

# From dots to faces: individual differences in visual imagery capacity predict the content of Ganzflicker-induced hallucinations

Ana Chkhaidze <sup>1</sup>, Reshane R. Reeder <sup>2</sup>, Connor Gag<sup>3</sup>, Anastasia Kiyonaga <sup>1</sup>, Seana Coulson <sup>1,4,\*</sup>

<sup>1</sup>Department of Cognitive Science, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, United States

<sup>2</sup>Department of Psychology, Institute of Population Health, University of Liverpool, Bedford Street South, Liverpool L69 7ZA, United Kingdom

<sup>3</sup>Department of Computer Science and Engineering, University of California San Diego, 3235 Voigt Drive, La Jolla, CA 92093, United States

<sup>4</sup>Kavli Institute for Brain and Mind, University of California San Diego, 2985 Muir Lane, La Jolla, CA 92093, United States

\*Corresponding author. Department of Cognitive Science, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093, United States. E-mail: [scoulson@ucsd.edu](mailto:scoulson@ucsd.edu)

## Abstract

A rapidly alternating red and black display known as Ganzflicker induces visual hallucinations that reflect the generative capacity of the visual system. Individuals vary in their degree of visual imagery, ranging from absent to vivid imagery. Recent proposals suggest that differences in the visual system along this imagery spectrum should also influence the complexity of other internally generated visual experiences. Here, we used tools from natural language processing to analyze free-text descriptions of hallucinations from >4000 participants, asking whether people with different imagery phenotypes see different things in their mind's eye during Ganzflicker-induced hallucinations. Topic modeling of descriptions revealed that strong imagers described complex, naturalistic content, while weak imagers reported simple geometric patterns. Using crowd-sourced sensorimotor norms, we also found that participants with stronger imagery used language with richer perceptual associations. These findings may reflect individual variation in coordination between early visual areas and higher-order regions relevant for the imagery spectrum.

**Keywords** aphantasia, individual differences, visual cortex, visual hallucinations, visual imagery

## Introduction

In 1819, the physiologist Jan Purkinje made a striking observation: While waving his hand between his closed eyes and sunlight, he perceived “beautiful regular figures that are initially difficult to define but slowly become clearer” (Purkinje 1819). His drawings of these experiences—geometric lattices, radiating patterns, and spiraling forms—represent the first scientific documentation of what we now call “flicker-induced visual hallucinations”: visual percepts that arise not from external stimuli but from the brain's response to rhythmic light stimulation without pharmacological intervention (Hewitt et al. 2025). Flicker stimulation has become a widely used experimental method for non-invasively and non-pharmacologically inducing “altered states of consciousness” (ASCs): transient shifts in subjective experience that differ qualitatively from ordinary waking awareness (Fort et al. 2025). From Purkinje's early introspections to modern laboratory studies (Fig. 1), the experienced visuals in flicker paradigms have been so consistent that they have been codified as “form

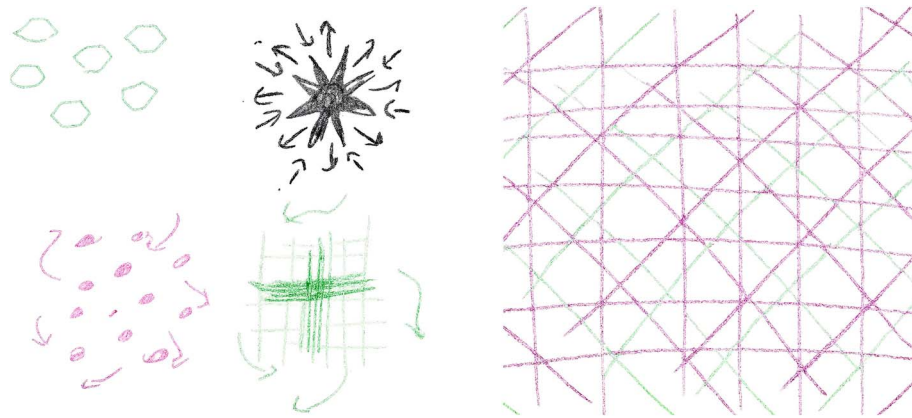
constants”, a universal vocabulary for simple visual hallucinations (Rule et al. 2011). This consistency suggests that rather than transient perceptual noise, flicker-induced hallucinations reflect the intrinsic architecture of the human visual system (Bressloff et al. 2001, 2002). As such, they offer a unique experimental window into other forms of internally generated visual experiences, including visual mental imagery, a fundamental yet poorly understood cognitive trait.

Mental imagery, the ability to internally simulate sensory experiences, is one of the most familiar phenomenological aspects of the human experience. Argued to play a central role in core cognitive and affective functions (Pearson 2019), visual mental imagery has been shown to relate to working memory (Keogh and Pearson 2014), mind wandering (Christian et al. 2013), navigation (Bird et al. 2012), and language comprehension (Muraki et al. 2023). However, visual mental imagery capacity varies dramatically across individuals along what we term the “visual imagery spectrum” (referred to below simply as the imagery spectrum), ranging from aphantasia (complete absence of visual imagery) to hyperphantasia (strong,

Received 17 July 2025. revised 9 March 2026. accepted 16 April 2026

© The Author(s) 2026. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).



**Figure 1.** Ganzflicker-induced hallucination drawings from a pilot participant in our lab. Verbal report: “Initially sort of a grid-like pattern, but with certain parts being more emphasized than others, and moving around across the screen. Then at some point sort of an array of pulsating dots, also moving around the screen. Then something like starbursts. Occasionally the grid-like pattern would return, and during some periods there was just a sort of non-descript dynamic patterns moving around quickly. At one point more toward the end, I was seeing something similar to chemical molecular structures. All of these were in constant motion and fluctuation.”

lifelike mental pictures). Such diversity raises fundamental questions about cognitive universality: if people differ in how they internally represent sensory information, then the underlying mechanisms of the cognitive phenomena that draw on these representations may also differ systematically. Indeed, recent work demonstrates that imagery differences can predict memory recall (Bainbridge et al. 2021), emotional reactivity (Wicken et al. 2021), and working memory strategies (Reeder et al. 2024), suggesting that understanding individual differences across the imagery spectrum may be crucial for developing more inclusive theories of human cognition (Lupyan et al. 2023). While research has extensively documented that imagery vividness varies across individuals (Zeman 2024), we know remarkably little about what people with different imagery phenotypes “see”—i.e. the content and structure of these internally generated visual experiences. Here we used tools from natural language processing to analyze participants’ descriptions of their visual hallucinations, asking whether people across the imagery spectrum see different things in their mind’s eye.

### The layered model of visual imagery

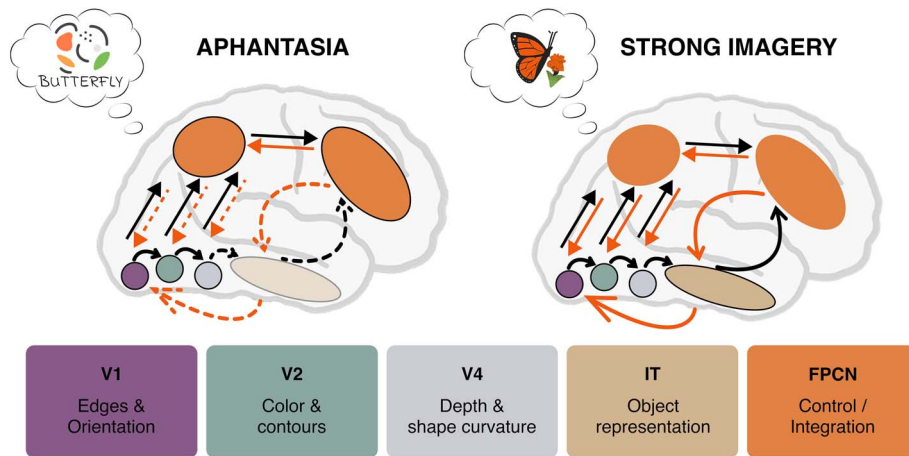
Historically, imagery research has predominantly focused on scalar measures of vividness (e.g. the Vividness of Visual Imagery Questionnaire; Marks 1973) rather than the compositional elements that constitute mental imagery (e.g. imagery manipulation and precision; Kosslyn et al. 1984). Initial neuroimaging studies of the imagery spectrum focused on early visual cortex, revealing that imagery vividness and precision correlate with the pattern and extent of activation in V1 during imagery tasks (Cui et al. 2007, Lee et al. 2012, Albers et al. 2013, Bergmann et al. 2016). However, more recent work has extended this investigation to the entire brain, confirming differences in early visual cortex when decoding imagined content (Cabbai et al. 2024), and revealing activation differences in insula (Cabbai et al. 2024, Silvanto and Nagai 2025), frontoparietal control networks (Spagna et al. 2021), and the fusiform gyrus (Liu et al. 2025), reflecting the distributed, multi-level nature of mental imagery processes.

Importantly, this distributed activity may translate into individual variation not just in imagery vividness, but in the content and structural complexity of what people can internally represent. While

individuals across the imagery spectrum may similarly represent low-level visual features (lines, edges, spatial patterns), those with weaker imagery or aphantasia may have reduced capacity for integrating these elements into complex, semantically meaningful images due to differences in top-down connectivity between control networks and visual areas during imagery (Milton et al. 2021, Liu et al. 2025). For example, when imagining a butterfly landing on a flower, early visual cortex might encode basic features like curved edges, color patches, and black lines across all individuals, but the successful construction of a vivid butterfly image may require higher-order semantic and object processing areas to integrate these features into meaningful shapes, along with the recruitment of color-selective areas to contribute the rich hues in the scene, as well as coordination via frontoparietal control networks to bind these elements into a unified percept (Mechelli et al. 2004, Dijkstra et al. 2019, Spagna et al. 2021). This coordination may be compromised in weaker imagers, resulting in fragmented or impoverished butterfly imagery despite intact basic feature representations. We refer to this framework as the “layered model of visual imagery” in which individuals across the imagery spectrum similarly represent low-level visual features in the early visual cortex, but differ in their capacity to integrate these elements into complex, semantically meaningful representations (for details see Fig. 2).

### Individual differences in imagery and the induction of visual hallucinations

If individual differences in visual imagery reflect variation in the hierarchical coordination of visual representations, then these differences should manifest not only during voluntary imagery but also in other situations where internally generated content shapes perception, such as during flicker-induced hallucinations (Reeder et al. 2024). Since Purkinje’s pioneering forays, experimental methods to induce visual hallucinations have evolved, systematized, and diversified. Visual hallucinations can be successfully induced through both sensory deprivation (e.g. blindfolding; Merabet et al. 2004) and sensory stimulation (e.g. stroboscopic light stimulation; Hewitt et al. 2025). A combination of both, the “Ganzfeld” paradigm involves presenting participants with a uniform colored field, inducing hallucinations as the brain



**Figure 2.** Layered model of visual imagery. Proposed differences in connectivity across the visual hierarchy during internally generated visual experiences between aphantasia and strong imagery. Black arrows indicate bottom-up feedforward connections; orange arrows indicate top-down modulatory connections. Solid lines represent strong connectivity; dashed lines represent reduced connectivity. Both individuals with aphantasia and with moderate to strong imagery show intact bottom-up flow from early visual cortex (EVC: V1, V2, V4) through parietal to frontal regions, explaining why people with no self-reported imagery still experience simple visuals during Ganzflicker-induced hallucinations (Königsmark et al. 2021, Reeder 2022). The key difference lies in the efficiency of coordinated activity across the hierarchy. In strong imagers, a combination of strong bidirectional connectivity between prefrontal areas and inferotemporal (IT) cortex and robust top-down modulation of visual cortical activity via the fronto-parietal control network (FPCN) enable integration of low-level features into coherent, semantically meaningful imagery. In aphantasia, these connections are weaker, limiting conscious imagery to simple content.

generates structure in the absence of patterned input (Metzger 1929, Pütz et al. 2006, Wackermann et al. 2008, Schmidt and Prein 2019, Schmidt et al. 2020). Stroboscopic light stimulation by contrast, delivers bright rhythmic stimulation, typically white light to closed eyes, producing hallucinations possibly through neural entrainment in the visual cortex (as reviewed in Hewitt et al. 2025). The “Ganzflicker”—a special case of the flicker-based paradigm—combines elements of both approaches: like stroboscopic light stimulation, it uses rhythmic flicker to drive visual experience; like Ganzfeld, it employs uniform chromatic stimulation, though unlike previous suggestions, this does not have to fill the entire visual field to elicit hallucinations (Reeder 2022). Specifically, the term refers to open-eye viewing of two uniform chromatic fields (typically red and black) alternating rhythmically (Königsmark et al. 2021). All these induction methods can successfully lead to ASCs (Fort et al. 2025), though the phenomenological character of the resulting experiences may vary (Shenyan et al. 2024, Fort et al. 2025).

