Journal of Perceptual Imaging® 7: 1–8, 2024.

Mid-level Characteristics of Drawings Made by Observers with Aphantasia

Benjamin Balas
Psychology Department, North Dakota State University, Fargo, ND, USA
E-mail: benjamin.balas@ndsu.edu

Abstract. *Individuals with aphantasia report either absent or dramatically reduced mental imagery compared to control participants. The image of an object or scene produced "in the mind's eye" lacks detail for these individuals or is simply not there. Line drawings made from memory are a straightforward way to assess the contents of visual imagery for aphantasic individuals relative to controls. Prior analyses of the Aphantasia Drawing Database have revealed specific impairments in visual memory for objects, but relatively spared scene accuracy, suggesting that the encoding of visual scenes in aphantasia is more complex than an overall reduction in imagery might suggest. Here, we examined the mid-level image statistics of line drawings from this database to determine how simpler visual feature distributions differed as a function of aphantasia and reliance on image recall rather than direct observation during image reproduction. We find clear differences across several different sets of mid-level properties as a function of aphantasia, which offers further characterization of the nature of visual encoding in this condition.*
Keywords: *Aphantasia, drawing, mid-level vision, image statistics*

[DOI: 10.2352/J.Percept.Imaging.2024.7.000402]

1. INTRODUCTION

Aphantasia is a condition in which individuals report either an inability to engage in mental imagery or report mental imagery that is lacking in vividness and clarity [28]. Although the condition may in some instances be multi-sensory (a subset of individuals with aphantasia report a lack of imagery across sensory modalities [8]), it has mostly been examined in the context of just one or a small subset of senses [5] with visual imagery being especially prominent [18]. Aphantasia can be assessed using subjective reports like the Vividness of Visual Imagery Questionnaire [21, 22], but recent research has focused on differences between individuals with aphantasia and those without on a variety of cognitive tasks. The goal of the majority of these studies is to establish objective perceptual and/or cognitive indices of the condition. For example, binocular rivalry is not affected by visual imagery instructions that do serve as a visual prime for control participants [18]. The absence of this priming effect suggests that individuals with aphantasia do not necessarily maintain a sufficiently vivid internal image of the prime stimulus to affect subsequent performance. In mental rotation tasks, individuals with severe aphantasia do not show the same effect of rotation distance on response time that control participants do, an effect that is typically interpreted as a reflection of some form of mental imagery used to manipulate an internal image of the target object(s) [23]. Despite this, individuals with aphantasia tend to perform as well as control participants in mental rotation tasks and other putatively visuospatial or working memory tasks that are assumed to depend on imagery, suggesting the adoption of some alternate strategy that does not depend on a pictorial representation of a target object that is absent [15, 20, 27]. In terms of physiological correlates of visual perception and visual imagery, individuals with self-reported aphantasia or reduced visual imagery as assessed by self-report also differ from controls in a manner consistent with objectively reduced or absent visual imagery. Kay et al. [17], for example, report that being asked to imagine achromatic patterns that are either bright or dark induces pupillary responses consistent with actually seeing these images if participants report vivid visual imagery, but this is reduced in participants reporting reduced vividness. During mental imagery, Zeman et al. also report that an aphantasic participant exhibited increased activity in frontal regions relative to control participants, coupled with reduced activity in parietal regions [27]. The parietal lobe contributes substantially to visuospatial processing, so reduced activity in this region is consistent with the absence of visual imagery and/or transformations of an internal image. Cortical excitability in areas V1–V3 is also lower as a function of imagery strength, as indexed using the binocular rivalry imagery paradigm in which instructions to imagine one member of the rivalrous stimulus pair affect subsequent rates of rivalry [19]. More broadly, imagery vividness in general correlates with the overlap between perceptual activity and neural activity during imagery, translating to a specific neural basis for imagery vividness in individuals with aphantasia and the general population [9]. Together, these results indicate that aphantasia does indeed reflect differences in visuospatial imagery between control participants and individuals subjectively reporting severely reduced visual imagery, both in terms of behavioral performance and neural responses.

1.1 *Aphantasia and Drawings*

Having established that reported aphantasia (a condition that may affect approximately 1 in 20 people to some degree [6]) does correspond to profound differences in visual imagery in terms of the maintenance of an internal image,

Received July 13, 2023; accepted for publication May 2, 2024; published online June 25, 2024. Associate Editor: Thrasyvoulos N. Pappas.
2575-8144/2024/7/000402/8/$00.00

it is natural to wonder about the contents of that internal image (if any) and how these differ across observers. We would of course assume that the fidelity of any internal image maintained by an individual with aphantasia would be greatly reduced compared to a typical participant, and that assumption is the basis of some brief assessments of aphantasic experience. In popular assessments of aphantasia offered online (websites including WikiHow, BuzzFeed, and other similar sites include such short quizzes) as a sort of personality or ability test, participants are asked to select an image that best approximates their internal image of an object that they imagine in their mind's eye. An apple, for example, may be depicted in these assessments as having color that varies in its saturation or an external contour that varies according to its sharpness. Though not especially rigorous, these popular assessments speak to the underlying subjective phenomenon that is both the basis of the aphantasic experience and also the most elusive: What does one see when one imagines something, and how is this experience different for an individual with aphantasia?

