Visual interference can help and hinder memory: Measuring memory fidelity using a novel circular shape space

Aedan Y. Li[1], Jackson Liang[1], Andy C. H. Lee[1,2], Morgan D. Barense[1,2]

[1]Department of Psychology, University of Toronto

[2]Rotman Research Institute at Baycrest Hospital

Corresponding author:
Aedan Y. Li
Department of Psychology
University of Toronto
100 St. George Street
Toronto, ON, Canada
M5S 3G3
aedanyue.li@mail.utoronto.ca
(416) 978-5464

Word Count: 11,913

Abstract

Previous work suggests that the similarity of distracting information differentially alters how memories are forgotten. Though these effects have been shown for color memory, it is unclear if they extend to other modalities such as shape. In a first experiment, we created the *Validated Circular Shape Space* (VCS space), the first perceptually uniform "Shape Wheel" whereby subjective similarity was empirically quantified by angular distance on a 2D circle. In a second experiment, we utilized VCS space to examine the impact of distractor similarity on shape memory. A converging set of mixture model and model-free analyses revealed that distractor similarity differentially impacted memory detail. Dissimilar distractors disrupted both fine- and coarse-grained information, rendering the memory inaccessible. In contrast, similar distractors disrupted fine-grained but increased the reliance on coarse-grained information, rendering the memory blurred. As these effects were consistent across domains for both color and shape, we suggest these findings provide converging evidence for a set of general principles regarding how the nature of distracting information impacts memory. More generally, VCS space can be used to not only quantify subjective similarity but to also quantify memory fidelity for shape, a currently underexplored feature type in memory experiments.

Keywords: Subjective similarity, interference, perceptually uniform shape space, memory fidelity, mixture model

Acknowledgements

We thank Sol Z. Sun and Bryan Hong for helpful discussion, Tim Brady for helpful discussion and his suggestion to use multidimensional scaling for validation, Keisuke Fukuda for commenting on an initial draft of this manuscript, and Elizabeth Page-Gould for statistical assistance. AYL is supported by a NSERC grant. This work is supported by a Scholar Award from the James S McDonnell Foundation, an Early Research Award from the Ontario Government, and a NSERC Discovery grant and Accelerator supplement to MDB. MDB is supported by the Canada Research Chairs program. This work is also supported by an NSERC Discovery grant to ACHL.

Earlier versions of this work were presented at the Lake Ontario Visionary Establishment in 2016, the Canadian Society for Brain, Behaviour and Cognitive Science in 2017, the Vision Sciences Society in 2017 and 2018, the Toronto Area Memory Group in 2018, and the Memory Disorders Research Society in 2018.

Visual interference can help and hinder memory: Measuring memory fidelity using a novel circular shape space

Certain items are more similar than others. For example, we can agree that a computer is more similar to a television than to a dog. This similarity judgment is made on the basis of a variety of characteristics, including perceptual (Wardle et al., 2016) and semantic properties (Jozwik, Kriegeskorte, & Mur, 2016). Here, the tendency for observers to group items together on the basis of some characteristic or set of characteristics is known as *subjective similarity*. An extension to the construct of subjective similarity is *subjective variability*, whereby the similarity within sets of items can range from low to high. For example, the change in subjective similarity within one set (e.g. laptop, computer, television, headset) can be less variable compared to another set (e.g. pen, dog, sofa, water bottle).

Depending on the task requirements, the similarity of distracting information can have different and sometimes opposing effects on memory. In classic list-learning paradigms, memory failure is more likely to occur when distracting items are similar to the target item (Baddeley, 1964; Baddeley & Dale, 1966; Gordon, Hendrick, & Levine, 2002; Logie, Della, Wynn, & Baddeley, 2000; Wickelgren, 1965; Wickelgren, 1966; Wickens, 1970). Supporting these results, experiments using complex objects find greater memory and discrimination failures as the similarity between objects increase (Barense et al., 2012; Yeung et al., 2013; Watson & Lee, 2013). In other cases, high similarity can instead benefit memory performance, an effect that holds for stimuli ranging from simple lines to complex faces (Jiang et al., 2016; Kahana, Zhou, Geller, & Sekuler, 2007; Lin & Luck, 2009; Mate & Baques, 2009; Morey, 2018; Sanocki & Sulman, 2011; Sims, Jacobs, & Knill, 2012). To reconcile this seemingly contradictory set of findings, distractor similarity has been proposed to differentially impact memory detail (Sun,

Fidalgo, et al, 2017). Instead of conceptualizing memory as an all-or-nothing process whereby items are only ever remembered or forgotten, a more accurate and sensitive measure of performance may be provided by capturing memory along a continuous metric. Indeed, emerging behavioral, computational, and neuroimaging evidence suggest that not only can memories range along a continuum of detail (Bays, Catalao, & Husain, 2009; Ma, Husain, & Bays, 2014; Yonelinas, 2013; Zhang & Luck, 2008; Zhang & Luck, 2009), but that there are also distinct neural signatures tracking memory fidelity (Brunec, Moscovitch, & Barense, 2018; Nilakantan, Bridge, VanHaerents, & Voss, 2018; Richter et al., 2016; Stevenson et al., 2018; Wais et al., 2017; Wais, Montgomery, Stark, & Gazzaley, 2018). These issues were empirically examined using a continuous retrieval task (*Figure 1*; Wilken & Ma, 2004), whereby distractor similarity was found to differentially impact memory detail for the color of objects (Sun, Fidalgo, et al, 2017). Dissimilar distractors disrupted both fine- and coarse-grained information about the target, rendering the memory inaccessible. In contrast, similar distractors disrupted fine-grained but increased the reliance on coarse-grained information about the target, rendering the memory blurred.

Within this perspective, distractor similarity may therefore help or hinder memory depending on the nature of the task (Sun, Fidalgo, et al., 2017). In some cases, similar distractors may help memory when the preserved coarse-grained information can serve as a reminder of the target item (Hanczakowski, Beaman, & Jones, 2017; Jiang et al., 2016; Lin & Luck, 2009; Mate & Baques, 2009; Morey, 2018; Sanocki & Sulman, 2011). In other cases, this benefit from similar distractors may instead manifest as a memory impairment when the coarse-grained information cannot be used to disambiguate between distractor and target items (Barense et al., 2012; Cann, McRae, & Katz, 2011; Roediger & McDermott, 1995; Watson & Lee, 2013; Yeung

et al., 2013). These differential effects of distractor similarity on memory detail were robust for color (Sun, Fidalgo, et al., 2017), but it is not yet known whether they will generalize to another domain. We thus aimed to replicate and extend previous color findings to shape memory, but in other to do so, we faced two nontrivial problems: how could we quantify the subjective similarity of shape, while simultaneously measure memory detail along a continuum?

Perceptually Uniform Spaces

To quantify subjective similarity, some experimenters laboriously collect similarity judgments (Hout, Goldinger, & Ferguson, 2013; Martin et al., 2018; Kriegeskorte & Mur, 2012). Other experimenters sample stimuli from a *perceptually uniform space* (McDermott & Webster, 2012). Perceptually uniform spaces contain stimuli mapped to a *n*-dimensional space, whereby the distances between stimuli have been previously validated to approximate subjective similarity. For example, CIELAB color space is a commonly used three-dimensional perceptually uniform space (e.g. colors are defined on the basis of a luminance dimension, a green-red dimension, and a blue-yellow dimension) that mimics human trichromatic color vision (McDermott & Webster, 2012). To simplify the stimulus set, experimenters typically reduce the dimensionality of CIELAB color space to a two-dimensional *circular space* (i.e. a “color wheel”; *Figure 1a*). By holding luminance constant, stimuli are sampled from the circumference of a circle so that hue varies incrementally (Ma, Husain, & Bays, 2014). Importantly, there are two ways to define similarity on circular color space (*Figure 1c*). First, the circular space can be said to be *perceptually uniform* as it was created from CIELAB space (i.e. colors sampled from any distance apart are about as similar as any other colors sampled from the same degrees’ distance). Second, the *magnitude* of similarity can be approximated by angular distance (i.e. colors sampled from closer distances tend to be more similar than colors sampled from further distances). Here,