Simple hallucinations—geometric shapes, color alterations, and elementary shapes—are the most commonly reported experiences in these light-induced hallucinations, across both Ganzfeld and flicker-based paradigms (Pütz et al. 2006, Wackermann et al. 2008, Allefeld et al. 2011, Rule et al. 2011, Schmidt and Prein 2019, Schmidt et al. 2020, Bartossek et al. 2021, Hewitt et al. 2025, Pistolas and Wagemans 2025). These experiences have been suggested to reflect the intrinsic dynamics of early visual cortex as computational models demonstrate that conditions of either rhythmic stimulation or reduced sensory input result in these patterned representations due to lateral connectivity and retinotopic organization in V1 (Ermentrout and Cowan 1979, Bressloff et al. 2001, 2002).

Despite the predominance of simple hallucinations, complex hallucinatory content has also been reported in both Ganzfeld (Pütz et al. 2006, Shenyan et al. 2024) and flicker-induced paradigms (Allefeld et al. 2011, Schwartzman et al. 2019, Königsmark et al. 2021, Reeder 2022, Shenyan et al. 2024) (Fig. 3). Complex hallucinations involve recognizable semantic content, like faces, human figures, animals, objects, landscapes, and structured scenes. Notably, the likelihood of

experiencing complex versus simple hallucinations appears to depend in part on the induction method. Shenyan et al. (2024) found that complex hallucinations were relatively more likely during Ganzfeld than Ganzflicker, suggesting that bottom-up rhythmic stimulation of the early visual cortex may not be sufficient to generate complex content. Instead, complex hallucinations may require additional involvement of higher-level visual areas—either through feedforward propagation up the visual hierarchy or through top-down interpretative processes (see e.g. Shenyan et al. 2024). Consistent with a role for top-down processing, Pistolas and Wagemans (2025) documented complex imagery (fish, sharks, ocean scenes) that appeared to be shaped by contextual cues such as illumination color and ambient sound. However, only a portion (~20%) of participants were susceptible to such context-driven influences. Complex hallucinations also occur in flicker paradigms: some participants undergoing stroboscopic light stimulation have reported semantically rich scenes such as wheat fields, faces of loved ones, and birds flying into space (Beauté et al. 2025). Thus, hallucination complexity appears to depend on more than just the induction method and may also vary as a function of stable individual traits such as visual imagery capacity (Beauté et al. 2025).

Indeed, in recent years visual imagery capacity has emerged as a reliable predictor of flicker-induced hallucinations. Large-scale studies using the Ganzflicker paradigm show that individuals with stronger imagery report more frequent and more vivid hallucinations during Ganzflicker stimulation (Königsmark et al. 2021, Reeder 2022). If complex hallucinations depend on layered mechanisms similar to those supporting voluntary visual imagery, then individuals with stronger imagery capacity should also be more likely to experience complex content. When asked to categorize their hallucinations using fixed response options, participants with strong imagery were more likely to report complex hallucinations than those with weak imagery (Reeder 2022).

While these studies established the first empirical link between visual imagery capacity and Ganzflicker-induced hallucinations, closed-ended response formats cannot fully reveal the experiential



**Figure 3.** Ganzflicker-induced simple and complex hallucinations. Drawings by an artist based on her firsthand Ganzflicker experiences, illustrating both simple and complex visual hallucinations, including colored geometric patterns, a tunnel, and human figures. Reproduced with permission; drawings © Kirsten Baron.

content of hallucinations. The layered model of imagery offers a testable prediction about hallucination content across the imagery spectrum: individuals with stronger imagery should report more complex, naturalistic hallucination content, while those with weaker imagery should predominantly report the simple geometric patterns linked to activity in the early visual cortex. In other words, the same processing differences that distinguish imagery phenotypes should be visible in the content of what people “see” during Ganzflicker—a prediction that can be tested by analyzing the semantic structure of open-ended phenomenological reports.

### The present study

To characterize variability in hallucination content across the imagery spectrum, we conducted a secondary analysis of an open dataset containing >6000 participants’ free-text descriptions of their experiences following Ganzflicker stimulation (Reeder 2022). After the Ganzflicker exposure, participants provided self-reported visual imagery vividness ratings to determine their position on the imagery spectrum, and answered structured questions about their experiences during Ganzflicker (Fig. 4a). While Reeder (2022) focused primarily on these structured questionnaire data, participants also provided rich narrative reports describing their visual hallucinations (see Supplementary Material 1 for example descriptions), a corpus that has not previously been analyzed for content differences across the imagery spectrum. Analyzing the content and structure of these free-text descriptions requires a methodological approach that can accommodate the variability and richness of open-ended phenomenological reports while enabling systematic, replicable analysis. To address this challenge, we employed two complementary methods (Fig. 4c).

First, we aimed to uncover the phenomenological content of Ganzflicker-induced hallucinations, asking whether it changes across the imagery spectrum. To locate latent themes in participants’ descriptions, we used Mapping of Subjective Accounts into Interpreted Clusters, a topic modeling pipeline that relies on machine learning to identify coherent themes in text corpora, allowing data-driven discovery of hallucination content (Beauté et al. 2025). Following prior literature (Reeder 2022, Hewitt et al. 2025), we expected a

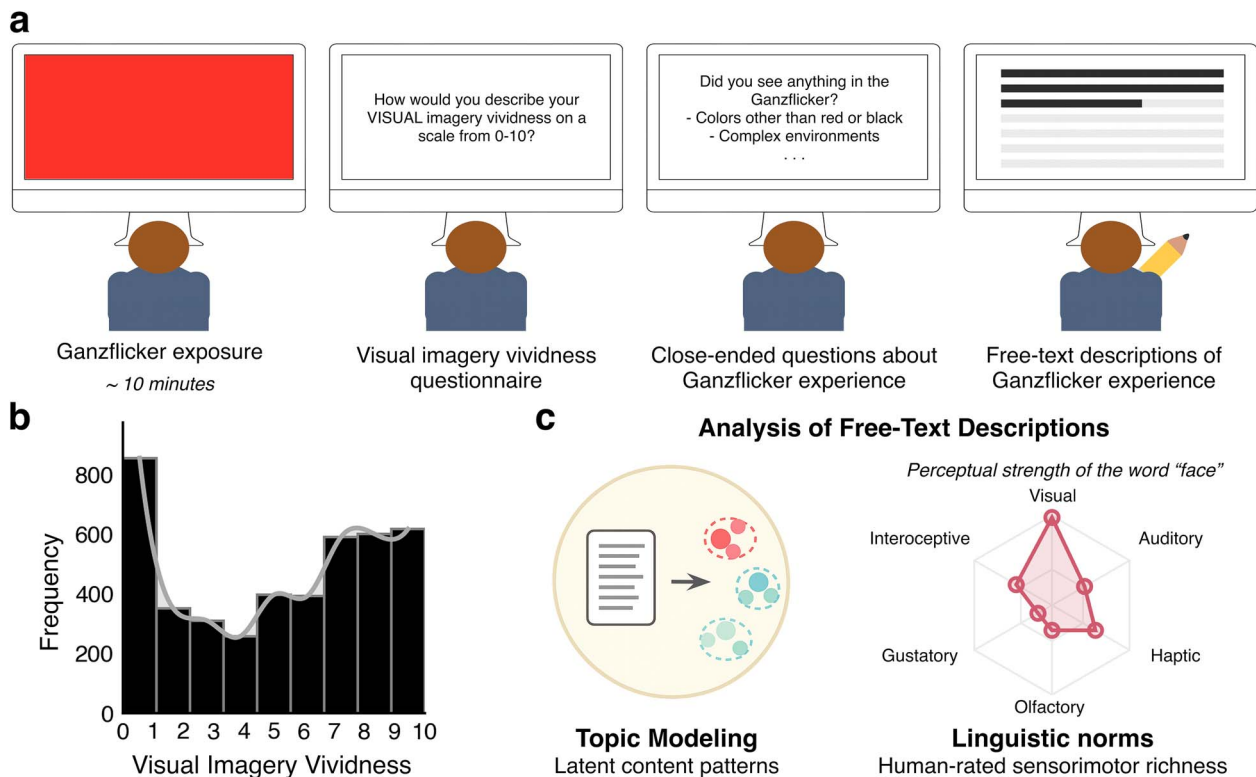
range of simple and complex themes to emerge. We hypothesized that participants with stronger imagery would describe more complex, naturalistic content (faces, scenes, structured hallucinations), while those with weaker imagery would report simpler, geometric patterns, arising from the individual differences in brain organization posited in the layered model of imagery.

Second, we used the Lancaster Sensorimotor Norms (Lynott et al. 2020) to assess explicit sensorimotor language in participants’ descriptions. These norms provide crowd-sourced ratings of how strongly words evoke different sensory and motor experiences—for example, rating “thunder” as highly auditory, “velvet” as highly tactile, “kick” as foot-related, or “clap” as hand-related. While topic modeling may reveal thematic content at the level of whole descriptions, the Lancaster norms complement this analysis by indexing explicit sensorimotor vocabulary at the word level. Moreover, as they are derived from human ratings, they offer a psychologically grounded measure with established links to behavioral indices of word processing, including word naming and the lexical decision task (Connell and Lynott 2012), as well as neural responses to words in the property verification task (Vinaya et al. 2025). Crucially, these norms allowed us to identify which sensory and motor dimensions emerge in Ganzflicker-induced hallucinations and test whether they vary systematically across the imagery spectrum. We expected participants with higher imagery vividness scores to use language with stronger sensorimotor associations, reflecting richer internal simulation.

## Materials and Methods

### Participants

We used a dataset from a Ganzflicker study (Reeder 2022) with 6664 individuals who underwent 10 minutes of continuous Ganzflicker stimulation and reported their experiences, alongside visual imagery vividness scores. Participants were recruited via a popular science article published in *The Conversation* describing the Ganzflicker phenomenon (Reeder 2021), which was republished by at least 14 media outlets including *Science Alert*, *The Daily Beast*, and *Big Think*, all of which contained a link to the Ganzflicker study hosted on Google Forms. This widespread media coverage led to a large, international



**Figure 4.** Ganzflicker task and analytical approach. (a) Experimental procedure from Reeder (2022). Participants viewed red-black flicker for ~10 minutes, completed a demographics questionnaire and rated their visual imagery vividness (0–10); afterwards, they answered whether they saw anything during Ganzflicker and provided free-text descriptions of what they saw, followed by additional closed-ended questions about their experience. (b) Distribution of visual imagery vividness ratings across participants in this study. Bars show participant counts at each discrete vividness rating; the curve shows the overall shape of the distribution. (c) Analysis pipeline: topic modeling to identify semantic themes and human-rated linguistic norms to quantify sensorimotor content in hallucinations.

volunteer sample, with the majority of readers based in the USA (73%), followed by the UK (5.2%), Canada (4.1%), and Australia (3.3%). A total of 68.1% participants reported their gender as male, 27.6% as female, 2.4% as other, and 1.9% preferred not to say. Five thousand seven hundred seven individuals reported their age ( $M=38.07$ ,  $SD=15.08$ ). The sample also included a higher-than-typical proportion of individuals at the extremes of the imagery spectrum, including self-reported aphantasia (Fig. 4b). We excluded participants whose descriptions were missing or who provided written descriptions in a language other than English, leaving 4365 participants. As reported by Reeder (2022), the study was approved by the ethics committee of the medical faculty of Otto-von-Guericke University and all participants gave informed consent.