One straightforward way to address this fundamental question is to simply ask individuals to report the contents of their internal image by making a record of it. Unlike standard forced-choice experiments or reaction time assessments of perceptual and cognitive performance, drawing tasks provide participants with an opportunity to directly report the contents of their memory [1] or their imagery. The resulting drawings are also a rich source of data that can be analyzed in a variety of ways [11]. Thus, if we are curious about the nature of internal images in aphantasia, why not ask individuals with aphantasia to draw what they can imagine as best as they can and compare their drawings to those by control participants? Bainbridge et al. [2] did exactly this, yielding a database comprising drawings made both from memory and from direct observation of a natural scene by individuals with aphantasia and control participants. In their analysis of high-level object and scene content, they report that the drawings of individuals with aphantasia revealed reduced object memory relative to controls but high spatial accuracy and few memory errors. These results indicate that the difference between an aphantasic individual's mental representation of a recalled scene and a control participant's representation of the same is nuanced—it is not the case that the aphantasic record is simply worse, but it is lacking in some specific aspects of scene structure. Moreover, the lack of any differences between groups when direct observation of images was permitted during drawing suggests that aphantasia does not affect the perceptual and motor processes that contribute to drawing itself. Instead, drawings in this study appear to reveal differences in the stored record of scene appearance observers have access to during recall when they attempt to reproduce the visual features present in complex scenes.

1.2 *Do Mid-level Features Differ in Aphantasic Drawings?*

In the current study, we examined another aspect of the drawings made by aphantasic individuals and control participants to further understand potential differences in the nature of visual encoding and visual imagery between these groups. Specifically, we chose to examine mid-level properties of the drawings made by these participants. By mid-level properties, we refer to aspects of visual structure that are neither as complex as nameable objects, materials, or scene labels nor as basic as color and intensity distributions or local measurements of contrast. To say more about what mid-level properties are rather than focus on what they are not, these are aspects of visual structure that include distributions of curvature, junction types, contour lengths, and other features that require some integration and/or organization of oriented contrast measurements across position. Our goal was to extend the analysis of the Aphantasia Drawing Database to test the hypothesis that aphantasic drawings may not just differ in terms of high-level object representations but also in terms of the fidelity of more basic visual features that reflect the encoding of scene structure at a finer grain.

To expand upon our motivation for using mid-level features as the focus of our analysis, it is not the case that we anticipate individuals with aphantasia will have some deficit in seeing or reproducing specific junction types, line curvatures, or line segment orientations. Instead, we chose to examine mid-level features in these drawings because these features are an important bridge between low-level visual features and high-level structure like recognizable objects, scene layout, and textures. In particular, the features we focus on here are important indicators of a range of relationships between 3D scene layout and the 2D pictorial projection of the same. For example, different types of junctions (T versus Y versus X) signal different depth relationships in line drawings [13]—the presence or absence of these junctions may indicate the extent to which these depth relationships are available to individuals with low vividness of visual imagery. Similarly, contour length can be used as a proxy for shape smoothness or an index of the amount of complexity [10], or fine textural detail included in a drawing. In this case, shifts in the distribution of contour length could indicate biases for visual structure at coarse versus fine spatial scales, independent of specific nameable objects or high-level scene features. Finally, both contour curvature and contour orientation are interesting to consider in terms of natural image statistics and potential biases in perception and imagery due to statistical regularities in the visual world [29]. Do observers with aphantasia rely more heavily on prior expectations of orientation distribution than other observers, for example, reproducing vertical and horizontal orientations in an even greater disproportion relative to oblique orientations? Likewise, are high-curvature contours less evident in the drawings made by aphantasic observers given an over-reliance on contour smoothness as a guiding statistical principle that dictates image reconstruction in the absence of vivid imagery? In each case, these mid-level features offer a specific opportunity to examine an aspect of visual structure that is not captured by analysis of high-level object and texture inclusion, nor adequately measured by

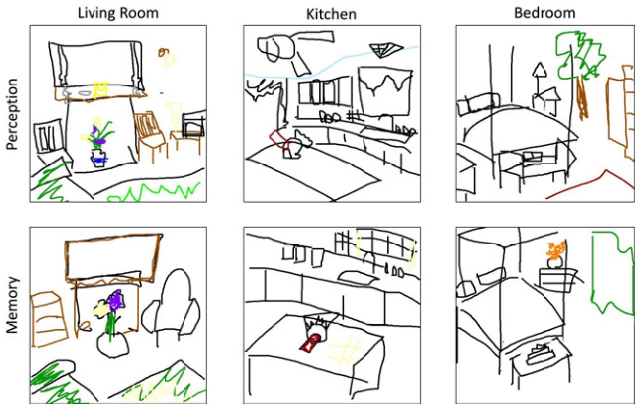

Figure 1. Examples of drawings depicting each of the three indoor scenes from the Aphantasia Drawing Database by a participant with typical visual imagery. The top row depicts drawings made while looking at the target image for reference and the bottom row depicts drawings made from memory.