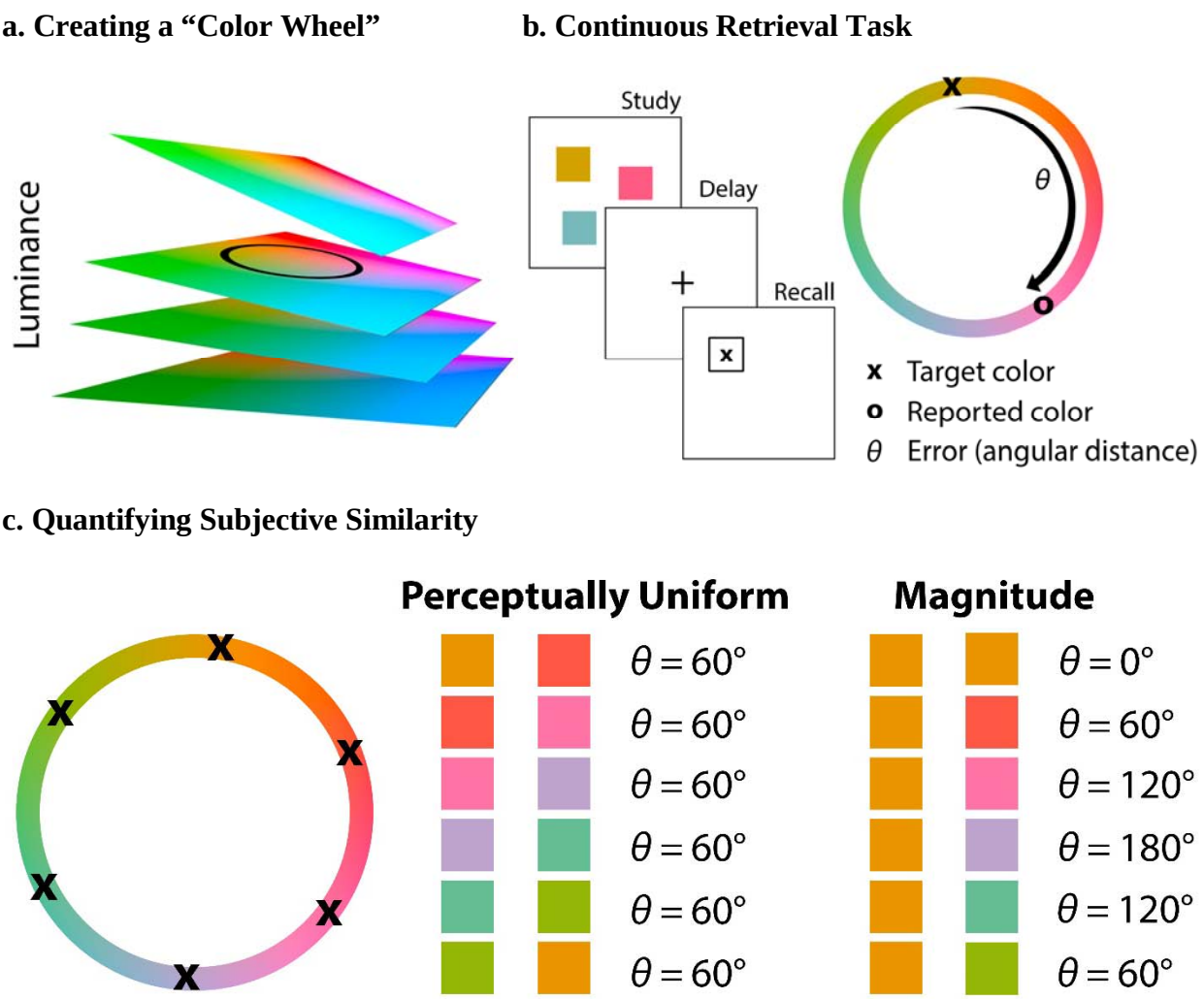


Figure 1. (*a*) Rendition of three-dimensional CIELAB color space, whereby subjective similarity can be operationalized by the Euclidean distance between colors. After controlling for luminance on the y-axis of the figure above, a 2D circle can be defined so that hue varies incrementally. (*b*) On a continuous retrieval task, memory performance (i.e. error) is typically operationalized as the angular distance (θ) between a target and a reported item on circular color space. (*c*) There are two ways to consider subjective similarity. An ideal circular space is *perceptually uniform*, meaning that two items sampled from any distance (60 degrees apart on the example above; see the color pairs displayed in each row) are about as similar as any two other items sampled from the same distance. Second, on an ideal circular space, the *magnitude* of subjective similarity can be approximated by angular distance. In the example above, pairs of colors (displayed in each row) sampled closer in angular distance tend to be more similar than pairs of colors sampled further away.

we wish to emphasize that on an ideal circular space, angular distance needs only to be an approximation of subjective similarity – and does not need to exactly reflect subjective similarity – even if the circular space is perceptually uniform (*Figure 1c*).

The properties of circular spaces are highly useful, as they can be used to not only quantify subjective similarity (Lin & Luck, 2009; Sun, Fidalgo, et al., 2017), but to also quantify memory detail along a continuous metric (Bae et al., 2014; Brady, Konkle, & Alvarez, 2011; Fukuda, Awh, & Vogel, 2010; Ma, Husain, & Bays, 2014; Sun, Fidalgo, et al., 2017). To measure memory detail, a circular space is typically used in conjunction with a continuous retrieval task (*Figure 1b*). Target items are randomly sampled from circular space every trial. After a delay, observers report the target by selecting an item along circular space. An index of memory detail, *error*, is then calculated by taking the angular distance (θ) between the target and reported item (*Figure 1b*). Responses that are low in error thus reflect a high-fidelity memory that closely approximates the target. Mathematical models, such as mixture models (Ma, Husain, & Bays, 2014; Suchow et al., 2013; Sun, Fidalgo, et al., 2017; Zhang & Luck, 2008), are often also used to estimate both the detail of memories as well as the probability of outlier responses that resemble random guesses (*Figure 2*).

Building a Circular Shape Space

A circular space can be used to examine how distractor similarity influences memory detail for shape. However, whereas experimenters have created and validated these stimulus spaces for color (Cheung, 2016; McDermott & Webster, 2012; Pointer, 1981; Robertson, 1977), a validated circular shape space does not yet exist. Instead, shape is a stimulus class frequently used in memory experiments without an explicit measure of subjective similarity. For example, a previous attempt to develop a shape space mathematically varied two sinusoidal components

(Zhang & Luck, 2008; 2009). Although an important first step, human observers do not necessarily perceive subjective similarity the same way they perceive mathematical differences (Busey, 1998; Haushofer et al., 2008; Larkey & Markman, 2005), and thus subjective similarity may be inadequately captured.

In *Experiment 1*, we addressed this issue by creating the *Validated Circular Shape Space* (VCS space), the first perceptually uniform shape space whereby angular distance is a proxy for subjective similarity (available on the Open Science Framework: https://osf.io/d9gyf/). A set of 360 shapes was mapped onto 360 degrees of a 2D circle. This space was empirically validated to confirm that it was perceptually uniform (i.e. shapes sampled from any degrees' distance were about as similar as any other shapes sampled from the same degrees' distance). We then precisely quantified the relationship between angular distance and subjective similarity, demonstrating that shapes sampled closer in angular distance were more similar than shapes sampled further away. In *Experiment 2*, we utilized VCS space to examine how the similarity of distracting information impacts shape memory.

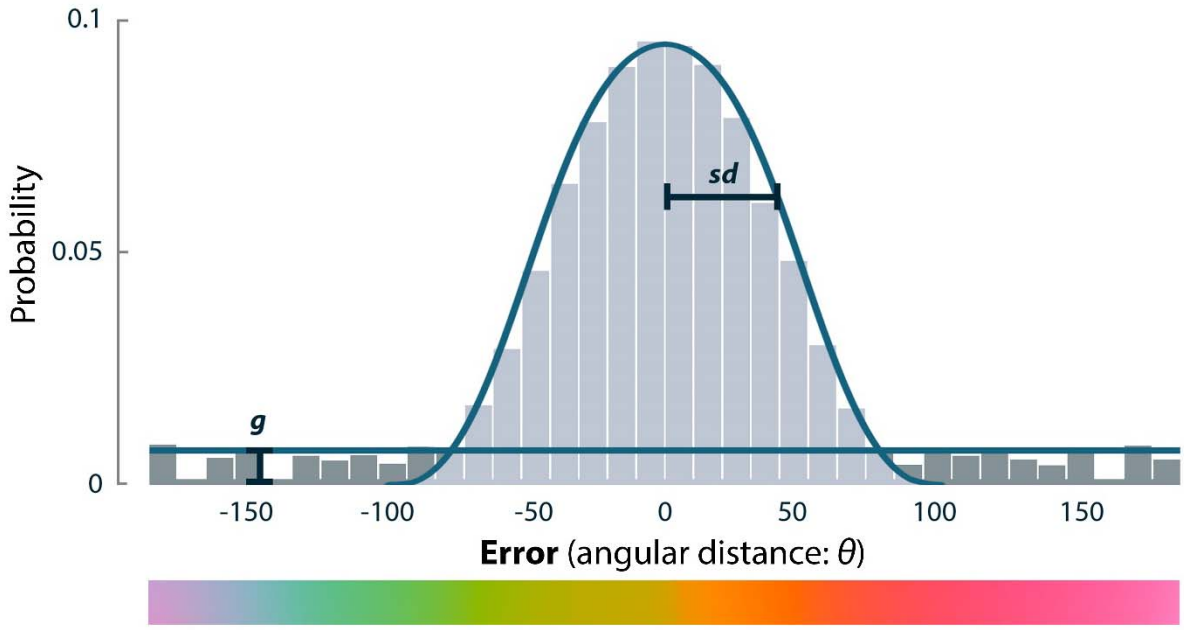


Figure 2. The mixture model often used to estimate memory detail. Error is first calculated as the signed angular distance between a target and a reported color on circular space (*Figure 1b*). Zero error on the x-axis reflects an exact match between the target and reported color, whereas the y-axis reflects the probability of selecting each error across all trials of the experiment. Two parameters are then fit to the histogram of errors using maximum likelihood estimation. A uniform distribution is thought to reflect the probability of random guesses (g). P_{mem} is subsequently defined as 1 – g, representing the probability some fine- or coarse-grained information about a target is remembered. In contrast, *precision* is defined as the inverse of the standard deviation ($1/sd$) of a von Mises distribution (i.e. the circular analogue of a Gaussian distribution), representing memory fidelity. Higher as opposed to lower precision reflects a memory containing more fine-grained information.

Experiment 1. Developing the Validated Circular Shape Space

We implemented a novel iterative validation procedure in which we: 1) designed a set of initial shapes, dubbed "prototypes" (*Figure 3a*); 2) morphed between pairs of these prototypes to generate 360 unique shapes mapped onto each degree of a circle (*Figure 3b*); 3) collected similarity ratings (*Figure 4a*) and collated these ratings across participants (*Figure 4b*) within a group similarity matrix (*Figure 4c*); 4) used this group matrix to reconstruct shape space with multidimensional scaling (MDS) (*Figure 4d*); and 5) identified problematic regions that were not circular (*Figure 4e*). Problematic shapes were iteratively corrected until a circular shape space could be created. For example, prototypes may have been discarded when the similarity between shapes was excessively similar. In other cases, additional shapes may have been included in regions where shapes were more dissimilar than expected. Through the iterative correction of problematic shapes, MDS was used to create a data-driven circular space without the need to mathematically define the dimensions underlying the representation of visual form. Thus, VCS space was made to be perceptually uniform (*Figure 1c*), in contrast to other approaches that did not empirically collect similarity judgments (Zhang & Luck, 2008; 2009). In step 6, for the final validation iteration we quantified the exact relationship between angular distance and subjective similarity using a *distance function*, mapping a single numerical similarity value onto each degree of VCS space. The creation of this distance function was inspired by previous work on circular color space (Schurgin, Wixted & Brady, 2018).

We define the start of one validation "iteration" by the creation of a set of prototype shapes. There were seven iterations in total. For brevity, in the main text we report the methods and results of the 7th iteration that ensured VCS space was circular, as well as an overview of the

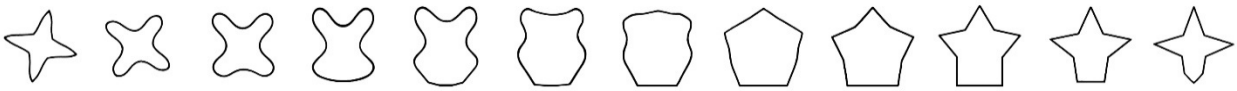
experimental procedure. See *Supplemental Material* for details on the earlier iterations, including how prototype shapes were selected and the specific criterion for defining problematic shapes.

Participant Recruitment

Sixty-seven participants (M_{age} = 21.58, *SD* = 2.76, *Females* = 48) were tested across all seven validation iterations, recruited from the undergraduate psychology student pool at the University of Toronto and from the community. Participants from the undergraduate student pool were compensated with course credit. Participants from the community were compensated $20 CAD for the final validation iteration and $10 CAD for the 1st – 6th iterations. The final validation iteration required up to 90 minutes to complete, which included time for instructions, a practice session, the actual experiment, and debriefing. Earlier validation iterations contained fewer stimuli to be rated and required under an hour to complete. We recruited new individuals for each validation iteration. This study was approved by the University of Toronto ethics board.