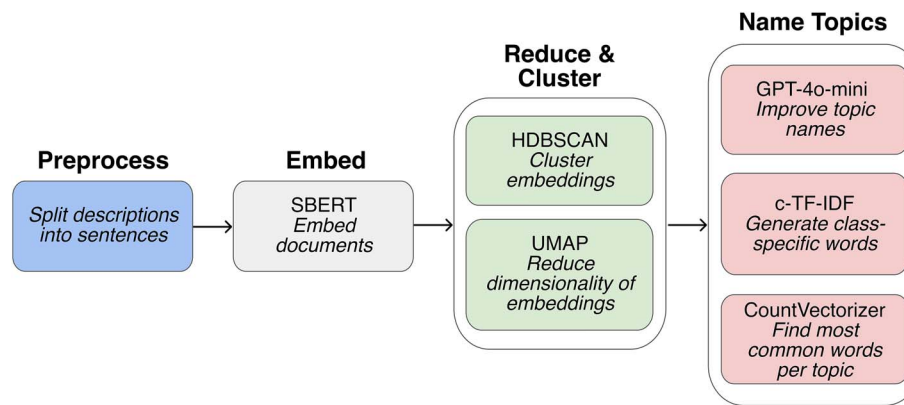
### Procedure

Participants viewed a Ganzflicker stimulus that alternated between red and black at 7.5 Hz for ~10 minutes. They were instructed to watch Ganzflicker on full-screen and under dim lighting, accompanied by white noise to homogenize the acoustic environment, and to avoid distractions. Approximately 73% of participants reported compliance with the white noise recommendation; however, the proportion of complex versus simple hallucinations did not differ between those who used white noise and those who did not (Königsmark et al. 2021, Reeder 2022), suggesting that auditory masking did not substantially influence hallucination content. Further details on environment control instructions are reported in Reeder (2022).

After the stimulus presentation, participants completed a demographic survey. To situate them on the imagery spectrum, they rated their visual imagery vividness on a single item [How would you describe your VISUAL imagery vividness on a scale from 0 (no mental imagery) to 10 (as vivid as real perception)]. Participants also answered closed-ended questions assessing their Ganzflicker experiences, and an open-ended free-text description of what they saw during exposure. In this paper, we focus on these free-text descriptions.

### Topic modeling

To uncover latent themes in participants' descriptions of Ganzflicker-induced hallucinations, we implemented a multi-step topic modeling pipeline based on BERTopic (Grootendorst 2022), an unsupervised method that integrates transformer-based embeddings, dimensionality reduction, density-based clustering, and class-based term weighting (Fig. 5). After clustering, each topic's most informative terms were extracted using class-based Term Frequency-Inverse Document Frequency (c-TF-IDF), which identifies words that best characterize a given cluster relative to the rest of the corpus. These keyword sets were used as the primary basis for human-interpretable topic labels. To further reduce researcher expectancy bias, we also employed GPT-4o-mini to generate independent topic labels based on the same representative keywords and exemplar sentences. Human- and model-generated labels showed close alignment (see Table 1). A more explicit description of the pipeline is detailed in Fig. 5.



**Figure 5.** Topic modeling pipeline. Topic modeling pipeline: Free-text hallucination descriptions were split into sentences, embedded using Sentence-BERT, reduced with UMAP, and clustered using HDBSCAN. Topics were labeled using a combination of most-frequent terms, c-TF-IDF, and GPT-4o-mini summarization.

### Sentence-level tokenization and preprocessing

Each hallucination report was first split into sentences using `sent_tokenize()` from the Natural Language Toolkit (NLTK) (Bird et al. 2009), which applies the “Punkt” algorithm (Kiss and Strunk 2006)—an unsupervised model that uses punctuation and contextual cues to identify sentence boundaries. This step allowed us to capture distinct experiential elements that often co-occur within a single report. Treating entire descriptions as a single document would have conflated diverse perceptual content under one topic. The resulting sentence-level documents were lightly cleaned—converted to lowercase and stripped of excess whitespace—but no stemming, lemmatization, or stopword removal was performed. This minimal preprocessing ensured the preservation of contextual information important for semantic modeling, resulting in a corpus of >10 000 sentence-level documents.

### Sentence embeddings and dimensionality reduction

We generated semantic embeddings using Sentence-BERT (Reimers and Gurevych 2019), which produces 384-dimensional vector representations optimized for capturing semantic similarity in short texts. This embedding process preserves subtle differences in meaning and context, allowing the model to distinguish between similar phrases with different phenomenological implications (e.g. “I felt time slow down” versus “I lost track of time”).

To prepare the embeddings for clustering, we applied uniform manifold approximation and projection (UMAP) to reduce their dimensionality (McInnes et al. 2018). UMAP preserves both local and global structure in the data, helping maintain the semantic relationships between sentences in the lower-dimensional space. We configured the reduced dimensionality (`n_components=10`), the neighborhood size used to balance local versus global structure (`n_neighbors=15`), and the minimum distance between embedded points (`min_dist=0.1`) so as to yield an embedding structure that supported coherent and interpretable topic clustering.

### Clustering with hierarchical density-based spatial clustering of applications with noise

The reduced sentence embeddings were clustered using hierarchical density-based spatial clustering of applications with noise (HDBSCAN)

(Campello et al. 2013, McInnes et al. 2017), a non-parametric algorithm that identifies clusters of varying density without requiring a predefined number of clusters. This feature is particularly advantageous when analyzing naturalistic text, where the number and size of coherent experiential categories are unknown. We set the minimum cluster size to 30, allowing the model to discover a diverse but stable set of themes while treating small or noisy clusters as outliers. Sentences not confidently assigned to any topic were labeled as outliers (`topic = -1`) and excluded from topic-based analyses to ensure the coherence and interpretability of topic clusters. A total of 32 sentences (<0.3% of the corpus) were labeled as outliers; these were short Turkish-language descriptions that had not been flagged during automated language detection and were excluded from topic analyses. The final model identified 28 topics spanning a wide range of visual and conceptual content, ranging from simple geometric shapes to complex scenery (Fig. 6, Table 1).

### Topic extraction and coherence evaluation

To identify the most distinctive terms characterizing each topic, we applied c-TF-IDF (Grootendorst 2022). This weighting scheme highlights words that are frequent within a topic and infrequent across other topics. Unlike standard TF-IDF, which considers individual documents, c-TF-IDF treats all documents within a topic as a single pooled entity, allowing for robust extraction of topic-defining terms. Specifically, our pipeline employed the BM25+ variant of c-TF-IDF, which provides additional smoothing and has been shown to improve performance on sparse or short-text data.

To evaluate the interpretability of topics, we computed the  $C_v$  coherence score (Röder et al. 2015), a widely used metric that assesses the degree of semantic similarity among the top-ranked words within each topic.  $C_v$  combines normalized pointwise mutual information, cosine similarity, and a sliding window approach to assess word co-occurrence. Our model achieved a coherence score of 0.48, which falls within the commonly reported range of 0.4 to 0.7 in topic modeling studies involving short, noisy, or experiential texts (O’Callaghan et al. 2015, Röder et al. 2015). Topic structure was further examined using a 2D topic embedding map (Fig. 6) and a hierarchical dendrogram (Fig. S1), which visually confirmed the separability and cohesion of extracted themes.

**Table 1.** Representative hallucination descriptions by topic.

Human labels	Model labels	Representative descriptions
Experiential descriptions	Trippy experience	"It was like watching a trippy video of colours and sounds"
Hallucination absence	Minimal content	"I didn't see anything substantial though I thought I would"
Noise	White noise	"The white noise and screen together gave me an uneasy feeling but nothing noteworthy just as if this was a horror movie this would be the part where someone dies feeling"
Faces and figures	Human faces	"A potato that turned into an old man's face and then some alien scenery that was pulsing"
Unpleasant images	Subjective experience	"It was not an enjoyable experience but i felt no overwhelming discomfort else i would have stopped the experiment"
Eyes	Eye patterns	"After several minutes i saw a left eye sketch-like which rotated and disappeared then reappeared several times"
Sensation duration	Duration events	"Images only lasted for a few seconds each"
Images	Images movement	"After a few seconds of this I usually would have to blink and the images would be gone"
Gaze	Eye variations	"As soon as I tried to look at them or focus on them they would disappear"
Webs	Spider webs	"patterns often in the form of spider webs"
Dots	Dots Pattern	"Tiny dots of other colours as part of the pattern like smaller than pixels and so many that they almost made up other colours"
Colors	Colors and Patterns	"The black ended up shifting between a greenish black and deep blue"
Movement	Visual distortion	"Mostly just different colors bouncing all around the screen in different shapes"
Flashes	Flashing lights	"Blue pink yellow flashes but mostly a continuous pink flash"
Lines	Lines and patterns	"Thick horizontal and vertical lines with fuzzy edges that intersected at the center of my field of view"
Patterns and shapes	Geometric Patterns	"I saw a bunch of shapes like squares circles lines"
Fractals	Fractals and patterns	"Like a kaleidoscope but much more faint and ever changing"
Spirals and rotation	Spiral rotation	"Everything constantly spinning"
Stars and galaxies	Stars and galaxies	"I saw purple dots that were circling around like stars in a galaxy also with spiral arms etc."
Flowers	Flower imagery	"Perceived i was looking at a field of tulips and daisies and flowers as if through a crumpling red fold of cellophane"
Fire	Fire imagery	"I saw a starburst firework style shape occasionally and a couple of little sparks like a sparkler"
City skyline	Cityscape views	"Reminded me of looking at a bright light with my eyes closed"
Forest and trees	Forest scenery	"On the far edge and wrapping around the shore was a tree line that consisted of birch like trees tall and thinner with plenty of canopy but were not birches and some juniper trees it was accompanied with average pond foliage like cattails and reeds and such"
Water	Ocean reflections	"It wasn't realistically colored or anything just felt like I was standing on the edge of a lake looking at something in the middle of it"
Space	Space travel	"Flying through a space station"
Tunnels	Tunnel experience	"A circle would appear then it turned into a long tunnel"
Hallways and doors	Hallway passage	"Also a very long dark corridor with light at the end it was moving around on the vertical axis"
Butterflies	Butterfly shape	"First was like a butterfly outline with neon lines"

### Topic labeling

To reduce researcher bias—i.e. the tendency for manual labeling to be influenced by theoretical priors—and to enhance interpretability of the keywords, we used "GPT-4o-mini" (OpenAI 2024) for LLM-assisted topic labeling. As outlined by Beauté et al. (2025), this method offers a scalable and unbiased alternative to manual labeling for large phenomenological corpora. For each topic, we supplied the model with its top 10–15 c-TF-IDF keywords and a few representative sentences, prompting it to generate concise, human-readable topic labels. We also independently reviewed the keywords and representative example sentences for each topic and manually assigned labels. The two approaches produced highly consistent results, differing mainly in surface phrasing (e.g. "city skyline" versus "cityscape views"). Topic modeling figures (Figs 5, 6, S1, S2, and S3) reflect the manually-generated labels; Table 1 includes both human-generated and LLM-generated labels for comparison.

### Participant-level topic features

Each sentence was assigned both a hard topic label and a soft probability distribution over all topics. We used BERTopic's standard hard-label assignment, where the most probable topic for each sentence was determined as the argmax of the topic-probability vector  $p(\text{topic} | \text{sentence})$ . In addition, BERTopic provides the full topic-probability distribution for each sentence, indicating the likelihood that it belongs to each topic. To construct participant-level features, we aggregated these probabilities by identifying, for each participant  $\times$  topic pair, the maximum probability that the model assigned to that topic across all of the participant's sentences. This maximum-based aggregation preserves sensitivity to distinct, high-confidence topic expressions and prevents unrelated sentences from diluting the signal. The resulting participant-by-topic matrix contained continuous topic salience scores, where each cell reflected the strongest evidence that a given participant expressed that experiential theme (scores near 1 indicating



Lancaster Sensorimotor (LS) norms (Lynott et al. 2020) as described in Chkhaidze et al. (2023) (see Fig. 8a for the pipeline).

### Description preprocessing

We first preprocessed hallucination descriptions using standard natural language processing techniques in python with NLTK (Bird et al. 2009). This involved: (i) tokenization, (ii) removal of punctuation and conversion to lowercase, (iii) removal of English stopwords using the NLTK stopwords corpus, (iv) word lemmatization to reduce words to their base forms, and (5) spell correction using the pypellchecker library (Barrus 2025). Description length was calculated as the total number of tokens after preprocessing.

### Quantifying descriptions using Lancaster Sensorimotor norms

The LS Norms provide human-derived ratings for ~40 000 English words across 11 dimensions: six perceptual modalities (visual, auditory, gustatory, olfactory, haptic, and interoceptive) and five motor/action dimensions (foot, hand, head, mouth, and torso). For each word, human raters indicated how strongly they experience that concept through each sensory modality or through actions involving each body part. In addition to these individual modality ratings, the LS Norms include two composite scores: “perceptual strength” and “action strength”. These reflect the maximum strength value assigned to the dominant modality for each word—e.g. if a word is rated highest on the “visual” dimension, its perceptual strength is that visual rating. This approach is designed to capture the most salient sensory or motor modality through which a concept is typically experienced, and is theorized to reflect conceptual concreteness (Banks and Connell 2022).

For each hallucination description, we matched preprocessed words to the LS database and calculated average scores across all six perceptual, five motor LS dimensions, and two composite measures. Descriptions with fewer than three valid LS norm-matched words (~7%) were excluded and we ended up with 4 057 participants for this analysis.