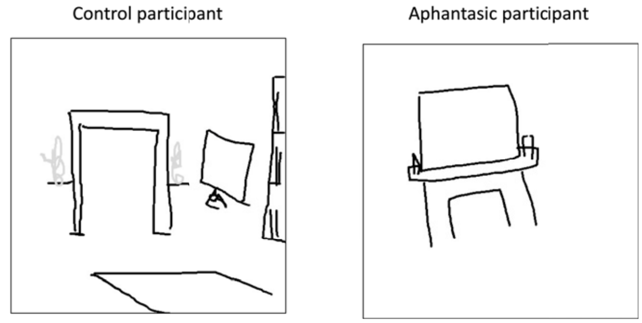

Figure 2. A drawing of the living room scene made from memory by a control participant (left) and a participant with aphantasia (right). Both contain some elements of the original scene, but multiple objects are missing from the aphantasic individual's drawing.

simple measures of local oriented contrast. While mid-level features are not independent of either low-level or high-level features, they also do provide unique information about visual structure at a specific level of granularity [29] that we think may yield insights into how drawings are made in the absence of strong visual imagery.

To achieve this goal, we applied tools for line drawing analysis from the Mid-Level Vision (MLV) Toolbox [26], an open-source Matlab Toolbox that supports measurement of a wide variety of low- to mid-level image features. For each set of candidate image statistics, we compared feature distributions extracted from the drawings of individuals with aphantasia and control participants to determine the extent to which these measurements were separable by group when drawings were made from memory or by direct observation. Briefly, we find that aphantasic individuals do represent mid-level structure differently from control participants in their recalled drawings, suggesting that the contents of their visual memory are limited in terms of simple image structures as well as high-level representations of objects.

2. METHODS

2.1 *Image Database*

The Aphantasia Drawing Database consists of color line drawings of natural scenes made by individuals with aphantasia ($N = 63$) and control participants ($N = 52$) with typical visual imagery. Each individual was asked to view three images depicting natural indoor scenes (a bedroom, a kitchen, and a living room) and make one set of drawings (one per scene) while using the photographs of these scenes as a reference (what we will call the *perception* condition) and another set of drawings from memory (what we will call the *recall* condition). We display examples of *perception* and *recall* drawings of each scene completed by the same participant in Figure 1.

Each original drawing is 500×500 pixels in size, and we rendered all images in two-tone black and white for our analyses of mid-level image statistics. We also removed any participant who did not have complete drawings for all three scenes in both conditions, leaving us with a final sample of 56 complete sets of drawings by individuals with aphantasia and 48 complete sets of drawings by individuals with typical visual imagery. In the manuscript introducing the Aphantasia Drawing Database, Bainbridge et al. [2] observed that the drawings made from recall by participants with aphantasia lacked object detail that was present in the drawings made by control participants, as assessed by human raters who coded each drawing for specific visual content. Indeed, viewing images drawn from memory by a participant with aphantasia alongside similar images drawn from memory by a control participant makes this difference clear: The drawing by the individual with aphantasia is overall more sparse and lacks details that are present in the control drawing (Figure 2). Our goal is to characterize this discrepancy and potentially identify others in terms of differences in mid-level visual features.

2.2 *Mid-level Feature Extraction*

In order to test our hypothesis that the drawings made by individuals with aphantasia would differ from control drawings in terms of mid-level feature distributions, we analyzed the line drawings from the Aphantasia Drawing Database using the MLV Toolbox for Matlab [26]. This toolbox contains multiple functions for measuring a wide range of low- to mid-level features in photographic images and line drawings. To apply these functions, each original image in the database must first be converted into a vector line drawing that is used as the basis for extracting the candidate feature sets we will describe in what follows. In the case of images like these that are already composed of well-defined lines on a plain background, the resulting images retain nearly all of the structure present in the original drawing and can be generated relatively quickly. Our routines for batch-processing the images in the database to turn them into MLV line drawings are available on OSF at the following link: https://osf.io/3fxd6/.

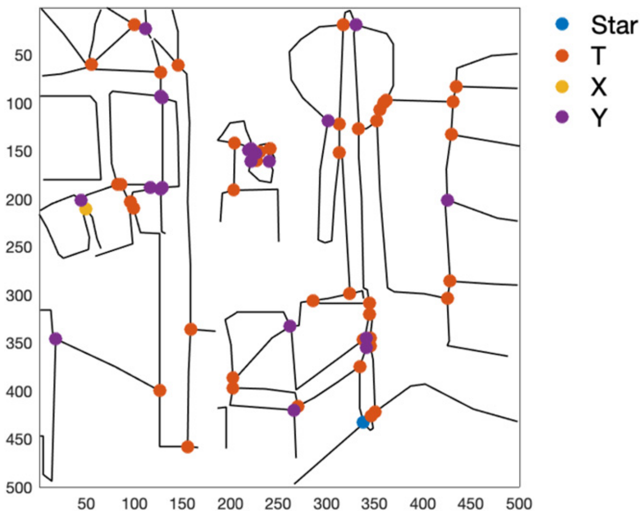

Figure 3. An example line drawing from the Aphantasia Drawing Database with junction type coded by the MLV toolbox. Note that while T, X, and Y junctions are generally plentiful, there are very few Star and Arrow junctions.