It was not possible to determine an a priori sample size because the quantitative measures of circularity were descriptive rather than inferential (*Figure 4e*). Instead, we developed VCS space over many iterations, whereby the stimulus space was successively replicated seven times. For example, in the 6th validation iteration (see *Supplemental Material. Validation 6*) a single problematic region on the reconstructed shape space was identified. Fixing that single problematic region in the final validation iteration produced a circular shape space. The successive replication of VCS space over each iteration combined with a novel procedure involving the alignment of shape space across two independent groups (described in *Validation Procedure*) was intended to maximize the generalizability of our findings. Overall, the final validation iteration tested a sample size of 21 individuals (M_{age} = 22.52, *SD* = 4.14, *Females* = 13).

a. Prototype Shape Stimuli



b. Simplified VCS Space

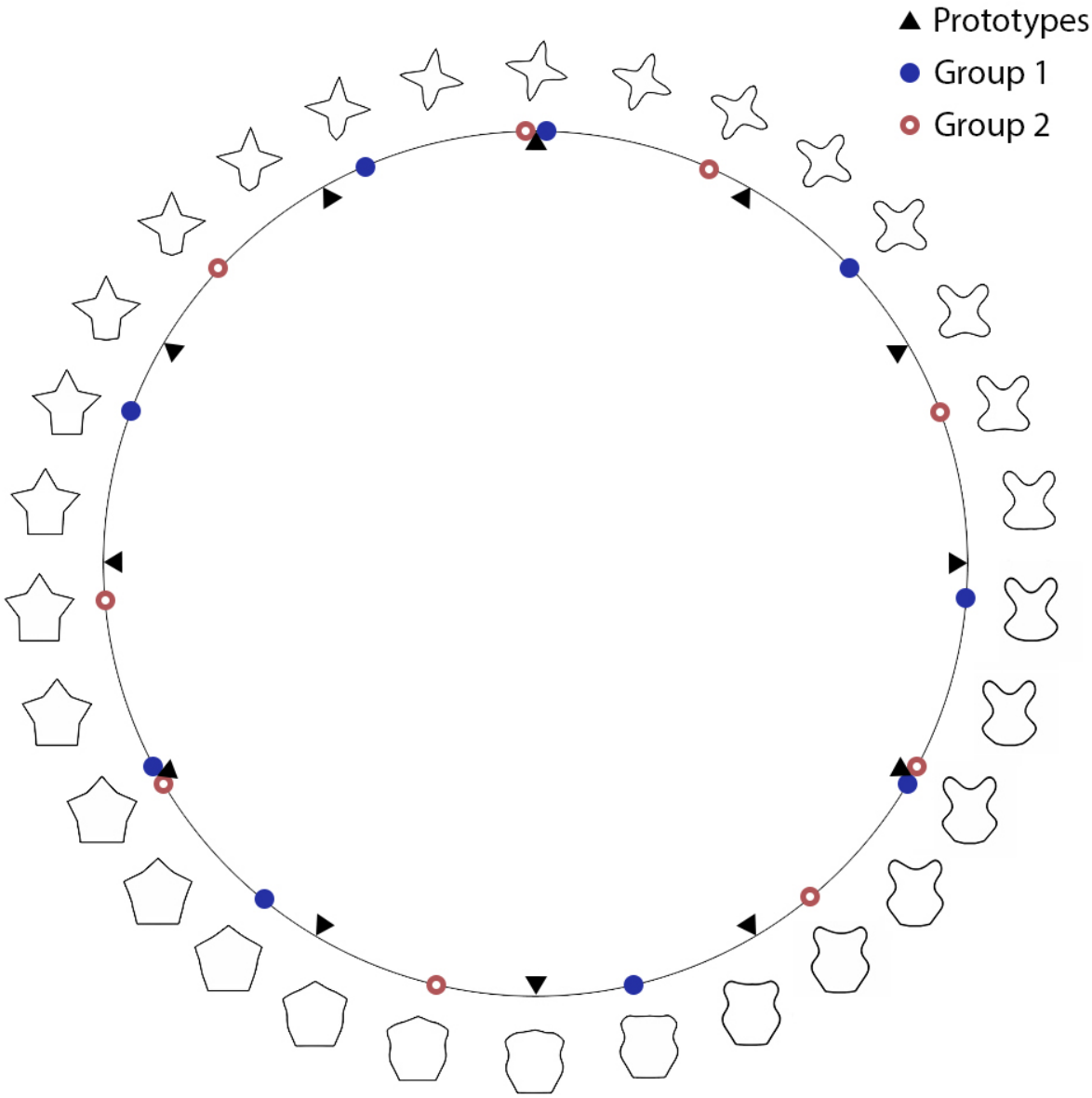


Figure 3. (*a*) The prototype shapes used to generate the stimulus space seperated by 30 degrees of angular distance (positions denoted by triangles in panel b). (*b*) Simplified VCS space with stimuli mapped to every 12 degrees of angular distance. The full circular stimulus space contained 360 stimuli, with a unique morphed shape mapped to each degree on the wheel. Morphs were obtained by blending the prototype shapes, the positions of which are denoted by black triangles. Closed blue denote shapes rated in group 1, whereas open red denote shapes rated in group 2. Three shapes were shared in group 1 and 2, representing the anchors used to align the shape space reconstructions from each group (described in the main text under *Validation Procedure*).

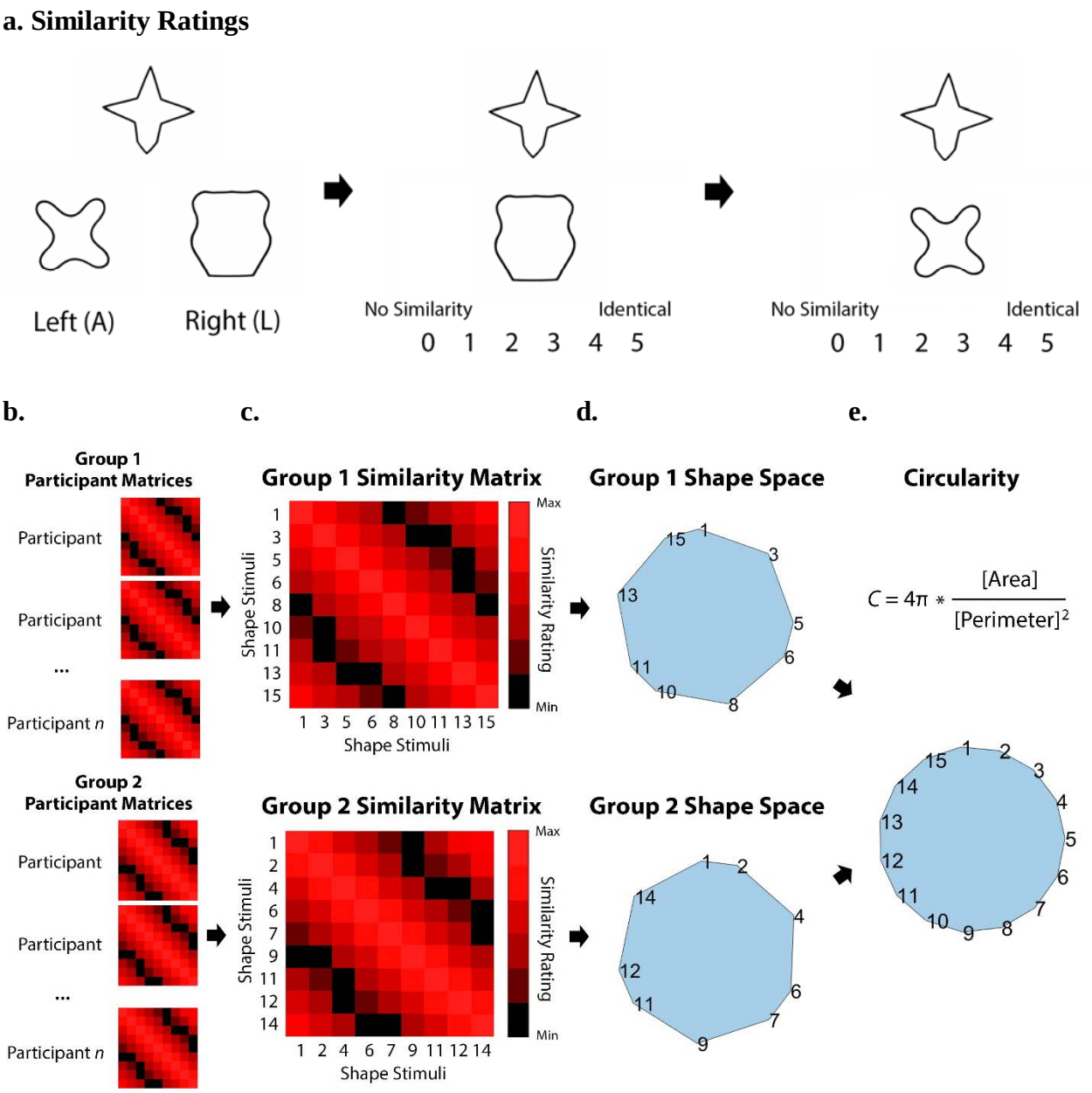


Figure 4. Similarity validation procedure using simulated data for illustrative purposes. (*a*) Participants were shown 3 shapes sampled from the set every trial, and judged which of the bottom shapes were more similar to the top one. After a response, participants rated the similarity between the top shape and bottom shapes. (*b*) The similarity judgments were averaged to create a similarity matrix for each participant. (*c*) All participant similarity matrices were then averaged to produce a similarity matrix for each group, whereby each cell represents the average similarity between two shapes sampled from the stimulus space. (*d*) MDS was used to reconstruct the subjective shape space from each group similarity matrix. (*e*) Last, the group 1 and group 2 shape spaces were aligned together to produce a single shape space. A quantitative circularity score characterized the extent to which the aligned subjective shape space was circular.

Apparatus

The validation experiment was created in MATLAB using psychtoolbox-3 (Kleiner, Brainard, & Pelli, 2007). All stimuli were displayed on the monitor of a Latitude 3460 Dell laptop with a resolution of 1920 x 1080 and a frame refresh rate of 60 Hz. The monitor was 12.18 inches tall by 6.85 inches wide. The distance from the monitor to the participant was approximately 50 cm. All shape stimuli were presented on top of a uniform white background, subtending approximately 3.44 degrees of visual angle. Participants responded to on-screen instructions using the keyboard keys.

Validation Procedure

Step 1 – Designing prototype shapes. In the first validation iteration (see *Supplemental Material*), a set of 2D shape line drawings was created as prototypes. These shapes were incrementally corrected across each validation iteration until no problematic shapes were present. In the final validation iteration, twelve prototype shapes (*Figure 3a*) were created by redesigning the problematic region from the previous validation iteration (see *Supplemental Material. Validation 6*). Problematic shapes were manually redesigned using Photoshop, taking care to keep the change in similarity approximately uniform across all prototypes.