### Regression models

We asked three research questions. First, we tested whether the overall perceptual and action richness of hallucination descriptions predicted participants’ self-reported imagery vividness. To do this, we fit a generalized linear model (GLM) predicting vividness from “perceptual strength” and “action strength”. The model was formulated as:

$$\text{visual vividness}_i = \beta_0 + \beta_1 \cdot \text{perceptualstrength}_i + \beta_2 \cdot \text{action strength}_i + \epsilon_i$$

Second, to clarify which specific perceptual dimensions were driving the overall perceptual strength effect, we fit a GLM including all six LS perceptual modality scores as simultaneous predictors:

$$\text{visual vividness}_i = \beta_0 + \beta_1 \cdot \text{visual}_i + \beta_2 \cdot \text{auditory}_i + \beta_3 \cdot \text{gustatory}_i + \beta_4 \cdot \text{olfactory}_i + \beta_5 \cdot \text{interoceptive}_i + \epsilon_i$$

Lastly, we asked whether the overall action strength effect was similarly driven by any specific motor modality. We fit a separate GLM using the five motor dimension scores to isolate their relative

contributions:

$$\text{visual vividness}_i = \beta_0 + \beta_1 \cdot \text{head}_i + \beta_2 \cdot \text{hand}_i + \beta_3 \cdot \text{mouth}_i + \beta_4 \cdot \text{foot}_i + \beta_5 \cdot \text{torso}_i + \epsilon_i$$

All predictors were z-scored prior to model fitting. GLM analyses were conducted in R using the “stats” package (R Core Team 2023).

### Controlling for response length via mediation

Initial analysis revealed that description length was a significant positive predictor of vividness ratings ( $\beta = .01$ ,  $SE = .00$ ,  $t = 6.80$ ,  $P < .001$ ), indicating that participants with more vivid imagery provided longer descriptions. To account for this potential confounding influence, we conducted a series of causal mediation analyses using the “mediation” package in R (Tingley et al. 2014). Specifically, for each set of predictors—(i) composite perceptual and action strength, (ii) the six perceptual modality scores, and (iii) the five motor modality scores—we assessed whether the effect of sensorimotor richness on visual vividness was mediated by description length.

Each mediation analysis included two models: a “mediator model” predicting standardized description length from the sensorimotor predictors of interest, and an “outcome model” predicting visual vividness from both the sensorimotor predictors and standardized description length.

Mediator model (example for a single predictor):

$$\text{description length}_i = \gamma + \gamma \cdot \text{sensorimotor predictor}_i + \eta_i$$

Outcome model:

$$\text{visual vividness}_i = \beta_0 + \beta_1 \cdot \text{sensorimotor predictor}_i + \beta_2 \cdot \text{description length}_i + \epsilon_i$$

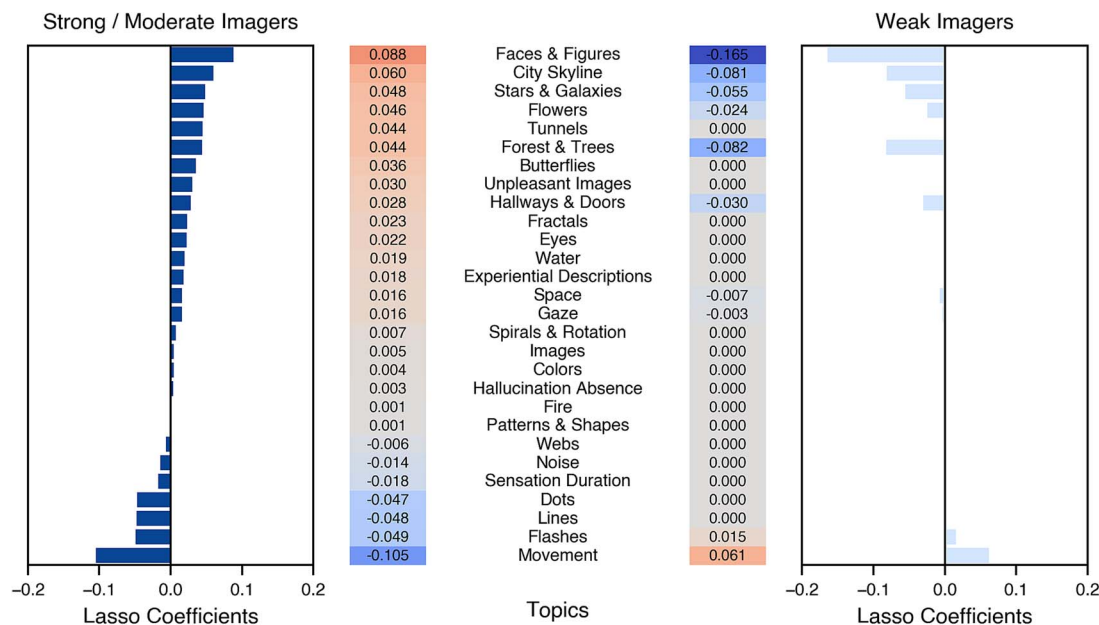
We estimated the average direct effect (ADE), average causal mediation effect (ACME), total effect, and proportion mediated for each predictor using 5000 nonparametric bootstrap simulations with percentile-based confidence intervals.

## Results

### Unsupervised topic modeling identifies systematic themes across the imagery spectrum

To map the latent content structure of participants’ hallucination descriptions, we used topic modeling. This unsupervised approach identifies patterns of word co-occurrence and groups semantically similar descriptions together, allowing us to reveal the types of experiences participants reported without imposing predetermined categories (Blei et al. 2003). Applying this method to the corpus yielded 28 distinct experiential themes, each reflecting a coherent hallucination type (e.g. geometric patterns, faces, tunnels, stars; Table 1).

The discovered topics encompassed a diverse range of hallucinations—from basic visual elements (e.g. lines, dots, color flashes), to scenery (e.g. forests and skylines), and socially-relevant objects (e.g. faces and figures) (Fig. 6, Table 1). Representative quotes from descriptions associated with each topic are in Table 1. Each row shows a free-text quote drawn from participant descriptions of their Ganzflicker-induced hallucinations, alongside the topic with which it was most strongly associated. For instance, the topic spirals and rotation included descriptions like “everything constantly spinning”,



**Figure 7.** Hallucination content patterns distinguishing weak from moderate/strong imagery groups. Lasso regression coefficients averaged across the moderate and strong imagery groups and contrasted with the weak group. The plot shows only topics that survived regularization and appeared in at least 60% of 1000 bootstrapped regressions. Colors indicate coefficient direction and stability, with positive values reflecting topics more prevalent in the moderate/strong imagery group and negative values reflecting topics more common in the weak group.

while the topic Space was associated with reports like “flying through a space station”. These examples show that our unsupervised model successfully captured linguistically distinct yet conceptually similar experiences, ranging from motion and geometric percepts to naturalistic scenes and affective content.

Semantic structure in the 2D UMAP (McInnes et al. 2018) embedding showed clearly delineated clusters, where spatial proximity reflects thematic similarity (Fig. 6). Each point represents a sentence from a participant’s hallucination description, embedded into a shared semantic space and colored by its assigned topic. The visualization revealed several broad regions: geometric and motion-related content (e.g. patterns and shapes, fractals, spirals and rotation, lines, dots, movement) formed one area; naturalistic and environmental imagery (e.g. forest and trees, water, flowers, space, city skyline, tunnels) formed another; and metacognitive domains (e.g. sensation duration and hallucination absence) appeared in a more separate region. Together, these analyses indicate that hallucination descriptions organize into coherent semantic domains with consistent relationships among topics. The resulting topic structure depicted in Fig. 6 offers a data-driven foundation to examine whether hallucination content varies systematically across individuals with different imagery phenotypes.

We hypothesized that participants with stronger imagery would describe more complex, naturalistic hallucination content, whereas weaker imagers would report simpler visual distortions and geometric forms. We observed systematic content differences consistent with this prediction: strong imagers more often reported structured, semantically meaningful content (e.g. faces, hallways, natural environments), while weak imagers tended to describe simpler features such as geometric patterns, visual distortions, and flashes (Fig. 7).

To test this quantitatively, we first modeled participants’ continuous visual imagery vividness scores using Lasso regression (Tibshirani 1996). Given any particular hallucination description, the topic model

output provides a probabilistic profile across all 28 topics, reflecting the estimated content composition of that report. We used these topic probabilities as features in our models to predict imagery scores. Topic probabilities explained ~5% of the variance in vividness scores ( $R^2 = .05$ ). From the initial set of 28 topics, 25 were retained as non-zero predictors (see Fig. S2). Given the noisy nature of free-text reports and the multifaceted character of subjective experience, this effect reflects meaningful structure in the data despite its modest size. As human psychology and behavior is inherently complex, small but reliable effects in psychological research are both expected and interpretable (Funder and Ozer 2019). Positive predictors included naturalistic or structured content—faces and figures ( $\beta = 0.29$ ), hallways and doors ( $\beta = 0.15$ ), stars ( $\beta = 0.14$ ), city skylines ( $\beta = 0.14$ ), and forest imagery ( $\beta = 0.13$ ). Negative predictors consisted mainly of simple perceptual content, including movement ( $\beta = -0.23$ ), flashes ( $\beta = -0.17$ ), and lines ( $\beta = -0.12$ ). These results indicate that richer, more naturalistic descriptions are associated with higher imagery vividness, while lower vividness is linked to simpler perceptual experiences.

To characterize the hallucination content profiles of distinct imagery phenotypes, we categorized participants into three vividness groups: weak (0–3), moderate (4–7), and strong (8–10). This binning followed conventions in the imagery literature distinguishing aphantasia, typical imagery, and hyperphantasia (Zeman 2024). For each group, we computed the average topic probability across participants (Fig. S3). To assess whether these profiles were discriminable, we then trained Lasso-regularized classifiers to predict group membership from topic probability features. We trained three separate one-vs-rest classifiers, where each classifier learned to identify one imagery group versus all others, revealing which topics are most diagnostic of each imagery phenotype. The classifiers were evaluated using the F1 score—the harmonic mean of precision and recall—which is well-suited to imbalanced classification tasks and helps assess how well models identify the minority class (Saito and Rehmsmeier 2015). For

statistical validation, we conducted permutation tests (Nichols and Holmes 2002) in which group labels were randomly shuffled and each model retrained 1 000 times. The weak-imagery classifier showed significant above-chance performance ( $F1 = 0.54$ ,  $P < .001$ ), indicating that those with the least vivid imagery can be distinguished from other participants based on their hallucination content. Classifiers for the moderate ( $F1 = 0.45$ ) and strong groups ( $F1 = 0.41$ ) did not exceed chance ( $P > .50$ ), likely reflecting similarity between these two groups (Fig. S3).

To interpret classifier outputs, we examined bootstrapped Lasso coefficients for the topics and retained only those that appeared in at least 60% of 1 000 bootstrapped models (Fig. 7). The weak-imagery profile was sparsely defined: only 8 of 28 topics were retained, mostly reflecting the absence of naturalistic content. Positive predictors for weak imagery included simple perceptual features (e.g. flashes, movement). In contrast, all 28 topics survived regularization and bootstrapping for the moderate and strong groups, and their strongest predictors overlapped with those from the continuous model. Notably, topics absent in weak-imagery descriptions (e.g. faces, natural scenes) were positive predictors of strong imagery, whereas topics characteristic of weak imagery were negative predictors of strong imagery (Fig. 7).

Together, these findings suggest that free-text descriptions of hallucinations contain discriminative signals related to individual imagery phenotypes. The semantic content of these descriptions offers a high-dimensional, quantifiable window into how internal visual experiences differ across the imagery spectrum.

### Human-assessed sensorimotor content of descriptions varies with imagery vividness

We expected that participants with stronger imagery vividness would use language with stronger sensorimotor associations when describing their hallucinations, particularly in visual and motor dimensions, reflecting richer internal simulation.

To quantify this content, we obtained Lancaster scores for each word in participants' descriptions across six perceptual modalities, five motor dimensions, and two composite measures, then averaged these scores within each participant to create 13 participant-level sensorimotor features (see Fig. 8a for analysis pipeline). Across the dataset, visual strength was the most prominent perceptual modality, reflected in its relatively high average score ( $M = 3.57$ ) and substantial individual variability (range = 3.87), followed by haptic ( $M = 1.32$ ) and auditory ( $M = 1.11$ ) dimensions. In the motor domain, head- ( $M = 2.66$ ) and hand-related ( $M = 1.56$ ) concepts were most frequently represented. The composite measures of perceptual and action strength also showed high mean scores and variability, supporting their interpretability as global indices of sensorimotor richness. Full score distributions and descriptive statistics are reported in Table S1 and Fig. S4.