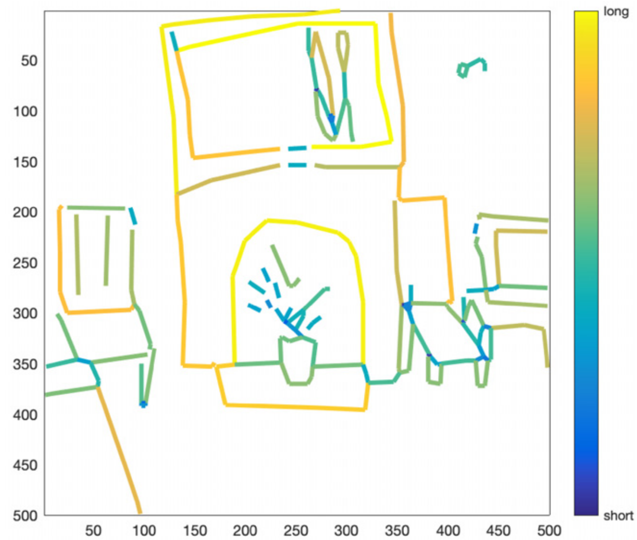

Figure 4. An example line drawing from the Aphantasia Drawing Database with line segment length coded by the MLV toolbox.

2.3 *Target Image Statistics*

We used the MLV toolbox to measure and compare four families of image statistics across *perception* and *recall* drawings made by both participant groups. In what follows, we describe each family of image statistics briefly and provide a visualization of a line drawing with these features extracted by the relevant toolbox functions.

3. JUNCTION TYPES

We used the MLV toolbox to measure the frequency of different junction types in each drawing. In general, a junction refers to a feature where line segments meet or cross and these are classified by the number of segments that contribute to the junction and the relative orientation of the segments. The MLV toolbox supports the extraction of T, Y, X, Arrow, and Star junctions, all of which we opted to measure in the database images. In Figure 3, we display an example drawing from the database alongside the MLV coding of junction types across this image. We measured junction frequency collapses across orientation, yielding five bins per image (one per junction type).

3.1 *Line Segment Length*

The MLV also supports the measurement of line segment length across the image. We measured the frequency of line segment length using eight logarithmically spaced bins, spanning a range from 2 pixels to twice the width of the image. In Figure 4, we display an example drawing from the database color-coded by segment length using the MLV.

3.2 *Contour Curvature*

We also extracted the distribution of contour curvature in participants' drawings using the MLV toolbox. Curvature as defined in the toolbox ranges from 0 to 90 degrees, and we measured curvature distribution using eight logarithmically spaced bins spanning this range, weighted by line segment length. In Figure 5, we display an example image with curvature color-coded per segment.

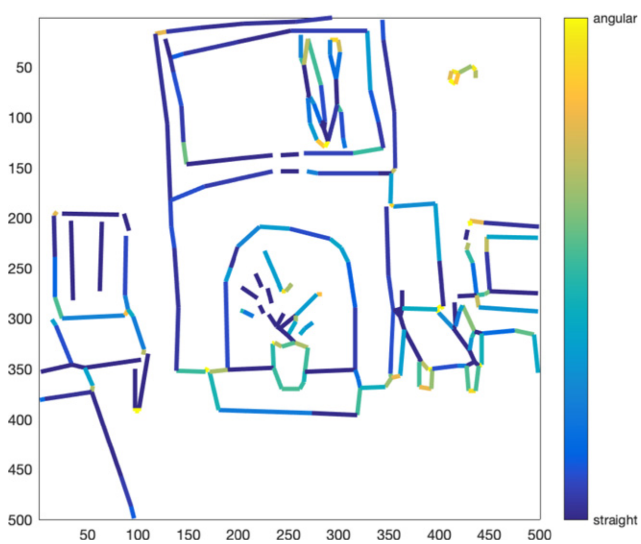

Figure 5. An example line drawing from the Aphantasia Drawing Database with curvature coded by the MLV toolbox. Most segments are straight or nearly so, with higher curvature values being increasingly infrequent.

3.3 *Line Segment Orientation*

Finally, we also extracted orientation histograms per image, weighted by line segment length. These histograms were computed using eight uniformly spaced bins.

3.4 *Comparing Feature Distributions*

In order to determine whether there were differences between the distribution of each of these feature types across participant groups in the *perception* and *recall* conditions, we adopted a classification analysis based on the histogram values recorded for each participant's images. Per image and per individual, we first converted the raw frequency histogram into a probability distribution by dividing each value in the histogram by the sum of all values across bins. We took this step because the sparsity of *recall* images produced by aphantasic individuals relative to control participants was bound to yield raw feature counts that were lower than those of control participants, an effect that is not of interest to us. Instead, we are interested in differences in the shape of the distribution across participant groups. These differences are retained in the probability distribution while baseline differences in the overall number of features present in different images are removed.