Step 2 – Morphing. Neighboring pairs of prototype shapes were then morphed using SqirlzMorph (Xiberpix, 2008). If visual artifacts were observed as a result of morphing, Photoshop was used to manually smooth the artifacts. This morphing procedure generated a circular stimulus space of 360 unique shapes (*Figure 3b*) from a set of 12 initial prototypes (*Figure 3a*).

Fifteen shapes spaced equidistantly – every 24 degrees – on this stimulus space were selected for validation. Participants were randomly separated either into group 1 (n = 11) or group 2 (n = 10); each group rated six unique shapes in addition to three "anchor shapes" that were identical to both groups (see *Figure 3b*, shapes denoted by closed blue and open red). Critically, the stimuli selected for validation were different from the prototype shapes used to build the space, with the exception of the anchor shapes which were prototypes. These three anchor shapes were specifically selected to be maximally distinct from each other, estimated from the results of the previous validation iteration. We aligned the reconstructed spaces from the two groups by the anchor shapes (described in step 4) based on the similarity ratings (described in step 3). Importantly, not all validation iterations required the selection of anchor shapes, as not all validation iterations required alignment (see *Supplemental Material*).

Step 3 – Similarity ratings. Participants were presented detailed instructions and completed a short practice version (5 trials) of the task prior to the procedure. The nine shapes selected for each group (denoted by closed blue and open red in *Figure 3b*) formed the stimuli to be rated for that group. Every trial, one shape was displayed near the top of the screen and two shapes were displayed near the bottom left and bottom right of the screen (*Figure 4a*). 800 ms later, participants were then cued with text to select the shape on the bottom left ("A" key) or the bottom right ("L" key) that was most visually similar to the shape positioned at the top of the screen. This initial judgment phase in which triads were presented (*Figure 4a*) allowed observers to calibrate to the shapes being rated on a given trial. 1000 ms after a response, the shape on the top and one of the two bottom shapes were randomly paired together. 800 ms later, participants rated the similarity of the paired shapes with 0 being "*no similarity*" and 5 being "*identical*". 1000 ms after a response, the top shape was displayed with the other bottom shape that had not

been rated. 800 ms later, participants again rated the similarity of the paired shapes with 0 being "*no similarity*" and 5 being "*identical*". After this last response, a screen appeared with the number of trials completed, and participants pressed the space key to begin the next trial. Importantly, short delays were included between the presentation of a shape and the presentation of the rating scale, so that participants studied the presented shapes rather than responding in a speeded way. We additionally emphasized to participants that they should use the entire rating scale during the experiment.

Over the course of the experiment, shapes were randomly sampled during the first triad of every trial such that relative to each individual shape in the set, all possible combinations of shape pairs were sampled. More specifically, participants made pairwise ratings between each shape relative to each other shape 14 times (e.g. shape 1 was rated relative to shape 2 in 14 different instances; of these 14 instances, shape 1 was presented at the top of the initial triad 7 times and at the bottom left or right of the initial triad 7 times). Each of the nine shapes was also rated on two trials relative to itself (e.g. shape 1 with shape 1). This created the diagonal in the similarity matrices; *Figure 4b, c*. These rules generated a list of three shapes every trial, and the presentation order of these trials were randomized. For these reasons, the final validation iteration required 270 trials. Earlier validation iterations required less trials (see *Supplemental Material*).

All pairwise ratings involving the same two shapes were averaged (e.g. the similarity values between shape 1 and shape 2 were averaged across all instances shown in the initial triad). These mean scores take into account the overall similarity of the set, and were represented within a similarity matrix for each participant (*Figure 4b*). These participant similarity matrices were then averaged across participants to produce a single similarity matrix for each group (*Figure*

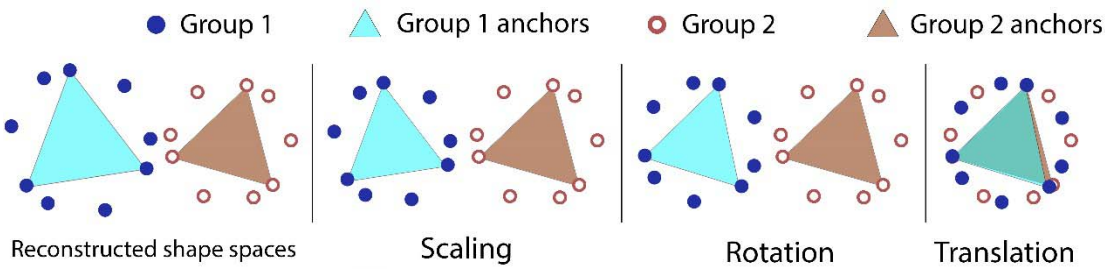
4c). All participant and group matrices were symmetrical due to the averaging procedure. Overall, this within-subject procedure was expected to increase power, reduce error, and allow MDS to recreate subjective space with greater precision.

Step 4 – Reconstructing shape space. From the averaged similarity matrix in each group (*Figure 4c*), shape space was reconstructed with MDS (*Figure 4d*) using the *mdscale* function in MATLAB. The anchor shapes (i.e. shapes that were the same in each group) were then used to align the reconstructed shape space from each group together (*Figure 4e*). Here, we used two separate alignment functions: Procrustes transformation (*Figure 5a*) and affine transformation (*Figure 5b*). Procrustes transformation preserves the relational structure of the space in both groups, which may result in imperfect alignment. In contrast, affine transformation includes an additional "skew" step and can perfectly align the two group shape spaces by the three anchor points, but may distort the relational structure of one of the spaces. As the affine transformation may alter the relational structure of one group space (*Figure 5b*), we visually compared the solutions obtained from the Procrustes and the affine transformations. Furthermore, only the affine aligned shape space could be used to quantitatively determine whether the final validation iteration was circular, as the Procrustes transformation does not perfectly align the group spaces (see *Step 5. Accessing circularity*).

We could not collect ratings for all 360 shapes from VCS space due to the exorbitant number of trials required for pairwise similarity judgments (64,620 trials just to collect a single rating for all pairwise combinations of 360 shapes). Researchers have previously decreased the number of trials needed by using a spatial similarity method (Kriegeskorte & Mur, 2012), in which participants "drag-and-drop" stimuli to locations on a computer screen. In this method, observers use the spatial proximity of items as an index of subjective similarity. However,

because our goal was to develop a particular spatial configuration (i.e. a circle), we did not want to bias participants into using a spatial similarity method which may unintentionally lead to circular configurations (Verheyen et al., 2016). Our novel procedure in which two independent groups each rated 6 unique shapes + 3 anchor shapes was instead used to infer the overall relational structure of VCS space. This method provides the most powerful demonstration that the space was indeed circular, because a circular space could have only been reconstructed across groups if subjective similarity changed uniformly and incrementally across the entire stimulus space.

a. Procrustes Transformation



b. Affine Transformation

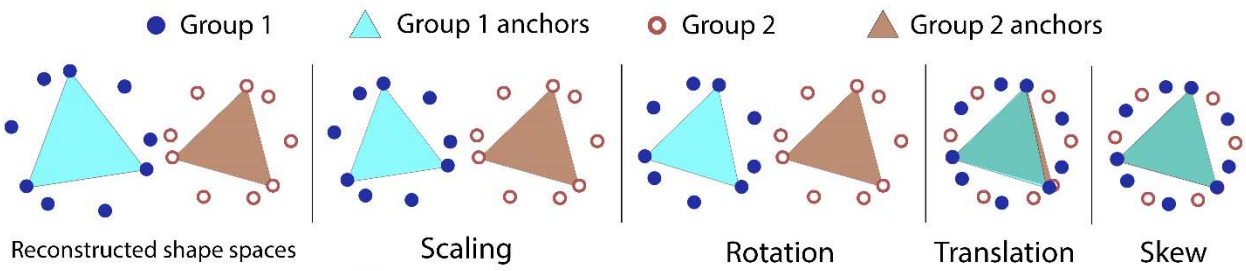


Figure 5. Visualization of the alignment procedure, whereby the transformation order of steps is simplified and the expected positional differences between spaces is exaggerated for illustrative purposes. First, MDS was used to reconstruct a separate shape space from each group similarity matrix. Separate alignment functions were then applied to the two shape spaces. Linear transformations are first determined for the anchor shapes, shown as a triangle, then applied to the rest of the space. (a) Procrustes transformation includes scaling, rotation, and translation, which preserves the relational structure of both spaces but can lead to imperfect alignment. (b) Affine transformation includes scaling, rotation, translation, and skew, which can distort the relational structure of one of the group shape spaces, but will perfectly align both groups together.

Statistical Analysis

Step 5 – Assessing circularity. It is standard practice for experimenters to qualitatively interpret reconstructed space, as MDS is a descriptive analysis (for examples of previous applications of MDS in the cognitive sciences, see Hout, Goldinger, & Ferguson, 2013; Hout, Papesh, Goldinger, 2013; Hout et al., 2016; Shepard, 1980). In addition to this qualitative assessment, we added a quantitative approach to assess whether or not a reconstructed shape space was circular. Circularity was defined using a simple mathematical ratio that captures the relationship between the area and perimeter of a spatial configuration. More specifically, the circle is the spatial configuration with the largest area relative to perimeter in the plane (see *isoperimetric inequality*; Osserman, 1978):

$$Circularity\ (C) = 4\pi * \frac{[Area]}{[Perimeter]^2}$$

To calculate C, borders were drawn between the projected coordinates of each shape in the aligned shape space (*Figure 4e*), following the relational structure of the original stimulus space (*Figure 3b*). For example, if the reconstructed position of shape 1 was at 1 degree on the stimulus space, shape 2 at 25 degrees, and shape 3 at 49 degrees, then a border was drawn from shape 1 to shape 2 to shape 3 (and so forth, with the final shape connected back to shape 1). Area and perimeter were then calculated to determine a value of C (*Figure 4e*). A space that is a perfect circle will have a C value equal to 1.0. As a space is increasingly less circular, C will approach 0.0. To complement this quantitative circularity score, we qualitatively assessed whether the positioning of the vertices within the reconstructed shape space approximately matched the vertices of the original stimulus space. For every validation iteration, a reconstructed shape space was defined quantitatively as sufficiently circular when C exceeded

0.90 and qualitatively as sufficiently circular when the relative positions of the vertices approximately matched the original shape space. See *Figure 6* for examples of how C changed from the first validation iteration to the final validation iteration.