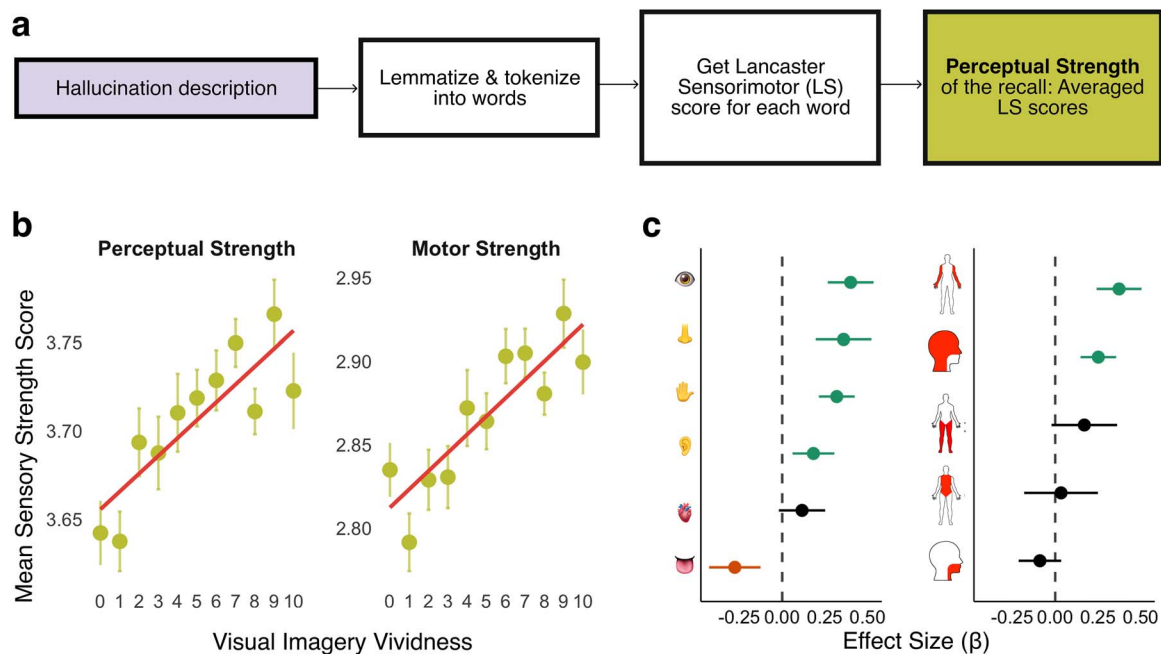
We first modeled imagery vividness using composite perceptual and motor (action) strength, controlling for description length. We entered the composite perceptual and motor (action) strength scores from the Lancaster Sensorimotor norms into a GLM, along with standardized description length as a covariate. Both predictors were significant: higher perceptual strength was associated with greater imagery vividness ( $\beta = 0.35$ ,  $SE = 0.05$ ,  $t = 6.54$ ,  $P < .001$ ), as was higher motor strength ( $\beta = 0.21$ ,  $SE = 0.05$ ,  $t = 4.17$ ,  $P < .001$ ) (Fig. 8b). Description length was also a significant positive predictor ( $\beta = 0.42$ ,  $SE = 0.05$ ,  $t = 8.39$ ,  $P < .001$ ), consistent with the idea that more vivid experiences elicit richer verbal responses.

To determine whether the effects of sensorimotor content were confounded or mediated by description length, we conducted causal mediation analyses with 5 000 bootstrap simulations. For perceptual strength, the total effect ( $\beta = 0.23$ ,  $P < .001$ ) was composed of a positive direct effect ( $ADE = 0.35$ ,  $P < .001$ ) and a significant "negative" indirect effect via description length ( $ACME = -0.12$ ,  $P < .001$ ), indicating a suppressive mediation. That is, the full influence of perceptual content was actually underestimated when not controlling for description length, with over half the total effect ( $-54\%$ ) masked by its association with longer responses. In contrast, action strength showed a small but significant positive mediation effect ( $ACME = 0.04$ ,  $P < .001$ ), contributing to a total effect of 0.25 ( $P < .001$ ), with 16% of this effect attributable to longer responses. These results suggest that both perceptual and action content are independently predictive of visual imagery vividness.

To identify which specific sensory modalities contributed to the observed effect of perceptual strength on visual imagery vividness, we fit a generalized linear model including all six Lancaster perceptual dimensions—visual, auditory, gustatory, olfactory, haptic, and interoceptive—as simultaneous predictors (see Fig. S5 for individual dimension trends). Description length was included as a covariate to control for potential verbosity differences.

The results revealed that five of the six perceptual modalities reliably predicted imagery vividness (Fig. 8c, left). Stronger visual content in hallucination descriptions was associated with higher vividness ratings ( $\beta = 0.40$ ,  $P < .001$ ), as was greater haptic ( $\beta = 0.32$ ,  $P < .001$ ), olfactory ( $\beta = 0.36$ ,  $P < .001$ ), and auditory content ( $\beta = 0.18$ ,  $P = .003$ ). Interestingly, gustatory strength was negatively associated with vividness ( $\beta = -0.28$ ,  $P < .001$ ), while interoceptive content showed a positive but non-significant effect ( $\beta = 0.12$ ,  $P = .093$ ).

To account for description length, we ran mediation analyses for each modality. These analyses confirmed that the core effects remained robust even after controlling for verbosity. Specifically, the direct effect of visual strength remained strong and positive ( $ADE = 0.40$ ,  $P < .001$ ), even though the total effect was slightly suppressed by a significant negative mediation via description length ( $ACME = -0.10$ ). This indicates that participants with more vivid imagery did not merely write longer descriptions; rather, the visual content of those descriptions was independently predictive of vividness. Haptic and olfactory strength also retained significant direct effects after mediation ( $ADEs = 0.32$  and  $0.36$ , respectively; both  $P < .001$ ), suggesting that these less commonly studied sensory modalities were meaningfully tied to self-reported individual differences in visual imagery. Although olfactory strength survived mediation analysis, its low mean and restricted range suggest that it should be interpreted with caution, as it may reflect isolated rather than widespread patterns of olfactory content. Auditory content also remained significant ( $ADE = 0.18$ ,  $P = .003$ ), though its mediated component was small but positive ( $ACME = 0.022$ ). The effect of auditory content on imagery vividness is mostly direct and robust, although there was a small, significant indirect effect of description length. The effect of interoceptive content remained non-significant after mediation ( $ADE = 0.12$ ,  $P = .10$ ; CI includes 0), indicating no additional influence of verbosity. While gustatory strength showed a significant negative coefficient in the regression model, its mediation results and real-data trend (Fig. S5) suggest interpretational caution. The model returned a negative direct effect ( $ADE = -0.28$ ,  $P < .001$ ), but raw data binned by vividness showed a slight upward trend in gustatory content, likely driven by sparse outliers. This dimension also



**Figure 8.** Lancaster sensorimotor analysis pipeline and results. (a) Analysis pipeline showing how hallucination descriptions were processed to extract sensorimotor content using the Lancaster sensorimotor norms. Text was lemmatized and tokenized, each word received Lancaster sensorimotor scores, and perceptual strength was calculated by averaging LS scores of individual words in each description. (b) Relationship between visual imagery vividness and composite sensorimotor dimensions. Both perceptual strength (left) and motor strength (right) showed significant positive associations with self-reported imagery vividness. Points represent mean scores  $\pm$  standard error for each vividness level. (c) Forest plots showing effect sizes ( $\beta$  coefficients) from generalized linear models predicting imagery vividness from individual sensorimotor dimensions, controlling for description length. Left panel shows perceptual modalities with icons representing visual (eye), olfactory (nose), haptic (hand), auditory (ear), interoceptive (heart), and gustatory (tongue) dimensions. Right panel shows motor dimensions with body part icons for head, hand, mouth, torso, and foot. Points show  $\beta$  estimates, and error bars represent 95% confidence intervals (CIs). The dashed vertical line marks  $\beta = 0$ ; effects whose CIs do not overlap this line are significant.

had the lowest mean (0.30) and smallest variance (0.035) among all predictors, suggesting that gustatory language was rarely present and potentially contributed noise. Given this, we refrain from interpreting its effect as meaningful. Together, these findings suggest that visual imagery vividness is reflected not only in explicitly visual language but also in modality-rich descriptions involving haptic, olfactory, and auditory references in addition to action-related language. The robustness of these effects after controlling for verbosity strengthens the claim that these are not mere artifacts of expressive ability or fluency.

To determine which body-related dimensions accounted for the effect of action strength on visual imagery vividness, we next modeled each of the five Lancaster motor dimensions—head, hand, mouth, torso, and foot strength—as simultaneous predictors of vividness (Fig. S5 shows individual trends for each motor dimension). Description length was included as a covariate. The model revealed significant positive effects for hand strength ( $\beta = 0.37$ ,  $P < .001$ ) and head strength ( $\beta = 0.25$ ,  $P < .001$ ), suggesting that vivid imagers were more likely to describe content referencing the head or hands (e.g. “looking around”, “reaching”, “hands in front of me”). The effects of foot, mouth, and torso strength were not statistically significant ( $P_s > .05$ ) (Fig. 8c, right).

We followed up this model with separate causal mediation analyses to account for the role of description length. As in earlier models, mediation analyses were conducted for each motor dimension using 5000 bootstrap simulations. Both hand strength and head strength remained robust predictors after controlling for length. For hand

strength, the direct effect (ADE) remained significant (ADE = 0.37,  $P < .001$ ), while the indirect effect (ACME =  $-0.056$ ,  $P < .001$ ) indicated a suppressive relationship with length: longer responses slightly diluted the observed association, but did not eliminate it. A similar pattern held for head strength (ADE = 0.25,  $P < .001$ ; ACME =  $-0.031$ ,  $P < .001$ ). In contrast, the effects of foot, mouth, and torso strength did not survive mediation analysis. Foot strength showed a small positive total effect (0.17), but neither the direct nor indirect paths reached significance ( $P = .07$  and  $.78$ , respectively). Torso strength had a positive indirect effect (ACME = 0.031,  $P < .001$ ), but its direct and total effects were non-significant, and its mediation estimate was highly unstable (Prop. Mediated = 0.48, 95% CI =  $-3.77$  to 4.10). Mouth strength showed a suppressive pattern (Prop. Mediated =  $-1.65$ ), but neither its direct nor total effect was significant ( $P = .16$ ), and the instability of this estimate again likely reflects noise due to sparse representation.

These results highlight head- and hand-related content as the most consistent motor predictors of imagery vividness. Their predictive power held even when accounting for description length and shared variance with other motor dimensions. This pattern suggests that vivid internal experiences are more likely to include references to sensorimotor processes involving head movements and hand interactions.

Prior work shows that women report higher visual vividness scores than men on the VVIQ (Richardson 1995, Aydin 2020) and self-reported vividness declines across the lifespan (Gulyás et al. 2022). To evaluate whether these factors influenced imagery ratings in our sample, we regressed vividness scores on age and gender. Age was

not a significant predictor ( $\beta = -0.0002$ ,  $P = .95$ ), but consistent with the prior literature, women reported reliably higher vividness than men ( $\beta = -0.53$ ,  $P < .001$ ). To ensure that the main effects reported above were not confounded by gender differences, we re-ran all Lancaster sensorimotor models including gender as a covariate. All sensorimotor predictors that were significant in the primary models remained significant with minimal change in effect size (Fig. S6).

## Discussion

### Overview

Through computational analysis of >4000 phenomenological reports, we characterized the content and structure of Ganzflicker-induced hallucinations and examined how this content varies across the visual imagery spectrum. Results from our complementary analytical approaches—topic modeling and sensorimotor norms—converged to reveal systematic differences in hallucination phenomenology consistent with the layered model of visual imagery. Below, we situate our findings within the broader literature on experimentally-induced hallucinations and ASC, and discuss what we learned about individual differences in visual imagery.

