of bi-stable figures in microgravity and hypergravity during parabolic flight. *Neuroscience letters*, *507*(2), 143-146. <https://doi.org/10.1016/j.neulet.2011.12.006>
- Clément, G., & Eckardt, J. (2005). Influence of the gravitational vertical on geometric visual illusions. *Acta Astronautica*, *56*, 911-917. <https://doi.org/10.1016/j.actaastro.2005.01.017>
- Davidson, G. D., & Pitts, M. A. (2014). Auditory event-related potentials associated with perceptual reversals of bistable pitch motion. *Frontiers in Human Neuroscience*, *8*(572). <https://doi.org/10.3389/fnhum.2014.00572>

- Denham, S., Bendixen, A., Mill, R., Tóth, D., Wennekers, T., Coath, M., ... & Winkler, I. (2012). Characterising switching behaviour in perceptual multi-stability. *Journal of Neuroscience Methods*, 210(1), 79-92. <https://doi.org/10.1016/j.jneumeth.2012.04.004>
- Gale, A., & Findlay, J. (1983). Eye-movement patterns in viewing ambiguous figures. In R. Groner, C. Menz, D. Fisher, & R. Monty (Eds.), *Eye movements and psychological functions: international views* (pp. 145-168). Hillsdale NJ: LEA.
- García-Pérez, M. (1989). Visual inhomogeneity and eye movements in multistable perception. *Perception & Psychophysics*, 46(4), 397-400. <https://link.springer.com/content/pdf/10.3758/BF03204995.pdf>
- García-Pérez, M. A. (1992). Eye movements and perceptual multistability. *Advances in Psychology* (88), 73-109. [https://doi.org/10.1016/S0166-4115\(08\)61743-4](https://doi.org/10.1016/S0166-4115(08)61743-4)
- Goolkasian, P., & Woodberry, C. (2010). Priming effects with ambiguous figures. *Attention, Perception & Psychophysics*, 72, 168-178. <https://link.springer.com/article/10.3758/APP.72.1.168>
- Grossmann, J. K., & Dobbins, A. C. (2006). Competition in bistable vision is attribute-specific. *Vision Research*, 46(3), 285-292. <https://doi.org/10.1016/j.visres.2005.06.002>
- Gulan, T., & Valerjev, P. (2010). Semantic and related types of priming as a context in word recognition. *Review of Psychology*, 17(1), 53-58. <https://hrcak.srce.hr/70661>
- Hsiao, J. Y., Chen, Y. C., Spence, C., & Yeh, S. L. (2010). Semantic congruency, attention, and fixation position modulate conscious perception when viewing a bistable figure. *Journal of Vision*, 10(7), 867. <https://doi.org/10.1167/10.7.867>
- Hsiao, J., Chen, Y., Spence, C., & Yeh, S. (2012). Assessing the effects of audiovisual semantic congruency on the perception of a bistable figure. *Consciousness and Cognition*, 21(2), 775-787. <https://doi.org/10.1016/j.concog.2012.02.001>
- Intaité, M., Koivisto, M., Rukšėnas, O., & Revonsuo, A. (2010). Reversal negativity and bistable stimuli: attention, awareness, or something else? *Brain and Cognition*, 74(1), 24-34. <https://doi.org/10.1016/j.bandc.2010.06.002>
- Intaité, M., Koivisto, M., & Castelo-Branco, M. (2014). Event-related potential responses to perceptual reversals are modulated by working memory load. *Neuropsychologia*, 56, 428-438. <https://doi.org/10.1016/j.neuropsychologia.2014.02.016>
- Intaité, M., Noreika, V., Šoliūnas, A., & Falter, C. M. (2013). Interaction of bottom-up and top-down processes in the perception of ambiguous figures. *Vision Research*, 89, 24-31. <https://doi.org/10.1016/j.visres.2013.06.011>