C was calculated specifically for the affine aligned shape space (*Figure 5b*). We used the affine aligned shape space to determine C rather than the Procrustes aligned shape space, because Procrustes transform does not always perfectly align the two group shape spaces together (*Figure 5a*). However, because the affine alignment procedure may distort the relational structure of one group shape space (*Figure 5b*), we conducted both Procrustes and affine transformations and visually compared the solutions. Permutation testing was conducted to determine whether the C value for the affine aligned shape space could have been obtained by chance. The shape labels of each participant matrix were shuffled to form a scrambled version of *Figure 4b*. These null participant matrices were averaged together to produce a null group similarity matrix (a scrambled version of *Figure 4c*). Null MDS coordinates were then projected from each null group similarity matrix, forming a null reconstructed space for each group. These null shape spaces were aligned together to produce a null affine aligned shape space. A null C score then determined the circularity of this null affine aligned shape space. This process was repeated 10,000 times to generate a null distribution of C scores. Statistical significance was defined as probability $p < 0.05$ for the observed C score relative to the null distribution of C scores. Critically, the observed C value derived from the reconstructed subjective space gives information regarding circularity, whereas the p value obtained from permutation testing gives information regarding whether the subjective space could be formed by chance. A subjective shape could be noncircular (e.g. $C = 0.75$, see *Figure 6*), yet could be statistically significant (e.g. $p < 0.05$) because participants may be homogenous in how they perceived a stimulus space.

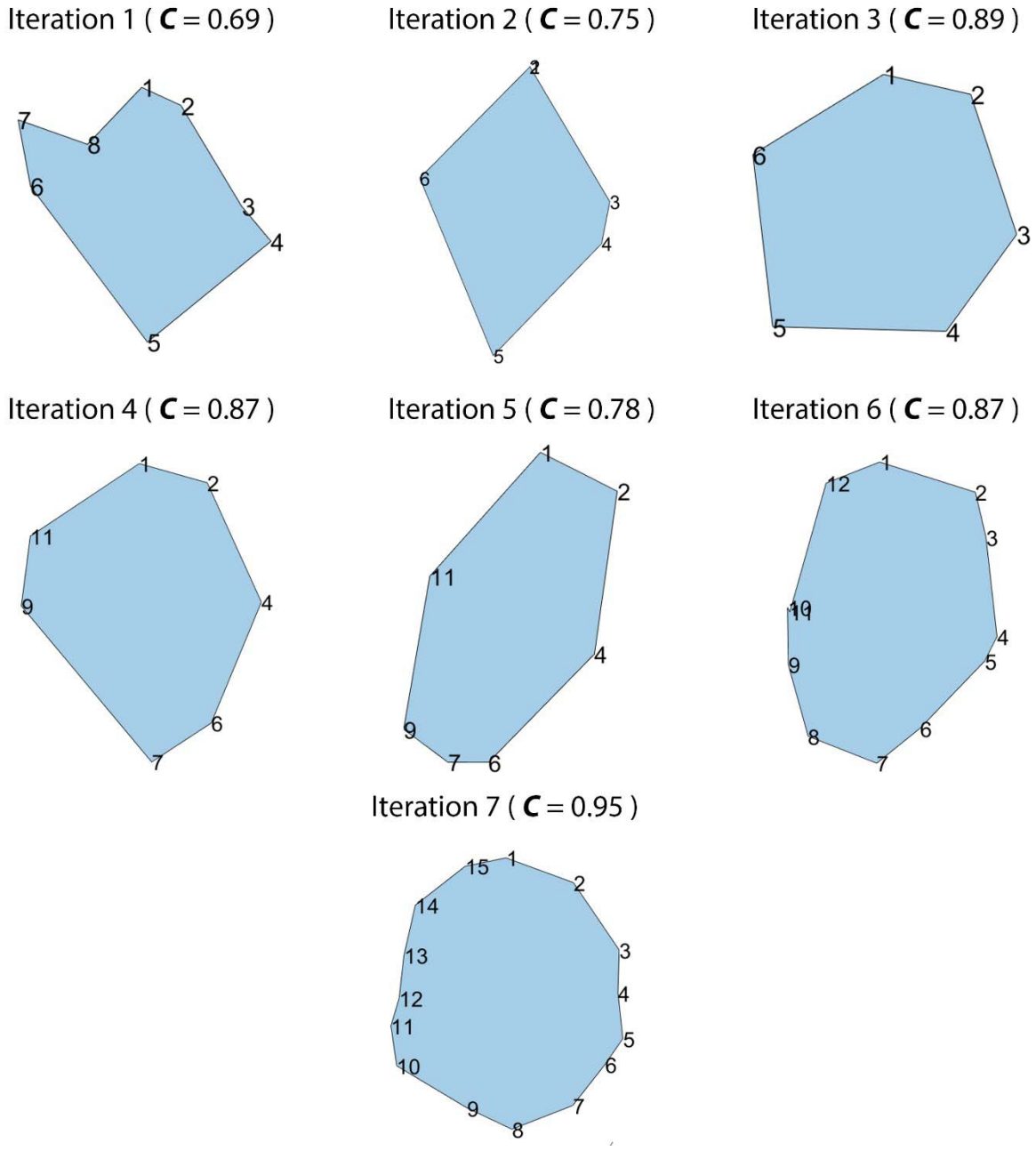


Figure 6. C was calculated for each validation iteration (see *Supplemental Material* for detailed results). All validation iterations were statistically significant using permutation testing. Numbers represent the position of each shape from the associated stimulus space.

We also examined the uniformity of individual shape spaces by calculating a *homogeneity score* representing the overall consistency between all the similarity judgments made for a given validation iteration. To calculate this score, each participant matrix (*Figure 4b*) was correlated with every other participant matrix. As participant matrices were symmetrical (see *Step 3 – Similarity ratings*), only the upper triangle of each participant matrix excluding the diagonal was correlated to each other. This created a new matrix with elements containing the correlation between each participant to every other participant. We first normalized these correlations by transforming to Fisher's z (Corey, Dunlap, & Burke, 1998). These z values were averaged then transformed back into a correlation value to produce a single score representing the degree to which individuals homogeneously perceived a shape space.

Permutation testing was conducted to compute the likelihood of observing the homogeneity score by chance. The labels of each participant matrix (*Figure 4b*) were shuffled and the upper triangle of these matrices were correlated to each other, forming a null correlation matrix. We then averaged the normalized correlations in the upper triangle of the null between-participant correlation matrix to produce a null homogeneity score. This procedure was repeated 10,000 times to produce a null distribution of homogeneity scores. Statistical significance was defined as probability $p < 0.05$ for the observed homogeneity score relative to the null distribution of homogeneity scores.

Step 6 – Distance function. Similarity can be considered in two ways for a circular space (*Figure 1c*). First, an ideal circular space is perceptually uniform, meaning that items sampled from any distance are about as similar as any other items sampled from that same distance. Second, the magnitude of similarity can be approximated by angular distance, meaning that items closer in angular distance are more similar than items further in angular distance. The

circularity value (*Figure 4e*) gives information regarding the degree to which the shape space is perceptually uniform. However, the circularity value does not give information regarding the magnitude of the numerical change in similarity from one shape to the next.

To determine if angular distance is an approximation of subjective similarity, we combined a quantitative and qualitative approach. Quantitatively, a function was used to predict the collected similarity ratings using the angular distance of shapes positioned on the stimulus space, akin to a regression analysis predicting subjective similarity from angular distance. This *distance function* mathematically describes the relationship between angular distance and subjective similarity on VCS space (see *Results* for a step-by-step description). In this way, a numerical similarity value was obtained for each possible pairing of the 360 shapes in VCS space, even though ratings were only collected for 15 shapes (*Figure 3b*). Qualitatively, we examined the change in subjective similarity with respect to angular distance in each group similarity matrix (*Figure 4c*). If angular distance approximated subjective similarity, then shapes closer in angular distance should be associated with a higher similarity rating than shapes further in angular distance.

Summary of analyses. We computed the following set of measurements: 1) C, representing the circularity of the affine aligned shape space, 2) the *homogeneity score*, representing the mean correlation between all participant similarity judgments, and in the final validation iteration, 3) a *distance function* quantified the relationship between angular distance and subjective similarity.

Two types of similarity (*Figure 1c*) on VCS space were examined: whether the space was perceptually uniform, and whether angular distance approximated subjective similarity. First, a reconstructed shape space was defined as sufficiently circular (i.e. perceptually uniform) when

C exceeded 0.90 and when qualitatively, the relative positions of all vertices of the reconstructed shape space tended to match the original stimulus space. Second, if angular distance approximated subjective similarity (*Figure 1c*), then shapes closer in angular distance should be more similar than shapes further away.

Results

Validation Iterations 1 – 6

Circularity. Seven validation iterations were required to exceed a C score of 0.90 (*Figure 7; Supplemental Material*). We found that the first validation iteration was not circular (C = 0.69), suggesting that stimulus spaces arranged by experimenters without explicitly collecting similarity judgments may not match subjective similarity. All C scores *(Figure 7a*) were significantly different from what would be expected by chance alone (at or below $p < 0.01$ using permutation testing; see *Methods – Step 5. Assessing circularity* for a description of the analysis). This suggests that the reconstructed shape space within each iteration could not have been obtained by chance, but rather reflects the degree to which the stimulus space is perceptually uniform.