### Phenomenological profile of Ganzflicker-induced hallucinations

Our topic modeling identified 28 distinct experiential themes, providing a detailed characterization of Ganzflicker-induced hallucination phenomenology. This granular phenomenological mapping extends prior work on flicker-induced experiences, which has typically relied on structured questionnaires or broad categorical distinctions (Bartossek et al. 2021, Hewitt et al. 2025). Consistent with decades of research on light-induced hallucinations (Klüver 1926, Bressloff et al. 2001, 2002, Hewitt et al. 2025), simple geometric content was among the most frequently reported experiences. Our participants described classical “form constants” (Klüver 1926): spirals, tunnels, and cobwebs that recur across diverse altered states and induction methods, from Ganzfeld (Wackermann et al. 2008, Pistolas and Wagemans 2025) to stroboscopic light stimulation (Bartossek et al. 2021, Hewitt et al. 2025). However, our analysis also documented complex, semantically meaningful content: faces, human figures, city skylines, forest scenery, stars, and galaxies—naturalistic imagery that has previously been reported in flicker paradigms (Allefeld et al. 2011, Schwartzman et al. 2019, Königsmark et al. 2021, Shenyan et al. 2024, Beauté et al. 2025), but not characterized at this level of detail.

The hierarchical clustering of these 28 topics (Fig. S1) revealed three experiential domains that mirror the visual processing hierarchy: simple geometric content (e.g. lines, flashes, patterns, dots, colors), naturalistic scenes (e.g. skylines, forests, water, butterflies), and a cluster of social (faces, figures, eyes) and metacognitive (experiential descriptions, sensation duration) topics. This data-driven organization aligns with theoretical frameworks proposing that hallucinations evolve from geometric forms to structured experiences as higher-order visual areas become engaged (ffytche 2008, Reeder et al. 2024).

We also examined whether Ganzflicker induced alterations in consciousness beyond visual hallucinations, following the taxonomy of ASC features introduced by Fort et al. (2025). While our results primarily implicated the Perception and Imagery dimensions, participants also reported experiences that may relate to other ASC features: feelings of being “transported to another place”, “dreamlike state”,

and “lost sense of space” share phenomenological overlap with the Self-Boundary feature (which includes depersonalization and derealization experiences), while “lost a sense of time” aligns with the Time Sense feature. Participants also reported Emotion changes ranging from enjoyable to unpleasant. These findings situate Ganzflicker within the broader ASC literature, demonstrating that, like other non-pharmacological induction methods (Berkovich-Ohana et al. 2013, Kübel et al. 2021), it produces alterations extending beyond perception to encompass spatial disorientation, dreamlike absorption, and emotional engagement.

Importantly, experienced ASCs varied systematically across the imagery spectrum. Vivid imagers were significantly more likely to report feeling self-boundary alterations, including feelings of being transported, dreamlike states, and spatial disorientation (Fig. S7). By contrast, weak imagers were more likely to report no changes in consciousness. Emotional experiences also varied across individuals: vivid imagers rated their experience as more enjoyable (Fig. S8), even while describing more affectively intense visual content (e.g. “unpleasant images”). This finding suggests that vivid hallucinations may be experienced as engaging even when their content is disturbing, and fits with the claim that the Emotion feature interacts with but is conceptually distinct from perceptual content (Fort et al. 2025). Imagery was not related to the alterations in time perception, suggesting that Time Sense might be a relatively stable and broadly shared feature of Ganzflicker experience, invariant across the imagery spectrum. Beyond these broader experiential differences, hallucination content itself varied systematically with imagery capacity.

### Hallucinated content maps onto the visual processing hierarchy across the imagery spectrum

Imagery research has traditionally treated vividness as a single continuum, capturing how clearly or intensely mental images are experienced. Our findings suggest that individuals differ not just in the intensity of their imagery, but the level of representational complexity they can achieve. Accordingly, these findings are consistent with the layered model of visual imagery outlined in the introduction (Fig. 2). Inspired by neuroimaging evidence that aphantasic individuals retain early visual representations during imagery (Cabbai et al. 2024, Chang et al. 2025) this model holds that all individuals represent basic visual features in the early visual cortex. Further, since aphantasics show reduced functional connectivity between frontoparietal control networks and higher-order visual regions (Milton et al. 2021, Spagna et al. 2021, Liu et al. 2025), the layered model suggests stronger imagers more effectively engage higher-order regions to integrate basic visual features into coherent, semantically meaningful representations.

If imagery vividness is a scalar intensity parameter affecting all visual processing equally, we would expect proportional scaling of both simple and complex content across the imagery spectrum. Instead, we observed a qualitative shift: weak imagers described the building blocks of visual experience, while strong imagers reported content requiring top-down integration across the full cortical hierarchy. We formalize this mapping in Table 2, which links hallucination topics to their predicted processing levels based on their association with imagery vividness.

Even individuals reporting the complete absence of imagery described seeing something during Ganzflicker—but what they saw differed qualitatively from what stronger imagers reported.

**Table 2.** Hallucination content mapped onto visual processing hierarchy across the imagery spectrum.

Processing level	Area	Weak imagers/aphantasia	Moderate–strong imagers
Visual primitives Intermediate	V1/V2 V3/V4, MT/V5	lines, dots, flashes movement, colors,	Lines, dots, flashes Movement, colors, patterns and shapes, spirals and rotation, tunnels, webs
Semantically meaningful objects	IT	—	Butterflies, flowers, stars and galaxies, faces and figures, eyes, city skyline, forests and trees, hallways and doors, space
Control/integration	Frontoparietal, insula	—	Experiential descriptions, sensation duration

Weak imagers predominantly described simple visual experiences like geometric forms: lines, dots, flashes, and movement—content consistent with accounts of the early visual cortex involvement during hallucinations, like form constants arising from V1/V2 dynamics (Ermentrout and Cowan 1979, Bressloff et al. 2002, Burke 2002, Rule et al. 2011) and motion-selective intermediate areas like V3/MT (Amaya et al. 2023, 2025). Strong imagers, by contrast, reported complex, naturalistic, and socially relevant content: faces and figures, city skylines, forests, butterflies, and immersive spatial experiences. This type of content engages later extrastriate and ventral temporal regions that represent more complex and multi-dimensional stimuli (Kanwisher et al. 1997, Epstein and Kanwisher 1998).

Neural accounts of flicker-induced hallucinations also support our proposed link between individual differences in visual imagery and the extent of hierarchical interaction in the visual system. For example, computational models propose that simple geometric hallucinations arise from dynamical interactions within early visual cortex, where local excitatory and long-range inhibitory connections generate repeating patterns that resemble Klüver's form constants (Ermentrout and Cowan 1979, Bressloff et al. 2002, Rule et al. 2011). Complex naturalistic content, by contrast, requires engagement of higher-order visual areas (Shenyan et al. 2024; Stoliker et al. 2025). In addition to cortical hierarchical models, neural accounts have also highlighted a potential role for thalamocortical interactions in flicker-induced hallucinations (see Hewitt et al. 2025 for a review).

In fact, more effective thalamocortical engagement of higher-order visual areas could explain the individual differences in hallucination complexity observed in the present study. The thalamus acts as a central hub regulating cortical excitability and integrating information across networks (Hwang et al. 2017, Shine 2021), with ventral projections targeting supragranular cortical layers implicated in conscious experience (Phillips et al. 2019, Suzuki and Larkum 2020). Just as enhanced top-down connectivity has been shown to be relevant for hyperphantasics' voluntary imagery (Milton et al. 2021), fMRI studies reveal that periodic flicker stimulation produces thalamocortical hyperconnectivity between ventral thalamic nuclei and higher-order visual cortices, with connectivity strength correlating with subjective experience of hallucination intensity (Amaya et al. 2023, 2025). Future studies could test whether there is a relationship between imagery capacity, thalamocortical connectivity, and the complexity of flicker-induced hallucinations, as well as the complexity of other experiences that rely on the generative capacity of the visual system.

### Sensorimotor content of hallucination descriptions may reflect multisensory simulation

To complement computational analyses with human-derived measures, we turned to the Lancaster Sensorimotor Norms to examine the perceptual and action-related content of participants' descriptions using crowd-sourced judgments. We found that descriptions from individuals with higher imagery vividness contained more words with strong visual, haptic, and auditory content, as well as head- and hand-related motor associations. Notably, stronger imagers used richer language not only in the visual domain but also for other perceptual content (see Fig. 8c), consistent with evidence that imagery differences often extend beyond vision, with many aphantasics reporting weak imagery across sensory modalities (Zeman et al. 2020). This suggests that what we measure as “visual” imagery vividness may index broader variation in multisensory simulation capacity (Dance et al. 2021). In short, strong imagers used language with richer perceptual connotations, suggestive of more grounded and content-dense internal simulations.

Potentially at odds with predictive-coding accounts proposing that interoception and agency play a central role in imagery generation (Silvanto 2025, Silvanto and Nagai 2025), we failed to find a relationship between interoceptive language and imagery strength. However, the Lancaster norms quantify explicit interoceptive language, whereas the relevant theoretical accounts emphasize implicit mechanisms like agency or bodily prediction. These may not surface in verbal descriptions when participants are instructed to report what they saw. When directly asked about their conscious experience, however, stronger imagers were more likely to report spatial disorientation and dreamlike absorption, and to evaluate the experience as emotionally enjoyable compared to those with weak visual imagery (Figs S7 and S8). This confirms that embodied and affective experience do vary with imagery vividness, even when not captured by explicit language norms.

The most robust predictors of individuals' visual imagery vividness were visual, head- and hand-related language, consistent with our layered model, in which stronger imagers more effectively engage high-level perceptual systems that support complex, socially meaningful internal content. Faces and hands are among the most socially and evolutionarily significant concepts for humans (Bracci et al. 2010, Lacruz et al. 2019), each with its own dedicated neural circuitry (Kanwisher et al. 1997, Bracci et al. 2010). When higher-order areas are recruited during complex hallucinations, these strongest “defaults” may surface first. As individuals with aphantasia show reduced activation of the fusiform gyrus during imagery tasks (Liu et al. 2025),

stronger imagers might more effectively recruit specialized regions such as the fusiform face area, during Ganzflicker-induced hallucinations, resulting in experiencing richer and more complex visual content, as well as possibly richer multisensory associations.

## Limitations

Our study drew on a large online sample (~4000 participants), allowing us to capture rare individual differences (e.g. aphantasia and hyperphantasia) and to examine Ganzflicker experiences across diverse contexts. However, online studies come with inherent constraints. For example, we could not directly control participants' viewing environments. Although all participants viewed the same 7.5 Hz red–black flicker stimulus, variations in the properties of their monitors (e.g. luminance, color saturation, timing stability) and viewing conditions (e.g. lighting, distractions, use of white noise) may have introduced heterogeneity into their perceptual responses. Without standardized conditions, the data could in principle represent a heterogeneous set of experiences rather than differences related to participants' position on the imagery spectrum.

However, several factors support the interpretation that our findings reflect genuine Ganzflicker phenomenology. Prior analyses of the impact of environmental variables in this dataset, including white noise use, lighting conditions, or display medium, failed to reveal significant effects on reported hallucination content (see Reeder 2022 for details; see also Königsmark et al. 2021). Further, the hallucination content documented here aligns with that reported in prior Ganzflicker studies using different samples and setups (Königsmark et al. 2021), including one conducted in a controlled laboratory environment (Shenyan et al. 2024). Nevertheless, online data collection on its own cannot fully rule out the contribution of uncontrolled environmental factors, and future work with larger controlled laboratory samples would strengthen confidence in these findings.

Another limitation of our study relates to participant recruitment via a public article about Ganzflicker. It seems likely that this may have attracted individuals already interested in hallucination phenomena as well as those at the extremes of the imagery spectrum, in keeping with the higher-than-typical proportion of self-reported aphantasia in our sample. This raises the possibility that observed results reflect demand or expectancy effects, with those interested in hallucinations eager to report ASCs and those who identify as aphantasic eager to produce descriptions reflecting their expectations rather than their actual experiences. However, Ganzflicker-induced hallucinations reliably emerge in laboratory settings without self-selection (Shenyan et al. 2024). Further, in the present study, even participants reporting no visual imagery described hallucinations, arguing against expectancy alone as an explanation (Reeder 2022). Finally, hallucination content changed predictably with imagery vividness—a pattern inconsistent with simple demand effects.

Yet another limitation of the present study is our use of participants' verbal descriptions as an index of their subjective visual experiences. Although language provides a powerful window into the content of participants' hallucinations, their descriptions necessarily conflate phenomenology with narrative expression. Individuals may differ not only in what they experience but also in how they verbalize those experiences. Notably, participants with more vivid imagery provided longer descriptions, raising the possibility that verbosity rather than experiential richness drives some effects. To address this, our analyses of Lancaster Sensorimotor Norms incorporated measures to control

for description length, finding that vividness effects remained significant. Corroborating evidence also comes from closed-ended measures: participants with higher imagery vividness reported greater alterations in consciousness and more positive affect during Ganzflicker stimulation on questions that did not require generating linguistic descriptions (Figs S7 and S8). That said, verbal reports alone cannot adjudicate between differences in what people experience and differences in how they describe it. Future work incorporating objective measures such as EEG, eye tracking, and physiological markers will be critical for directly assessing perceptual dynamics during Ganzflicker stimulation.