Having calculated these probability distributions per image and per participant for each set of candidate mid-level statistics, we used a pattern classification analysis to determine how separable aphantasic and control participant feature distributions were. Specifically, we measured the classification error for labeling each feature distribution as "Aphantasia" or "Control" in the *perception* and *recall* conditions separately using k-nearest neighbors classification. A low error rate thus reflects strong separability of the feature distributions across participant groups (substantial differences between aphantasic and control individuals' drawings) while a high error rate reflects weak separability and little difference across participant groups. A similar analysis was reported in [7] to compare abstract depictions of emotion categories in terms of similar sets of mid-level features. In each analysis, we compare the observed classification accuracy to the distribution of values obtained from a permutation analysis in which we scrambled the labels of our data points 10,000 times and measured classification accuracy with these shuffled labels. We used a 99th percentile criterion for considering the measured outcome to be significantly better than chance meaning that the observed value (or higher) is only observed in fewer than 1% of the bootstrapped values.

An advantage of this procedure is that it obviates the need to consider some potential complicating factors in our data. For example, the low overall occurrence of some features (e.g. Arrow and Star junctions, or highly curved contours) could induce a sort of "floor effect" that could yield highly significant statistical interactions as assessed by tools like ANOVA, but that would not be especially meaningful. Also, although some of our feature distributions could be compared across groups by fitting a low-dimensional function to the data (e.g. a power-law model of curvature), for other features it is not at all obvious how to do this (e.g. junction types). By comparison, this analysis allows us to remain agnostic about the underlying form of the data and characterize differences between our two participant groups across conditions via pattern recognition. A key limitation of the method, however, is that we are not able to examine interactions across features: We cannot comment on the distribution of short line segments with high curvature, for example, or other conjunctions one could imagine across feature classes. Although such an analysis may yield interesting differences across groups, we have chosen not to pursue such conjunctions for two reasons: (1) there is a combinatorial problem in terms of the number of possible conjunctions and comparisons we could make, which is not easy to address without strong prior hypotheses to constrain the analysis; (2) there is potential for very sparse data in some of these joint bins of feature counts. For now, we instead focus on examining each feature class in isolation from the others.

4. RESULTS

For each set of feature distributions, we carried out our classification analysis using the k-nearest neighbors algorithm implemented in the JASP v3.0 Machine Learning module [16]. Each of the drawings made by all participants was included in the full data set for both our *perception* and *recall* analyses. We used a 50/50 split of the data to randomly assign data points to training versus test samples and held out 20% of the training data for validation. The number of nearest neighbors used for classification was optimized per analysis but capped at a maximum value of 10 neighbors.

With regard to the drawings made in the *perception* condition, we found that *Junction Types, Contour Curvature,* and *Line Segment Orientation* did not support robust classification according to participant group. The distribution of *Junction Type* features supported a test set accuracy of 52.6%, *Contour Curvature* features yielded a test set accuracy of 54.5%, and *Line Segment Orientation* resulted in a test set orientation of 51.3%. In all three of these cases, comparison against the distribution of values obtained from our permutation analysis suggested that these accuracies did not differ significantly from what we might expect by chance (50th, 80th, and 64th percentiles, respectively). We did find, however, that the distribution of *Line Segment Length* did support reliable classification of drawings by group membership with a test accuracy of 78.8%. Compared to the distribution of values obtained from our classification analysis, this accuracy was well above our 97.5th percentile cut-off.

The data from the *recall* condition was very similar to the results reported previously for the *perception* condition. In both cases, *Junction Types, Contour Curvature,* and *Line Segment Orientation* did not support robust classification according to participant group. The distribution of *Junction Type* features yielded a test set accuracy of 51.9%, *Contour Curvature* features yielded a test set accuracy of 58.3%, and *Line Segment Orientation* features yielded a test set orientation of 56.4%. Again, none of these values reached our percentile threshold determined from our permutation analysis, though Contour Curvature did reach the 90th percentile and Line Segment Orientation reached the 95th. Though these results are suggestive of some systematic differences between our participant groups, we do not

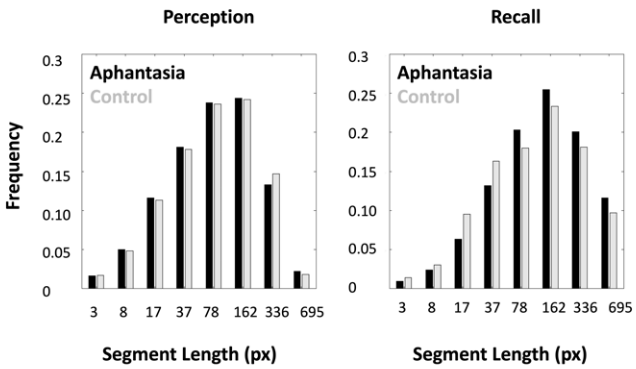

Figure 6. Average feature distributions for individuals with aphantasia and control participants in the *perception* and *recall* conditions. Although both distributions support group classification, the group difference is more evident in the *recall* condition.