Homogeneity score. The mean correlation between participant similarity judgments was consistently very high across all validation iterations (all homogeneity scores were statistically significant using permutation testing at $p < 0.0001$; *Figure 7b*). Overall, this result suggests that individuals were very similar to each other in their representation of the stimulus space across validation iterations, even when the stimulus space in a specific validation iteration was not circular.

a. Circularity (C) **b. Homogeneity Score**

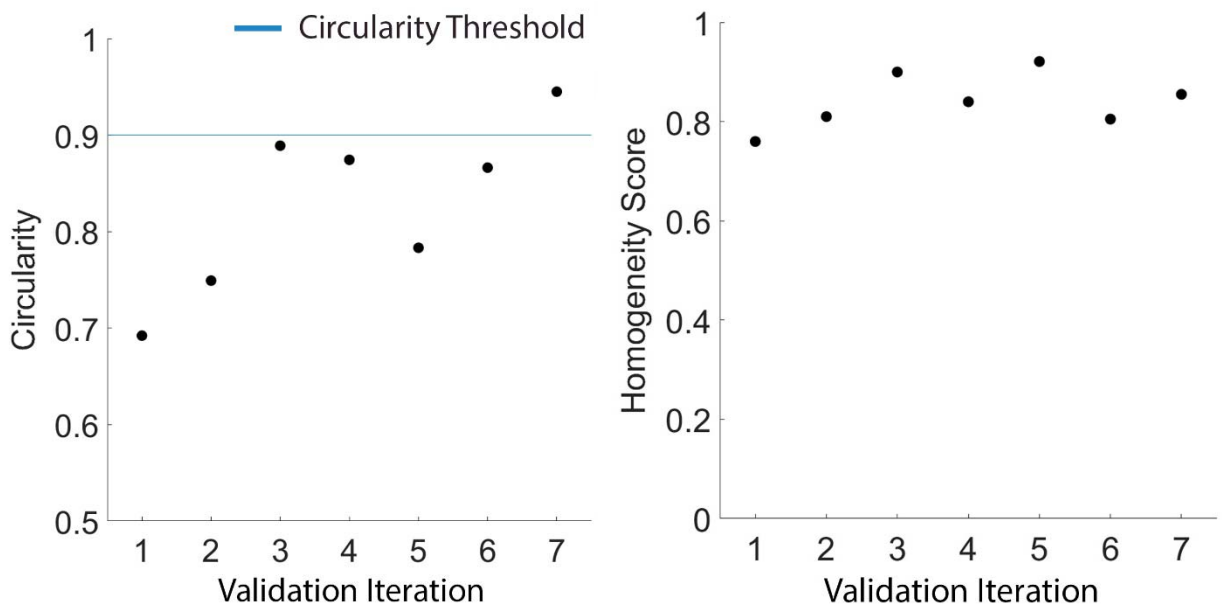


Figure 7. Overall results across validation iterations. (*a*) C by validation iteration, representing the degree to which each stimulus space was perceptually uniform. Higher values reflect a space that is closer to a perfect circle, meaning a space that is more perceptually uniform. (*b*) Homogeneity scores by validation iteration, reflecting the mean correlation between participant similarity judgments. The homogeneity score determined for each group was averaged to produce a single value for a given validation iteration in the plot above. These results suggest that even if a stimulus space in a validation iteration was not circular, participants were homogeneous in how they perceived shape similarity.

Final Validation Iteration

Circularity. In the 7th and final validation iteration, *C* of the affine aligned shape space was 0.95 ($p < 0.0001$ using permutation testing), suggesting that it was circular (*Figure 8*). Qualitatively, the positions of shapes on both the Procrustes (*Figure 8b*) and affine aligned (*Figure 8c*) shape space approximately matched the positions of shapes sampled from VCS space. These results confirmed that VCS space was circular, as the shapes to be rated in each group were sampled from arbitrary positions on the original stimulus space. For detailed group results, see *Supplemental Material. Validation 7*.

Complimenting the alignment procedure, when the group 2 similarity matrix was overlaid with the group 1 similarity matrix (*Figure 9*), a striking trend could be observed. Angular distance approximated (though was not identical to) subjective similarity: subjective similarity tended to decrease as angular distance increased. For example, shape 1 and shape 2 were separated by 24 degrees on VCS space and were associated with a similarity value of 3.8 (see the second element of the first row on *Figure 9*). Increasing the angular distance between shapes by another 24 degrees and comparing shapes 1 and 3 decreased the observed similarity rating to a value of 3 (see the third element of the first row on *Figure 9*). Importantly, this matrix was created from combining the similarity ratings *across groups* with no adjustment whatsoever. That is, we overlaid the similarity matrices from each group (see *Supplemental Material. Validation 7*) into a single matrix without interpolation across groups, further confirming that the change in similarity across VCS space was uniform.

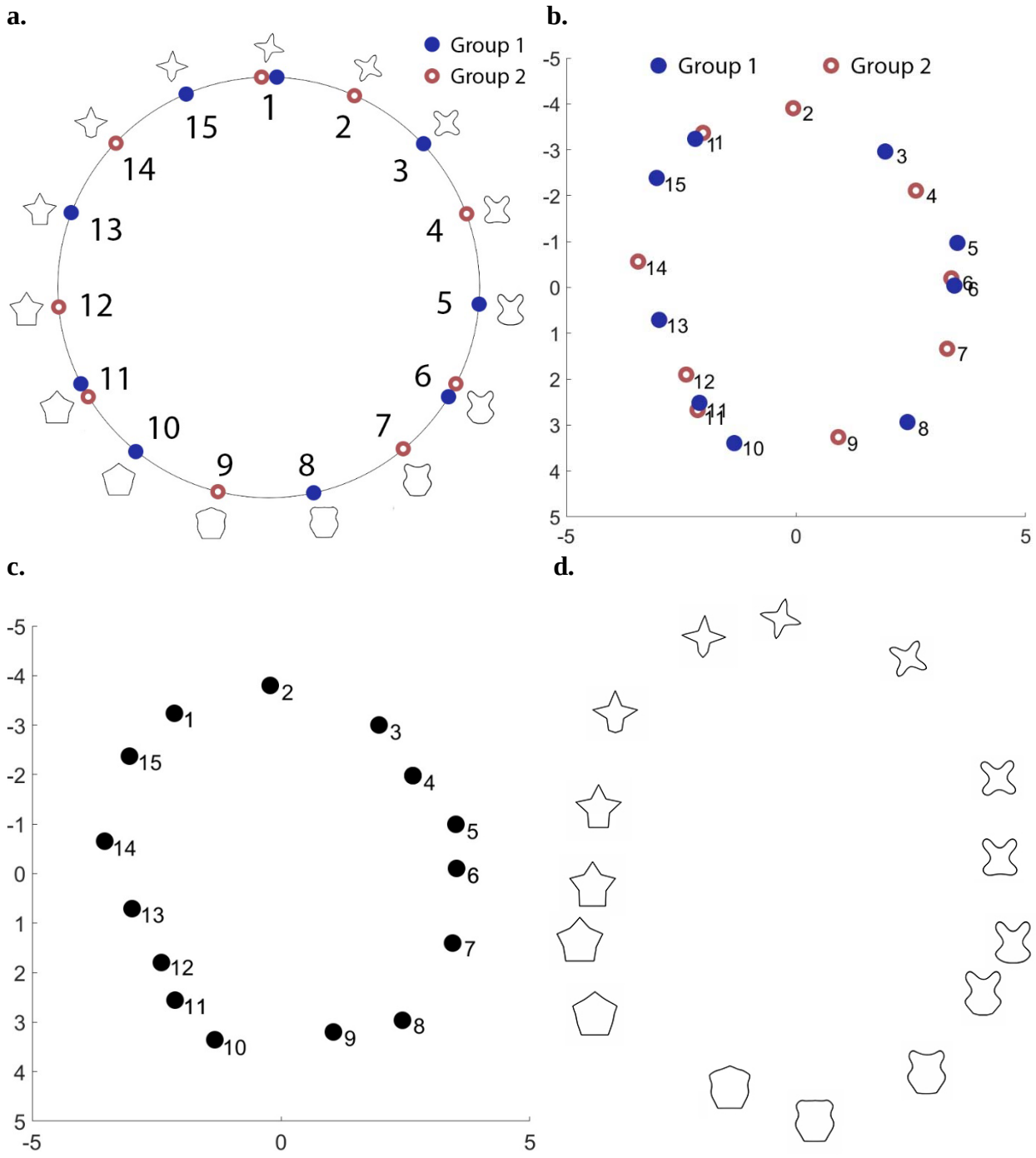


Figure 8. Results of the final validation iteration. Numbers represent the sequential position of the rated shapes sampled from VCS space (e.g. each rated shape was 24 degrees apart, so shape 1 would be at 1 degree, shape 2 at 25 degrees, shape 3 at 49 degrees, and so on). (*a*) Simplified VCS space labeled with the shapes rated during the final validation iteration. (*b*) Procrustes alignment of group 1 (closed blue) and group 2 (open red). (*c*) Affine alignment between group 1 and 2. Both alignment functions were highly similar and both resulted in circular reconstructions. (*d*) The final shape space reconstructed by affine transformation on the MDS solutions from group 1 and 2.

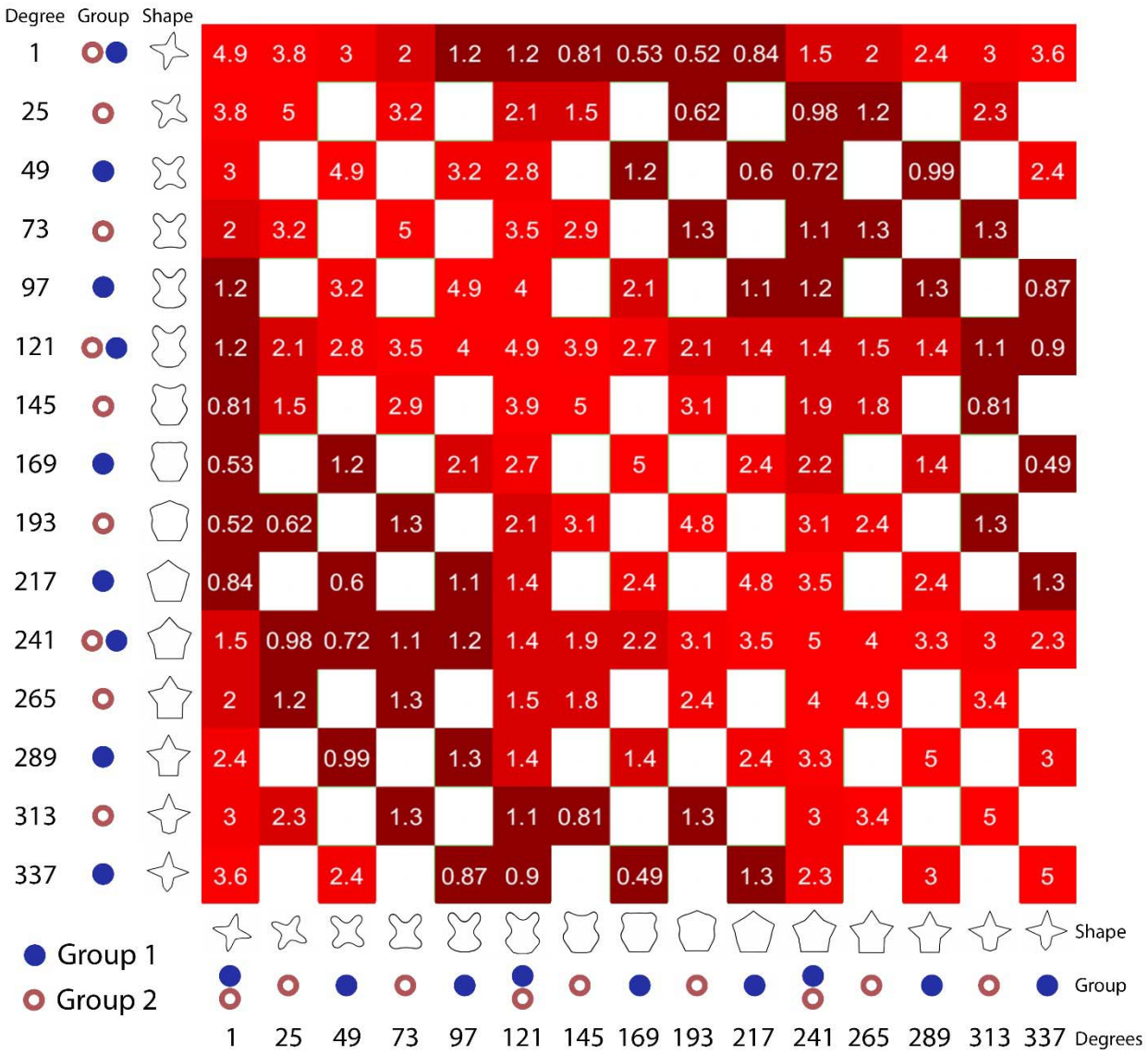


Figure 9. The similarity matrix for group 2 was simply combined with the similarity matrix for group 1. Elements with a value in both group 1 and group 2 were averaged together. Blank elements refer to pairs of shapes that did not have a similarity rating as they were in distinct groups. Each element refers to the averaged similarity rating across all participants, bounded from 0 (*no similarity*) to 5 (*identical*). The x- and y- axis include the rated shapes, the group, and the position on VCS space (*Figure 3b*) from which the shape was sampled. Here, we see a striking trend in similarity across groups, whereby shapes closer in angular distance were rated more similarly compared to shapes further in angular distance.