## Conclusion

Together, these findings offer a rich characterization of Ganzflicker-induced experience, from the diversity of visual hallucinations to broader alterations in consciousness, and reveal how both vary systematically across the visual imagery spectrum. Overall, results align with the theory of divergent predictive perception (Reeder et al. 2024), which proposes that individuals across the imagery spectrum differ in how strongly top–down predictions shape perceptual experience. Under this account, strong imagers generate richer predictive models that more readily engage higher-order visual representations, while weak imagers rely more heavily on bottom–up input. The Ganzflicker paradigm—which provides rhythmic stimulation with minimal structured input—may be particularly sensitive to these differences: when bottom–up information is impoverished, individual variation in top–down generative capacity becomes visible in the content of conscious experience.

Beyond Ganzflicker, these findings may have broader implications for understanding human cognition. Individual differences in imagery are not merely subjective curiosities but have been shown to influence a wide range of cognitive processes, from shaping our emotions (Wicken et al. 2021) and moral judgments (Amit and Greene 2012) to affecting how we adopt scientific theories (Sulik et al. 2025). By incorporating phenomenological variability into our models, we move toward a more complete and inclusive science of the mind. Further, this work demonstrates that computational analysis of free-text reports can reveal the hidden structure of complex mental phenomena, offering a scalable approach for studying individual differences in cognitive processes that are otherwise difficult to access (Feuerriegel et al. 2025). By treating language as a window into the mind's eye, we can begin to capture how people with different imagery phenotypes internally construct their visual world.

## Acknowledgements

The authors thank Sean Trott for helpful suggestions regarding statistical analysis, and the Brain & Cognition Lab and the Cognitive Neuroscience Lab at UC San Diego for valuable feedback and discussion.

## Author contributions

Ana Chkhaidze (Conceptualization [equal], Formal analysis [lead], Software [lead], Visualization [lead], Writing—original draft [lead], Writing—review & editing [equal]), Reshane R. Reeder (Conceptualization [equal], Investigation [equal], Writing—review & editing [equal]), Connor Gag (Formal analysis [supporting], Software [supporting], Visualization [supporting], Writing—review & editing

[equal]), Anastasia Kiyonaga (Conceptualization [equal], Writing—review & editing [equal]), Seana Coulson (Conceptualization [equal], Supervision [equal], Writing—original draft [supporting], Writing—review & editing [equal])

## Supplementary data

Supplementary data is available at *Neuroscience of Consciousness* online.

## Conflict of interest

The authors declare no competing interests.

None declared.

## Funding

This research was supported by an Innovative Research Grant from the Kavli Institute for Brain and Mind at UC San Diego awarded to A.C., S.C., and A.K.

## Data availability

The human dataset of the Ganzflicker study is from (Reeder 2022), and is accessible at: <https://osf.io/6dvh9/>, and the cleaned dataset we used in this paper is available at: [https://github.com/anachkhaidze/ganzflickerNLPalooza\\_paper/tree/main/data](https://github.com/anachkhaidze/ganzflickerNLPalooza_paper/tree/main/data).

The Lancaster Norms is from (Lynott et al. (2020)), and the data can be found at: [https://github.com/anachkhaidze/ganzflickerNLPalooza\\_paper/tree/main/data](https://github.com/anachkhaidze/ganzflickerNLPalooza_paper/tree/main/data).

## Materials availability

Ganzflicker experience and questionnaires can be found at: <https://forms.gle/tDKRKhva3uqC68tS9>

## Code availability

Analysis scripts (in R and Python) for all components of the study, including topic modeling and Lancaster Sensorimotor Norms analysis, are available at: [https://github.com/anachkhaidze/ganzflickerNLPalooza\\_paper](https://github.com/anachkhaidze/ganzflickerNLPalooza_paper).

The full pipeline is procedurally reproducible: all scripts, configuration files, and random seeds are provided to enable replication. Minor numerical variation may occur due to stochasticity in embedding models; however, the overall topic structure and model-derived similarity patterns remain stable across runs.

## References

- Albers AM, Kok P, Toni I *et al*. Shared representations for working memory and mental imagery in early visual cortex. *Curr Biol* 2013;**23**:1427–31. <https://doi.org/10.1016/j.cub.2013.05.065>
- Allefeld C, Pütz P, Kastner K *et al*. Flicker-light induced visual phenomena: frequency dependence and specificity of whole percepts and percept features. *Conscious Cogn* 2011;**20**:1344–62. <https://doi.org/10.1016/j.concog.2010.10.026>
- Amaya IA, Behrens N, Schwartzman DJ *et al*. Effect of frequency and rhythmicity on flicker light-induced hallucinatory phenomena. *PLoS One* 2023;**18**:e0284271. <https://doi.org/10.1371/journal.pone.0284271>
- Amaya IA, Nierhaus T, Schmidt TT. Thalamocortical interactions reflecting the intensity of flicker light-induced visual hallucinatory phenomena. *Netw Neurosci* 2025;**9**:1–17. [https://doi.org/10.1162/netn\\_a\\_00417](https://doi.org/10.1162/netn_a_00417)
- Amit E, Greene JD. You see, the ends don't justify the means: visual imagery and moral judgment. *Psychol Sci* 2012;**23**:861–8. <https://doi.org/10.1177/0956797611434965>
- Aydın Ç. Gender differences in visual imagery: object and spatial imagery. *Dokuz Eylül Üniversitesi Sosyal Bilimler Enstitüsü Dergisi* 2020;**22**:1045–64. <https://doi.org/10.16953/deusosbil.531463>
- Bainbridge WA, Pounder Z, Eardley AF *et al*. Quantifying aphantasia through drawing: those without visual imagery show deficits in object but not spatial memory. *Cortex* 2021;**135**:159–72. <https://doi.org/10.1016/j.cortex.2020.11.014>
- Banks B, Connell L. Multi-dimensional sensorimotor grounding of concrete and abstract categories. *Philos Trans R Soc B Biol Sci* 2022;**378**:20210366. <https://doi.org/10.1098/rstb.2021.0366>
- Barrus T. *pyspellchecker [Python software / GitHub repository]* San Francisco, CA: GitHub; 2025. <https://github.com/barrust/pyspellchecker>
- Bartossek MT, Kemmerer J, Schmidt TT. Altered states phenomena induced by visual flicker light stimulation. *PLoS One* 2021;**16**:e0253779. <https://doi.org/10.1371/journal.pone.0253779>
- Beauté R, Schwartzman DJ, Dumas G *et al*. Mapping of subjective accounts into interpreted clusters (MOSAIC): topic modelling and LLM applied to stroboscopic phenomenology. *Neurosci Conscious* 2026;**2026**:niag008. <https://doi.org/10.1093/nc/niag008>
- Bergmann J, Genç E, Kohler A *et al*. Smaller primary visual cortex is associated with stronger, but less precise mental imagery. *Cereb Cortex* 2016;**26**:3838–50. <https://doi.org/10.1093/cercor/bhv186>
- Berkovich-Ohana A, Dor-Ziderman Y, Glicksohn J *et al*. Alterations in the sense of time, space, and body in the mindfulness-trained brain: a neurophenomenologically-guided MEG study. *Front Psychol* 2013;**4**:912. <https://doi.org/10.3389/fpsyg.2013.00912>
- Bird S, Klein E, Loper E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*. Sebastopol, CA: O'Reilly Media, Inc, 2009.
- Bird CM, Bisby JA, Burgess N. The hippocampus and spatial constraints on mental imagery. *Front Hum Neurosci* 2012;**6**:142. <https://doi.org/10.3389/fnhum.2012.00142>
- Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res* 2003;**3**:993–1022.
- Bracci S, Ietswaart M, Peelen MV *et al*. Dissociable neural responses to hands and non-hand body parts in human left extrastriate visual cortex. *J Neurophysiol* 2010;**103**:3389–97. <https://doi.org/10.1152/jn.00215.2010>
- Bressloff PC, Cowan JD, Golubitsky M *et al*. Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences* 2001;**356**:299–330. <https://doi.org/10.1098/rstb.2000.0769>
- Bressloff PC, Cowan JD, Golubitsky M *et al*. What geometric visual hallucinations tell us about the visual cortex. *Neural Comput* 2002;**14**:473–91. <https://doi.org/10.1162/089976602317250861>
- Burke W. The neural basis of Charles bonnet hallucinations: a hypothesis. *J Neurol Neurosurg Psychiatry* 2002;**73**:535–41. <https://doi.org/10.1136/jnnp.73.5.535>
- Cabbai G, Racey C, Simner J *et al*. Sensory representations in primary visual cortex are not sufficient for subjective imagery. *Curr Biol* 2024;**34**:5073–5082.e5. <https://doi.org/10.1016/j.cub.2024.09.062>
- RJGB Campello, D Moulavi, J Sander (2013). Density-based clustering based on hierarchical density estimates. In J Pei, VS Tseng, L Cao, H Motoda, G Xu (Eds.), *Advances in Knowledge Discovery and Data Mining*