Table I. Full set of classification results across *perception* and *recall* conditions for each of our four mid-level features. Asterisks indicate outcomes different from chance according to a permutation test analysis.

Feature Set	Perception	Recall
Junction Types	52.6%	51.9%
Contour Curvature	54.5%	58.3%
Line Orientation	51.3%	56.4%
Line Length	78.8%*	89.1%*

discuss them further as they did not reach our criterion for significance. By comparison, the distribution of *Line Segment Length* supported much more robust classification of aphantasic versus control drawings: We observed a test set accuracy of 89.1%, which was far above our 99th percentile cut-off based on permutation test analysis. In Figure 6, we display the average distribution of *Line Segment Length* values in both the *perception* and *recall* conditions for both aphantasia and control participants.

Although we find that both distributions of line segment length allow for successful categorization of drawings by participant group, we see in the classification accuracy results and the plots in Fig. 6 that *recall* condition separability is particularly good. Indeed, the distributions of values in the *perception* condition are extremely close between our two participant groups, with only some subtle differences making accurate classification more likely. By comparison, the *recall* condition distributions are more clearly different and indicate a bias for individuals with aphantasia to include more longer line segments than shorter lines in their drawings compared to control participants. In Table I, we summarize the classification results from all feature sets in the *perception* and *recall* tasks.

We chose to continue with an exploratory classification analysis of the separability of drawings from the *perception* and *recall* tasks on the basis of the different mid-level feature distributions described earlier. For this analysis,

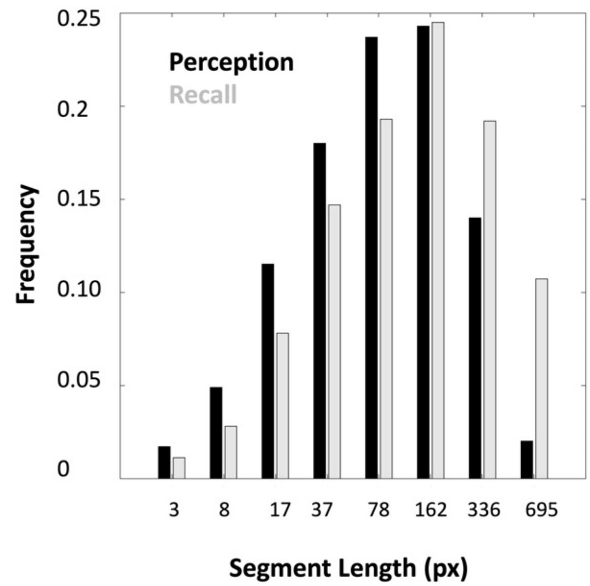

Figure 7. Average feature distributions of segment length for drawings made in the *perception* and *recall* conditions, collapsed across participant groups.

we combined features from individuals with aphantasia and control individuals to focus on differences in line drawings resulting from direct observation versus memory. We used the same k-nearest neighbors parameters described previously and conducted separate analyses for each set of features. Like we observed in the previous analysis, we found that classification accuracy was much better when *Line Segment Length* distributions were used than any of the other feature sets. The distribution of *Junction Type* features yielded a test set accuracy of 59.3%, *Contour Curvature* features yielded a test set accuracy of 64.4%, and *Line Segment Orientation* features led to a test set accuracy of 62.2%. By comparison, accuracy with *Line Segment Length* led to a test set accuracy of 84.0%. We display the average distribution of line segment length measurements in both the *perception* and *recall* tasks in Figure 7.

Similar to the difference between individuals with aphantasia and control participants, drawings made in the *recall* condition tend to differ from those in the *perception* condition due to a bias favoring longer line segments in *recall* drawings. A preliminary analysis indicated that this was the case for both individuals with aphantasia and control participants, but due to the smaller sample size involved in these within-group comparisons, we are potentially unable to measure differences in classification accuracy robustly in this case.

5. DISCUSSION

Our results demonstrate that there are differences in mid-level image features found in drawings of natural scenes that depend on observers' ability to engage in mental imagery. We also found that there are similar differences that are evident when comparing images made while directly observing a target scene to those made when attempting to recall that scene from memory. A unifying perspective on these two outcomes is that both the presence of aphantasia and the requirement to draw a scene from memory lead to relatively impoverished internal images. What we think is particularly interesting is that the lower fidelity of the internal images maintained under these conditions is only expressed robustly through one class of mid-level statistics considered here (namely, *Line Segment Length* distributions) and not through the others. This means that the previously reported effect of aphantasia on high-level features like object detail [2] does not manifest in some uniform cost to mid-level or low-level features that are associated with that detail. To put it another way, the fact that individuals with aphantasia tend to make sparser drawings from memory than control participants does not impact all simpler visual features in the image. Instead, the vividness of visual imagery only affects the expression of a limited class of mid-level structures.