Homogeneity score. The homogeneity score was very high in both groups (group 1 = 0.87, group 2 = 0.86), suggesting that participant similarity judgments were highly correlated to each other (*Figure 7c*). The observed correlation across participants could not be based on chance alone, as both homogeneity scores were statistically significant at $p < 0.0001$ using permutation testing.

Distance function. We quantified the exact relationship between angular distance and subjective similarity on VCS space, inspired by previous work on a distance function for circular color space (Schurgin, Wixted, & Brady, 2018). Using the collected similarity judgments across both groups (*Figure 9*), the subjective similarity of each shape relative to the set was plotted by degrees on VCS space. This is shown on *Figure 10a*, where every 24 degrees reflects the similarity value of one shape relative to every other shape in the set. For example, the similarity between shape 1 rated with itself is plotted at 0 degrees, shape 1 rated with shape 2 is plotted at 24 degrees, shape 1 rated with shape 3 is plotted at 48 degrees, and so forth. Likewise, the similarity between shape 2 rated with itself is plotted at 0 degrees, shape 2 rated with shape 3 is plotted at 24 degrees, shape 2 rated with shape 4 is plotted at 48 degrees, and so forth until all similarity values from *Figure 9* are plotted on *Figure 10a*.

These collected similarity ratings were averaged and then replotted by angular distance (*Figure 10b*). More specifically, the angular distance between shape 1 relative to shape 2 as well as shape 1 relative to shape 15 is 24 degrees' distance (because shapes are mapped along 360 degrees of a circle). All instances between pairs of shapes with equivalent angular distances were averaged to obtain a single similarity value at every 24 degrees' distance. To predict a similarity value for each degree on VCS space given our observed similarity ratings, a nonlinear exponential function successfully fit the relationship between subjective similarity and angular

distance (R^2 = 0.998; *Figure 10b*). This nonlinear function was normalized to create the line in *Figure 10c*. Normalization converted the y-axis of this function from our similarity rating scale (0 – 5) to a percentage scale (0% - 100%). Taken together, this distance function predicted a numerical similarity value – a percentage match in similarity – between any two shapes sampled from VCS space. Angular distance was an approximation of subjective similarity, as the subjective similarity between two shapes incrementally decreased as angular distance increased (*Figure 10c*).

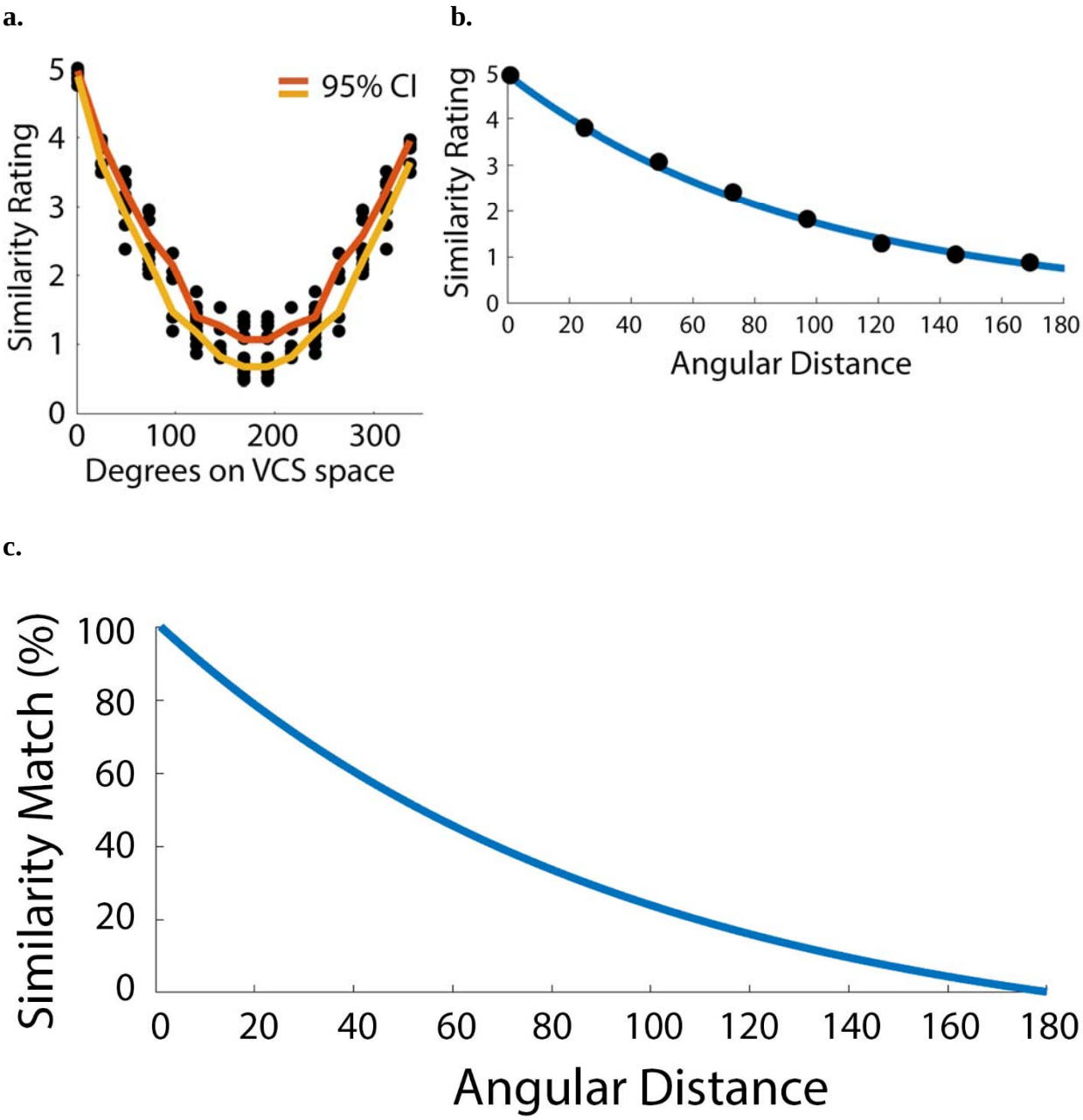


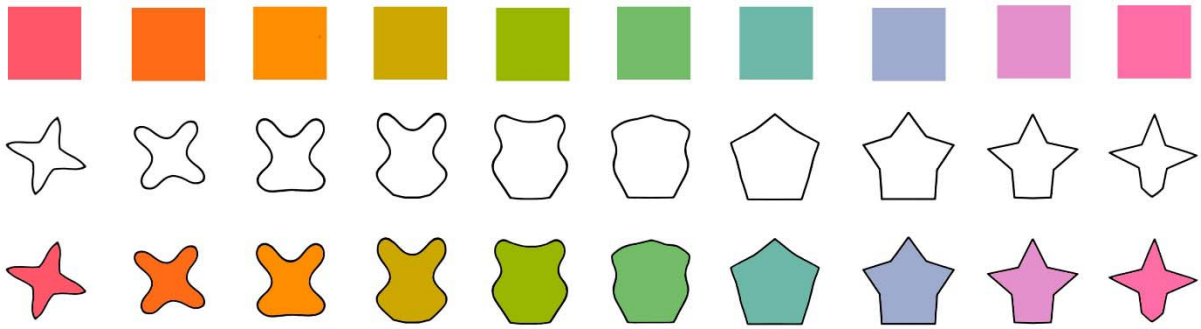
Figure 10. VCS space distance function. (*a*) Relative to each rated shape, all similarity judgments from *Figure 9* were replotted as a function of degrees on VCS space. Error bars reflect the 95% CI for the mean similarity rating as a function of degrees on VCS space. (*b*) Average subjective similarity as a function of angular distance. Angular distance approximated subjective similarity, as a nonlinear exponential function (R^2 = 0.998) well predicted the relationship between angular distance and the observed similarity ratings. (*c*) The nonlinear exponential function was normalized in order to directly compare between shapes along a percentage scale, bounded by 100%, representing a perfect match in similarity, and 0%, representing maximal dissimilarity with respect to the shape space.

Experiment 1 Discussion

We provide and characterize VCS space, the first perceptually uniform shape space whereby the angular distance between 360 shapes mapped on a 2D circle is a proxy for subjective similarity (available on the Open Science Framework: https://osf.io/d9gyf/). To achieve this property, VCS space was developed across a series of seven validation iterations. In the final validation iteration, MDS reconstructions from the similarity ratings of two separate groups were aligned to recreate the original circular stimulus space. VCS space is analogous to a "color wheel" built from CIELAB color space, previously validated to approximate the way in which human observers perceive color similarity (McDermott & Webster, 2012; Pointer, 1981; Robertson, 1977).