- (pp. 160–72). Berlin, Heidelberg: Springer. [https://doi.org/10.1007/978-3-642-37456-2\\_14](https://doi.org/10.1007/978-3-642-37456-2_14)
- Chang S, Zhang X, Cao Y *et al*. Imageless imagery in aphantasia revealed by early visual cortex decoding. *Curr Biol* 2025;**35**:591–599.e4. <https://doi.org/10.1016/j.cub.2024.12.012>
- Chkhaidze A, Coulson S, Kiyonaga A. Individual differences in preferred thought formats predict features of narrative recall. *Proc Annu Meet Cogn Sci Soc* 2023;**45**:1624–1629. <https://escholarship.org/uc/item/0jm2d2rm>
- Christian BM, Miles LK, Parkinson C *et al*. Visual perspective and the characteristics of mind wandering. *Front Psychol* 2013;**4**:699. <https://doi.org/10.3389/fpsyg.2013.00699>
- Connell L, Lynott D. Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition* 2012;**125**:452–65. <https://doi.org/10.1016/j.cognition.2012.07.010>
- Cui X, Jeter CB, Yang D *et al*. Vividness of mental imagery: individual variability can be measured objectively. *Vis Res* 2007;**47**:474–8. <https://doi.org/10.1016/j.visres.2006.11.013>
- Dance CJ, Ward J, Simner J. What is the link between mental imagery and sensory sensitivity? Insights from Aphantasia. *Perception* 2021;**50**:757–82. <https://doi.org/10.1177/03010066211042186>
- Dijkstra N, Bosch SE, van Gerven MAJ. Shared neural mechanisms of visual perception and imagery. *Trends Cogn Sci* 2019;**23**:423–34. <https://doi.org/10.1016/j.tics.2019.02.004>
- Epstein R, Kanwisher N. A cortical representation of the local visual environment. *Nature* 1998;**392**:598–601. <https://doi.org/10.1038/33402>
- Ermentrout GB, Cowan JD. A mathematical theory of visual hallucination patterns. *Biol Cybern* 1979;**34**:137–50. <https://doi.org/10.1007/BF00336965>
- Feuerriegel S, Maarouf A, Bär D *et al*. Using natural language processing to analyse text data in behavioural science. *Nat Rev Psychol* 2025;**4**:96–111. <https://doi.org/10.1038/s44159-024-00392-z>
- ffytche DH. The hodology of hallucinations. *Cortex* 2008;**44**:1067–83. <https://doi.org/10.1016/j.cortex.2008.04.005>
- Fort LD, Costines C, Wittmann M *et al*. Classification schemes of altered states of consciousness. *Neurosci Biobehav Rev* 2025;**175**:106178. <https://doi.org/10.1016/j.neubiorev.2025.106178>
- Funder DC, Ozer DJ. Evaluating effect size in psychological research: sense and nonsense. *Adv Methods Pract Psychol Sci* 2019;**2**:156–68. <https://doi.org/10.1177/2515245919847202>
- Grootendorst M. BERTopic: neural topic modeling with a class-based TF-IDF procedure. *CoRR*. 2022;abs/2203.05794. <https://arxiv.org/abs/2203.05794v1>
- Gulyás E, Gombos F, Sütöri S *et al*. Visual imagery vividness declines across the lifespan. *Cortex* 2022;**154**:365–74. <https://doi.org/10.1016/j.cortex.2022.06.011>
- Hewitt T, Amaya I, Beauté R *et al*. Stroboscopically induced visual hallucinations: historical, phenomenological, and neurobiological perspectives. *Neurosci Conscious* 2025;**2025**:niaf020. <https://doi.org/10.1093/nc/niaf020>
- Hwang K, Bertolero MA, Liu WB *et al*. The human thalamus is an integrative hub for functional brain networks. *J Neurosci* 2017;**37**:5594–607. <https://doi.org/10.1523/JNEUROSCI.0067-17.2017>
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 1997;**17**:4302–11. <https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997>
- Keogh R, Pearson J. The sensory strength of voluntary visual imagery predicts visual working memory capacity. *J Vis* 2014;**14**:7. <https://doi.org/10.1167/14.12.7>
- Kiss T, Strunk J. Unsupervised multilingual sentence boundary detection. *Computational Linguistics* 2006;**32**:485–525. <https://doi.org/10.1162/coli.2006.32.4.485>
- Klüver H. Mescal visions and eidetic vision. *Am J Psychol* 1926;**37**:502–15. <https://doi.org/10.2307/1414910>
- Königsmark VT, Bergmann J, Reeder RR. The Ganzflicker experience: high probability of seeing vivid and complex pseudo-hallucinations with imagery but not aphantasia. *Cortex* 2021;**141**:522–34. <https://doi.org/10.1016/j.cortex.2021.05.007>
- Kosslyn SM, Brunn J, Cave KR *et al*. Individual differences in mental imagery ability: a computational analysis. *Cognition* 1984;**18**:195–243. [https://doi.org/10.1016/0010-0277\(84\)90025-8](https://doi.org/10.1016/0010-0277(84)90025-8)
- Kübel SL, Fiedler H, Wittmann M. Red visual stimulation in the ganzfeld leads to a relative overestimation of duration compared to green. *Psych J* 2021;**10**:5–19. <https://doi.org/10.1002/pchj.395>
- Lacruz RS, Stringer CB, Kimbel WH *et al*. The evolutionary history of the human face. *Nat Ecol Evol* 2019;**3**:726–36. <https://doi.org/10.1038/s41559-019-0865-7>
- Lee S-H, Kravitz DJ, Baker CI. Disentangling visual imagery and perception of real-world objects. *NeuroImage* 2012;**59**:4064–73. <https://doi.org/10.1016/j.neuroimage.2011.10.055>
- Liu J, Zhan M, Hajhajate D *et al*. Visual mental imagery in typical imagers and in aphantasia: a millimeter-scale 7-T fMRI study. *Cortex* 2025;**185**:113–32. <https://doi.org/10.1016/j.cortex.2025.01.013>
- Lupyan G, Uchiyama R, Thompson B *et al*. Hidden differences in phenomenal experience. *Cogn Sci* 2023;**47**:e13239. <https://doi.org/10.1111/cogs.13239>
- Lynott D, Connell L, Brysbaert M *et al*. The Lancaster sensorimotor norms: multidimensional measures of perceptual and action strength for 40,000 English words. *Behav Res Methods* 2020;**52**:1271–91. <https://doi.org/10.3758/s13428-019-01316-z>
- Marks DF. Visual imagery differences in the recall of pictures. *Br J Psychol* 1973;**64**:17–24. <https://doi.org/10.1111/j.2044-8295.1973.tb01322.x>
- McInnes L, Healy J, Astels S. hdbscan: hierarchical density based clustering. *J Open Source Softw* 2017;**2**:205. <https://doi.org/10.21105/joss.00205>
- McInnes L, Healy J, Melville J. UMAP: uniform manifold approximation and projection for dimension reduction. *J Open Source Softw*. 2018;**3**:861. <https://doi.org/10.21105/joss.00861>
- Mechelli A, Price CJ, Friston KJ *et al*. Where bottom-up meets top-down: neuronal interactions during perception and imagery. *Cereb Cortex* 2004;**14**:1256–65. <https://doi.org/10.1093/cercor/bhh087>
- Merabet LB, Maguire D, Warde A *et al*. Visual hallucinations during prolonged blindfolding in sighted subjects. *J Neuroophthalmol* 2004;**24**:109–13. <https://doi.org/10.1097/00041327-200406000-00003>
- Metzger W. Optische Untersuchungen am ganzfeld. II. Zur Phänomenologie des homogenen ganzfelds. *Psychol Forsch* 1929;**13**:6–29. <https://doi.org/10.1007/BF00406757>
- Milton F, Fulford J, Dance C *et al*. Behavioral and neural signatures of visual imagery vividness extremes: Aphantasia versus Hyperphantasia. *Cereb Cortex Commun* 2021;**2**:tgab035. <https://doi.org/10.1093/texcom/tgab035>
- Muraki EJ, Speed LJ, Pexman PM. Insights into embodied cognition and mental imagery from aphantasia. *Nat Rev Psychol* 2023;**2**:591–605. <https://doi.org/10.1038/s44159-023-00221-9>
- Nichols TE, Holmes AP. Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum Brain Mapp* 2002;**15**:1–25. <https://doi.org/10.1002/hbm.1058>
- O’Callaghan D, Greene D, Carthy J *et al*. An analysis of the coherence of descriptors in topic modeling. *Expert Syst Appl* 2015;**42**:5645–57. <https://doi.org/10.1016/j.eswa.2015.02.055>

- OpenAI. *GPT-4o-Mini [Large Language Model]*. San Francisco, CA: OpenAI; 2024. <https://platform.openai.com/>.
- Pearson J. The human imagination: the cognitive neuroscience of visual mental imagery. *Nat Rev Neurosci* 2019;**20**:624–34. <https://doi.org/10.1038/s41583-019-0202-9>
- Pedregosa F, Varoquaux G, Gramfort A *et al*. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011;**12**:2825–2830. <https://inria.hal.science/hal-00650905>
- Phillips JW, Schulmann A, Hara E *et al*. A repeated molecular architecture across thalamic pathways. *Nat Neurosci* 2019;**22**:1925–35. <https://doi.org/10.1038/s41593-019-0483-3>
- Pistolas E, Wagemans J. And then there was light in the ganzfeld: clarifying the methods, experiences, and modulating factors of hallucinations and decays. *Neurosci Conscious* 2025;**2025**:niaf021. <https://doi.org/10.1093/nc/niaf021>
- Purkinje JE. *Beiträge Zur Kenntniss des Sehens in Subjectiver Hinsicht*. Johann Gottfried Calve, 1819. 176 pp. <http://archive.org/details/beitrgezurkenn00purk>
- Pütz P, Braeunig M, Wackermann J. EEG correlates of multimodal ganzfeld induced hallucinatory imagery. *Int J Psychophysiol* 2006;**61**:167–78. <https://doi.org/10.1016/j.ijpsycho.2005.09.002>
- Python Software Foundation. Python programming language. [software] 2023. <https://www.python.org/>
- R Core Team. R: A language and environment for statistical computing. [software] Vienna, Austria: R Foundation for Statistical Computing, 2023. <https://www.R-project.org/>
- Reeder R. Pseudo-hallucinations: why some people see more vivid mental images than others—test yourself here. *The Conversation*. 2021. <http://theconversation.com/pseudo-hallucinations-why-some-people-see-more-vivid-mental-images-than-others-test-yourself-here-163025>.
- Reeder RR. Ganzflicker reveals the complex relationship between visual mental imagery and pseudo-hallucinatory experiences: a replication and expansion. *Collabra: Psychology* 2022;**8**:36318. <https://doi.org/10.1525/collabra.36318>
- Reeder RR, Sala G, van Leeuwen TM. A novel model of divergent predictive perception. *Neurosci Conscious* 2024;**2024**:niae006. <https://doi.org/10.1093/nc/niae006>
- Reimers N, Gurevych I. Sentence-BERT: sentence embeddings using Siamese BERT-networks. In: *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*. Hong Kong, China: Association for Computational Linguistics; 2019. p. 3982–3992. <https://doi.org/10.18653/v1/D19-1410>
- Richardson JTE. Gender differences in the vividness of visual imagery questionnaire: a meta-analysis. *J Ment Imag* 1995;**19**:177–87.
- Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. *Proceedings of the Eighth ACM International Conference on Web Search and Data Mining* 2015;399–408. <https://doi.org/10.1145/2684822.2685324>
- Rule M, Stoffregen M, Ermentrout B. A model for the origin and properties of flicker-induced geometric phosphenes. *PLoS Comput Biol* 2011;**7**:e1002158. <https://doi.org/10.1371/journal.pcbi.1002158>
- Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 2015;**10**:e0118432. <https://doi.org/10.1371/journal.pone.0118432>
- Schmidt TT, Prein JC. The ganzfeld experience—a stably inducible altered state of consciousness: effects of different auditory homogenizations. *Psych J* 2019;**8**:66–81. <https://doi.org/10.1002/pchj.262>
- Schmidt TT, Jagannathan N, Ljubljanac M *et al*. The multimodal ganzfeld-induced altered state of consciousness induces decreased thalamo-cortical coupling. *Sci Rep* 2020;**10**:18686. <https://doi.org/10.1038/s41598-020-75019-3>
- Schwartzman DJ, Schartner M, Ador BB *et al*. Increased spontaneous EEG signal diversity during stroboscopically-induced altered states of consciousness. bioRxiv. 2019:511766. <https://doi.org/10.1101/511766>
- Shenyan O, Lisi M, Greenwood JA *et al*. Visual hallucinations induced by Ganzflicker and Ganzfeld differ in frequency, complexity, and content. *Sci Rep* 2024;**14**:2353. <https://doi.org/10.1038/s41598-024-52372-1>
- Shine JM. The thalamus integrates the macrosystems of the brain to facilitate complex, adaptive brain network dynamics. *Prog Neurobiol* 2021;**199**:101951. <https://doi.org/10.1016/j.pneurobio.2020.101951>
- Silvanto J. Interoception, insula, and agency: a predictive coding account of aphantasia. *Front Psychol* 2025;**16**:1564251. <https://doi.org/10.3389/fpsyg.2025.1564251>
- Silvanto J, Nagai Y. How interoception and the insula shape mental imagery and Aphantasia. *Brain Topogr* 2025;**38**:27. <https://doi.org/10.1007/s10548-025-01101-6>
- Spagna A, Hajhajate D, Liu J *et al*. Visual mental imagery engages the left fusiform gyrus, but not the early visual cortex: a meta-analysis of neuroimaging evidence. *Neurosci Biobehav Rev* 2021;**122**:201–17. <https://doi.org/10.1016/j.neubiorev.2020.12.029>
- Stoliker D, Preller KH, Novelli L *et al*. Neural mechanisms of psychedelic visual imagery. *Mol Psychiatry* 2025;**30**:1259–66. <https://doi.org/10.1038/s41380-024-02632-3>
- Sulik J, Rim N, Pontikes E *et al*. Differences in psychologists' cognitive traits are associated with scientific divides. *Nat Hum Behav* 2025;**9**:1147–61. <https://doi.org/10.1038/s41562-025-02153-1>
- Suzuki M, Larkum ME. General Anesthesia decouples cortical pyramidal neurons. *Cell* 2020;**180**:666–676.e13. <https://doi.org/10.1016/j.cell.2020.01.024>
- Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Methodol* 1996;**58**:267–88. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Tingley D, Yamamoto T, Hirose K *et al*. Mediation: R package for causal mediation analysis. *J Stat Softw* 2014;**59**:1–38. <https://doi.org/10.18637/jss.v059.i05>
- Vinaya H, Trott S, Pecher D *et al*. Words and worlds both: dynamic effects of distributional and sensorimotor information in semantic processing. *Open Mind* 2025;**9**:2149–74. <https://doi.org/10.1162/OPMI.a.316>
- Wackermann J, Pütz P, Allefeld C. Ganzfeld-induced hallucinatory experience, its phenomenology and cerebral electrophysiology. *Cortex* 2008;**44**:1364–78. <https://doi.org/10.1016/j.cortex.2007.05.003>
- Wicken M, Keogh R, Pearson J. The critical role of mental imagery in human emotion: insights from fear-based imagery and aphantasia. *Proc Biol Sci* 2021;**288**:20210267. <https://doi.org/10.1098/rspb.2021.0267>
- Zeman A. Aphantasia and hyperphantasia: exploring imagery vividness extremes. *Trends Cogn Sci* 2024;**28**:467–80. <https://doi.org/10.1016/j.tics.2024.02.007>
- Zeman A, Milton F, Della Sala S *et al*. Phantasia—the psychological significance of lifelong visual imagery vividness extremes. *Cortex* 2020;**130**:426–40. <https://doi.org/10.1016/j.cortex.2020.04.003>