- Kogo, N., Hermans, L., Stuer, D., van Ee, R., & Wagemans, J. (2015). Temporal dynamics of different cases of bi-stable figure-ground perception. *Vision Research*, *106*, 7-19. <https://doi.org/10.1016/j.visres.2014.10.029>
- Laubrock, J., Engbert, R., & Kliegl, R. (2008). Fixational eye movements predict the perceived direction of ambiguous apparent motion. *Journal of Vision*, *8*(14), 13. <https://doi.org/10.1167/8.14.13>
- Kornmeier, J., & Bach, M. (2005). The Necker cube - an ambiguous figure disambiguated in early visual processing. *Vision Research*, *45*(8), 955-960. <https://doi.org/10.1016/j.visres.2004.10.006>
- Kornmeier, J., Hein, C. M., & Bach, M. (2009). Multistable perception: when bottom-up and top-down coincide. *Brain and Cognition*, *69*(1), 138-147. <https://doi.org/10.1016/j.bandc.2008.06.005>
- Leopold, D. A., & Logothetis, N. K. (1999). Multistable phenomena: changing views in perception. *Trends in Cognitive Sciences*, *3*(7), 254-264. [https://doi.org/10.1016/S1364-6613\(99\)01332-7](https://doi.org/10.1016/S1364-6613(99)01332-7)
- Long, G. M., & Toppino, T. C. (2004). Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychological Bulletin*, *130*(5), 748-768. <https://doi.org/10.1037/0033-2909.130.5.748>
- Meng, M., & Tong, F. (2004). Can attention selectively bias bistable perception? Differences between binocular rivalry and ambiguous figures. *Journal of Vision*, *4*(7), 539-551. <https://doi.org/10.1167/4.7.2>
- Moreno-Bote, R., Rinzel, J., & Rubin, N. (2007). Noise-induced alternations in an attractor network model of perceptual bistability. *Journal of Neurophysiology*, *98*(3), 1125-1139. <https://doi.org/10.1152/jn.00116.2007>
- Okazaki, M., Kaneko, Y., Yumoto, M., & Arima, K. (2008). Perceptual change in response to a bistable picture increases neuromagnetic beta-band activities. *Neuroscience Research*, *61*(3), 319-328. <https://doi.org/10.1016/j.neures.2008.03.010>
- Patiño, V. (2007). La memoria semántica: modelos neuropsicológicos y alteraciones a causa del daño cerebral. *Neuropsicología*, *2*(1), 10-17. <http://www.imbiomed.com/1/PDF/Ns071-02.pdf>
- Pressnitzer, D., & Hupé, J.-M. (2006). Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization. *Current Biology*, *16*(13), 1351-1357. <https://doi.org/10.1016/j.cub.2006.05.054>

- Rodríguez, G. (2016). La reconfiguración perceptual de imágenes aplicada al desarrollo del pensamiento divergente en el aula de clase. *Revista Q*, 11(21), 61-81. <https://expeditiorepositorio.utadeo.edu.co/handle/20.500.12010/8737>
- Rodríguez, G., & Castillo, H. (2018a). Bistable perception: neural bases and usefulness in psychological research. *International Journal of Psychological Research*, 11(2), 63-76. <https://doi.org/10.21500/20112084.3375>
- Rodríguez, G., & Castillo, H. (2018b). Tareas de búsqueda visual: modelos, bases neurológicas, utilidad y prospectiva. *Universitas Psychologica*, 17(1), 1-12. <https://doi.org/10.11144/Javeriana.upsy17-1.tbvm>
- Runnova, A. E., Hramov, A. E., Grubov, V. V., Koronovskii, A. A., Kurovskaya, M. K., & Pisarchik, A. N. (2016). Theoretical background and experimental measurements of human brain noise intensity in perception of ambiguous images. *Chaos, Solitons & Fractals*, 93, 201-206. <https://doi.org/10.1016/j.chaos.2016.11.001>
- Schauer, G., Kanai, R., & Brascamp, J. W. (2016). Parietal theta burst TMS: Functional fractionation observed during bistable perception not evident in attention tasks. *Consciousness and Cognition* 40, 105-115. <https://doi.org/10.1016/j.concog.2016.01.002>
- Smith, E. L., Grabowecky, M., & Suzuki, S. (2007). Auditory-visual crossmodal integration in perception of face gender. *Current Biology*, 17(19), 1680-1685. <https://doi.org/10.1016/j.cub.2007.08.043>
- Sterzer, P., Kleinschmidt, A., & Rees, G. (2009). The neural bases of multistable perception. *Trends in Cognitive Sciences*, 13(7), 310-318. <https://doi.org/10.1016/j.tics.2009.04.006>
- Sterzer, P., & Rees, G. (2009). Bistable Perception and Consciousness. *Encyclopedia of Consciousness* (pp. 93-106). <https://doi.org/10.1016/B978-012373873-8.00011-6>
- Sterzer, P., Russ, M. O., Preibisch, C., & Kleinschmidt, A. (2002). Neural correlates of spontaneous direction reversals in ambiguous apparent visual motion. *NeuroImage*, 15(4), 908-916. <https://doi.org/10.1006/nimg.2001.1030>
- Van Loon, A. M., Knapen, T., Scholte, H. S., St. John-Saaltink, E., Donner, T. H., & Lamme, V. A. (2013). GABA Shapes the Dynamics of Bistable Perception. *Current Biology*, 23(9), 823-827. <https://doi.org/10.1016/j.cub.2013.03.067>
- Vatakis, A., & Spence, C. (2007). Crossmodal binding: Evaluating the “unity assumption” using audiovisual speech stimuli. *Perception & Psychophysics*, 69(5), 744-756. <https://doi.org/10.3758/BF03193776>

- Weilhammer, V., Stuke, H., Hesselmann, G., Sterzer, P., & Schmack, K. (2017). A predictive coding account of bistable perception - a model-based fMRI study. *PLoS Computational Biology*, 13(5), 1-21 <https://doi.org/10.1371/journal.pcbi.1005536>
- Xiaogang, W., Sang, N., Hao, L., Zhang, Y., Bi, T., & Qiu, J. (2017). Category Selectivity of Human Visual Cortex in Perception of Rubin Face-Vase Illusion. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.01543>
- Yamamoto, S., & Yamamoto, M. (2006). Effects of the gravitational vertical on the visual perception of reversible figures. *Neuroscience Research*, 55(2), 218-221. <https://doi.org/10.1016/j.neures.2006.02.014>