First, why should *Line Segment Length* statistics be impacted in the way that we have observed as a function of aphantasia and the imposition of a recall constraint? The bias that we observed in both conditions that we are associating with a weaker internal image favored longer line segments over shorter line segments, which could be the result of a number of pictorial differences. Shorter line segments tend to contribute to texture patterns, small-scale objects, and surface details, or could be used to render complex external outlines of objects or surfaces via piecewise drawing of an extended contour. One feature of the Aphantasia Drawing Database that is interesting to consider here is the use of text to label objects in scenes (e.g. putting the word "table" inside an otherwise indistinct shape), which was more prevalent in drawings by individuals with aphantasia. This practice should inflate short line segments relative to long ones, but obviously does not do so to a sufficient degree to overcome this more pronounced bias. Instead, the bias favoring longer contours in aphantasic drawings and drawings made from memory likely reflects the lack of surface detail in particular. Piecewise drawing of complex contours would likely result in orientation histogram differences that we do not observe in our data, leaving the presence (or relative absence) of surface detail as the main contributing factor. We note that this need not be directly tied to object detail, but may be a more general feature of how surfaces are rendered under these different conditions.

Besides asking why one of our feature classes is impacted by these conditions, we can also ask why the others are not. Though the other classes of mid-level features were better at separating *perception* drawings from *recall* drawings than at separating aphantasic versus control drawings, accuracy in the former classification analysis was still fairly low. One possibility that to our knowledge has not been explored is that some of the feature distributions we considered here may be lawful in the sense that constraints on the nature of line drawings lead to a narrow space of possible shapes for these distributions. By analogy, we are proposing something akin to the known properties of photographs of natural scenes like the approximate $1/f$ power spectrum of natural images [12] or the distribution of cardinal orientations relative to oblique orientations in natural scenes [3]. The former property of natural scenes is a consequence of the scaling of 3D shapes and surfaces at different positions when projected into two dimensions [24], while the latter is a function of gravity favoring contours either parallel to the ground plane or perpendicular to it [4]. Any lawful properties of line drawings would not be subject to the same kind of physical and optical constraints but might instead reflect motor constraints or biases on drawing. A broader examination of how these mid-level statistics may vary across other types of drawings or how they may vary developmentally may yield important insights into how curvature, orientation, and junctions are typically expressed in line drawings of complex scenes. Though there have been several recent discussions of why line drawings "work" perceptually [14], what we are suggesting is a close analysis of how line drawings tend to be made by both trained and naïve participants. As Sayim and Cavanaugh observe [25], there have been surprisingly few changes in some aspects of line drawing production over very long periods of human history, which may indicate the presence of something like the natural modes in line drawing statistics we are suggesting as the basis for the null results in these different feature classes.

Our analyses do have some limitations that are also important to consider alongside our discussion of these interesting aspects of our results. One aspect of the database that we did not include as an additional factor is the different images created by each participant of the three different indoor scenes. The advantage of the aggregate analysis presented here is that we end up with thrice the number of data points, but the downside is that there is within-participant variability we are not considering. On one hand, there may be interesting individual variations in how some of these mid-level features are expressed, which could potentially be examined in a mixed-model analysis that includes image as a factor. To the extent that individual participants do draw with a unique style that is evident in the feature distributions we considered here, this could inflate our classification accuracy somewhat given that other drawings made by the same participant may be particularly useful nearest neighbors for classification. There are also parameter values regarding the normalization of these different features by contour length, the granularity of feature histograms, and other properties of mid-level feature measurement that we did not explore in depth here. Although we do not anticipate that our results would be sensitive to these parameters, it is potentially worth investigating the robustness of our results along these lines. Finally, our choice of classifier was motivated largely by a desire for simplicity rather than maximizing our classification accuracy. It is possible that discriminant methods, including SVMs, could yield higher classification rates for our other feature distributions. Again, we doubt that this is likely to be the case, but we have not explored this possibility in any depth. Nonetheless, despite these avenues for further refinement of our approach, we think that the current study offers new and interesting insights into how impoverished internal images do and do not lead to effects on mid-level visual statistics.

6. CONCLUSIONS

Attempting to reproduce natural scenes via line drawings with an internal image that is not especially vivid leads to differences in the mid-level statistics of line length favoring longer line segments. This result indicates that both aphantasia and drawing from memory lead to reduced fidelity of surface detail and other small-scale features in complex scenes. The absence of effects on other aspects of mid-level visual statistics may suggest either that these properties of natural scenes are robust to the strength of an internal image or that these features are constrained in some manner by the nature of line drawing as a perceptual task and/or a motor task. Examining mid-level characteristics of line drawings more broadly may reveal lawful properties of line drawings or other differences in the structure of line drawings made under different constraints and of different kinds of natural objects, surfaces, and scenes.

REFERENCES

[1] W. A. Bainbridge, W. Y. Kwok, and C. I. Baker, "Disrupted object-scene semantics boost scene recall but diminish object recall in drawings from memory," Mem. Cogn. **49**, 1568–1582 (2021).

[2] W. A. Bainbridge, Z. Pounder, A. F. Eardley, and C. I. Baker, "Quantifying Aphantasia through drawing: Those without visual imagery show deficits in object but not spatial memory," Cortex **135**, 159–172 (2021).

[3] D. Coppola, H. R. Purves, A. N. McCoy, and D. Purves, "The distribution of oriented contours in the real world," Proc. Natl. Acad. Sci. USA **95**, 4002–4006 (1998).

[4] C. J. Dakin and A. Rosenberg, "Gravity estimation and verticality perception," Handb. Clin. Neurol. **159**, 43–59 (2018).

[5] C. J. Dance, J. Ward, and J. Simner, "What is the link between mental imagery and sensory sensitivity? Insights from aphantasia," Perception **50**, 757–782 (2021).

[6] C. J. Dance, A. Ipser, and J. Simner, "The prevalence of aphantasia (imagery weakness) in the general population," Consciousness Cogn. **97**, 103243 (2022).

[7] C. Damiano, P. Gayen, M. Rezanejad, A. Banerjee, G. Banik, P. Patnaik, J. Wagemans, and D. B. Walther, "Anger is red, sadness is blue: Emotion depictions in abstract visual art by artists and non-artists," J. Vis. **23**, 1 (2023).

[8] A. J. Dawes, R. Keogh, T. Andrillon, and J. Pearson, "A cognitive profile of multi-sensory imagery, memory and dreaming in aphantasia," Sci. Rep. **10**, 10022 (2020).

[9] N. Dijkstra, S. E. Bosch, and M. A. van Gerven, "Vividness of visual imagery depends on the neural overlap with perception in visual areas," J. Neurosci. the Official J. Soc. Neurosci. **37**, 1367–1373 (2017).

[10] D. C. Donderi, "Visual complexity: A review," Psychol. Bull. **132**, 73 (2006).

[11] J. E. Fan, W. A. Bainbridge, R. Chamberlain, and J. D. Wammes, "Drawing as a versatile cognitive tool," Nat. Rev. Psychol. **2**, 556–568 (2023).

[12] D. J. Field and N. Brady, "Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes," Vis. Res. **37**, 3367–3383 (1997).

[13] A. Guzman, "*Computer Recognition of Three-dimensional Objects in a Visual Scene*" Ph.D. Dissertation, (Department of Electrical Engineering, MIT, 1968).

[14] A. Hertzmann, "Why Do Line Drawings Work? A Realism Hypothesis," Perception **49**, 439–451 (2020).

[15] C. Jacobs, D. S. Schwarzkopf, and J. Silvanto, "Visual working memory performance in aphantasia," Cortex; A J. Devoted Study Nervous Syst. Behav. **105**, 61–73 (2018).

[16] JASP Team JASP (Version 0.17.2) [Computer software] (2023).

[17] L. Kay, R. Keogh, T. Andrillon, and J. Pearson, "The pupillary light response as a physiological index of aphantasia, sensory and phenomenological imagery strength," eLife **11**, e72484 (2022).

[18] R. Keogh and J. Pearson, "The blind mind: No sensory visual imagery in aphantasia," Cortex; A J. Devoted Study Nervous Syst. Behav. **105**, 53–60 (2018).

[19] R. Keogh, J. Bergmann, and J. Pearson, "Cortical excitability controls the strength of mental imagery," eLife **9**, e50232 (2020).

[20] R. Keogh, M. Wicken, and J. Pearson, "Visual working memory in aphantasia: Retained accuracy and capacity with a different strategy," Cortex; A J. Devoted Study Nervous Syst. Behav. **143**, 237–253 (2021).

[21] D. F. Marks, "Visual imagery differences in the recall of pictures," Br. J. Psychol. **64**, 17–24 (1973).

[22] D. F. Marks, "New directions for mental imagery research," J. Mental Imagery **19**, 153–167 (1995).

[23] Z. Pounder, J. Jacob, S. Evans, C. Loveday, A. Eardley, and J. Silvanto, "Only minimal differences between individuals with congenital aphantasia and those with typical imagery on neuropsychological tasks that involve imagery," Cortex; A J. Devoted Study Nervous Syst. Behav. **148**, 180–192 (2022).

[24] D. L. Ruderman, "Origins of scaling in natural images," Vis. Res. **37**, 3385–3398 (1997).

[25] B. Sayim and P. Cavanagh, "What line drawings reveal about the visual brain," Front. Hum. Neurosci. **5**, 118 (2011).

[26] D. B. Walther, D. Farzanfar, S. Han, and M. Rezanejad, "The mid-level vision toolbox for computing structural properties of real-world images," Front. Psychol. **14**, 1322 (2023).

[27] A. Z. Zeman, S. Della Sala, L. A. Torrens, V. E. Gountouna, D. J. McGonigle, and R. H. Logie, "Loss of imagery phenomenology with intact visuo-spatial task performance: a case of 'blind imagination'," Neuropsychologia **48**, 145–155 (2010).

[28] A. Zeman, M. Dewar, and S. D. Sala, "Lives without imagery - Congenital aphantasia," Cortex; A J. Devoted Study Nervous Syst. Behav. **73**, 378–380 (2015).

[29] C. Zetzsche, "Natural scene statistics and salient visual features," Neurobiol. Attention **37**, 226–232 (2005).