Experimenters using VCS space may consider similarity in two ways. VCS space is *perceptually uniform*, meaning that shapes separated by any distance are about as similar as any other shapes separated by the same distance. By sampling stimuli from equidistant angular positions on VCS space, the variability in similarity from one shape to the next is consistent across the set. In this way, the experimenter could utilize VCS space alone or in conjunction with CIELAB color space so that the similarity of multiple features within an object can be empirically manipulated (*Figure 11a*). The second way to consider similarity is with respect to *magnitude*, meaning that shapes closer in angular distance tend to be more similar than shapes further in angular distance. If the magnitude of similarity must be quantified with numerical precision, then the researcher may consider using our distance function (available on the Open Science Framework: https://osf.io/d9gyf/; see *Figure 10c*). In *Figure 11b*, each shape was sampled using the distance function so that subjective similarity changed incrementally by 20%.

a. Perceptually Uniform Stimuli



b. Quantifying the Magnitude of Subjective Similarity

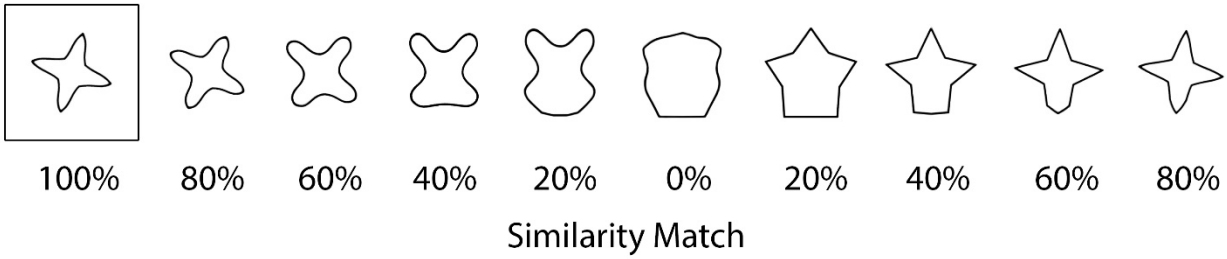


Figure 11. (*a*) Example of how the *perceptually uniform* property of VCS space can be used in an experiment. Colors sampled from equidistant positions on CIELAB space (on a circle centered around L = 70, a = 20, b = 38, with radius = 60 units) and shapes sampled from equidistant positions on VCS space can be used to create objects with multiple features that are well-controlled in subjective similarity. (*b*) The *magnitude* of subjective similarity is approximated by angular distance on VCS space. To quantify the magnitude of subjective similarity with numerical precision, we used our distance function to sample shapes. Each shape shown above differs by a 20% match in subjective similarity relative to the shape displayed in the box.

The successful creation of VCS space directly implies the existence of a fundamental shape space coded by the visual system. This conclusion is additionally supported by the high homogeneity scores across each validation iteration (*Figure 7c*), suggesting that participants were highly correlated in how they perceived a stimulus space, even when a given stimulus space was not circular. Researchers have previously proposed that neurons sensitive to curvature may be important for processing visual form, with resultant models based on curvature detecting units (Bell et al., 2010; Loffler, 2008; Pasupathy, 2006; Salmela, Henriksson, & Vanni, 2016; Schmidtmann & Fruend, 2019; Wilson & Wilkinson, 2015). Though we have established that VCS space is circular, we do not yet know the dimensions underlying the representation of our shape space. One future direction of investigation may be to determine whether existing shape curvature models can describe the organization of VCS space. Another direction of investigation may involve neuroimaging to examine whether the circular organization of VCS space is contained within patterns of neural activity. Akin to how neural representations of color in V4 are organized in a way similar to subjective color space (Brouwer & Heeger, 2009), and object representations in IT are organized in a way similar to subjective semantic space (Charest et al., 2014), neural representation of visual form may follow a similar organization to subjective shape space. One candidate region is LO, known to be important for shape processing (Drucker & Aguirre, 2009; Haushofer, Livingstone, & Kanwisher, 2008; Op de Beeck, Torfs, & Wagemans, 2008).

In addition to understanding how shapes are processed by the visual system, VCS space can be utilized to study how shapes are processed during memory. Current visual short-term memory experiments typically use color and orientation as the stimuli of choice (Ma, Husain, & Bays, 2014), and while this research is immensely informative, our well-validated perceptually

uniform set of shapes can allow for novel investigation of a feature dimension currently not well understood. These research possibilities include perceptual averaging (Cant & Xu, 2012), feature binding (Bays, Wu, & Husain, 2011; Fougnie & Alvarez, 2012; Pertzov & Husain, 2014; Schneegans & Bays, 2017), and more broadly, measuring memory detail along a continuous metric (Bae et al., 2014; Brady, Konkle, & Alvarez, 2011; Fukuda, Awh, & Vogel, 2010; Ma, Husain, & Bays, 2014; Sun, Fidalgo, et al., 2017).

In *Experiment 2*, we used VCS space in conjunction with a continuous retrieval task (*Figure 12*) to examine how the similarity of distracting information impacts shape memory. The methodology was based on a previous study assessing memory for color (Sun, Fidalgo, et al., 2017). In this study, dissimilar color distractors disrupted both fine- and coarse-grained color information, rendering the memory inaccessible. In contrast, similar color distractors disrupted fine-grained but increased the reliance on coarse-grained color information, rendering the memory blurred. Although this study showed that the distractor similarity of colors differentially altered memory detail, it is not yet known whether these effects reflect a general set of principles regarding how the type of interference impacts memory. We therefore aimed to replicate and extend previous color findings to shape memory. To estimate memory detail, we used a mixture model (*Figure 2*) as well as a set of model-free analyses.

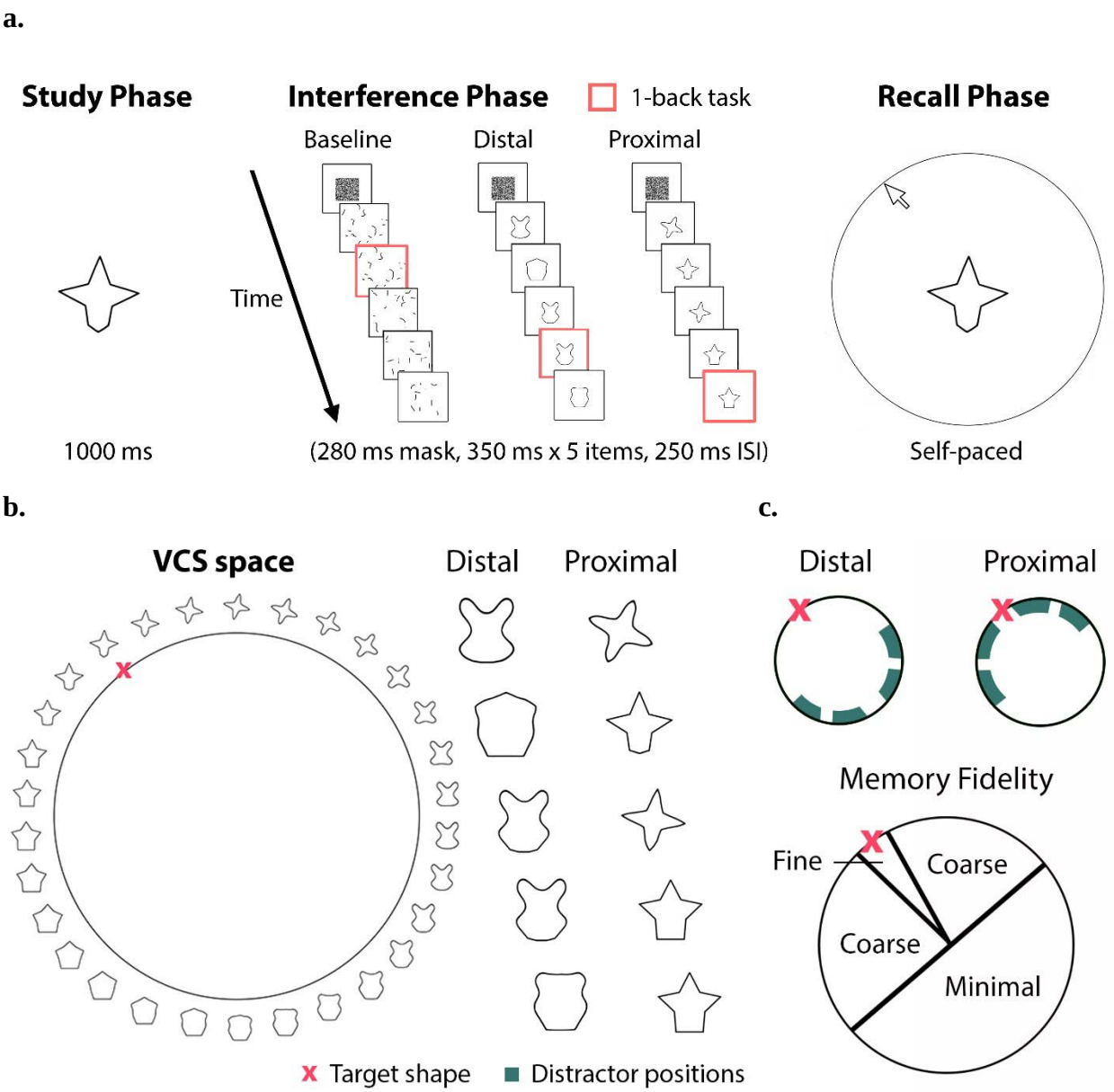


Figure 12. Experiment 2: Continuous retrieval task. (*a*) During the study phase, individuals remembered a shape randomly sampled from VCS space. Next, distractors were presented sequentially during the interference phase. There were three conditions, baseline (scrambled shapes), distal ("subjective dissimilar"), and proximal ("subjectively similar"). During the recall phase, participants reported the target shape on a response circle. Each degree of the response circle was associated with a shape from VCS space (*b*), which appeared at the center of the screen when the mouse cursor moved. (*b*) For distal trials, distractors were sampled from the opposite side of VCS space. For proximal trials, distractors were sampled from the same side of VCS space. In this way, angular distance on VCS space was used to approximate the similarity of distractors. (*c*) Distractors were sampled from four subdivisions in each condition, separated by buffer zones, with one shape repeated as a measure of attention. Last, these subdivisions were used in a model-free estimate of memory fidelity (see *Experiment 2. Model-free analysis*).

Experiment 2. The Differential Impact of Distractor Similarity

Participant Recruitment

Thirty participants (M_{age} = 20.13, *SD* = 2.16, *Females* = 21) were recruited from the undergraduate psychology student pool at the University of Toronto and from the community. This sample size was selected based on past research using the same methodology, as we aimed to replicate previous color findings (Sun, Fidalgo, et al., 2017) to a novel stimulus class. A sample size of thirty was required to achieve a probability of 0.95 of detecting a true main effect in a three-condition, one-way repeated measures analysis of variance (rm-ANOVA), assuming a medium effect size (η^2 = 0.10) and a modest mean correlation between repeated measures (r = 0.40). All power analyses were conducted using G*Power (Faul et al., 2009). To increase power, we also increased the number of trials from 75 trials per interference condition (Sun, Fidalgo, et al., 2017) to 100 trials per interference condition (see *Procedure*). Participants from the undergraduate psychology student pool were compensated with course credit and participants from the community were compensated with $20 CAD. This experiment (n = 330 trials) required approximately 90 minutes to complete, which included consent, instructions, a short practice session of the task (12 trials), the actual experiment, and debriefing. Participants were allotted an optional break approximately every 15 minutes. This study was approved by the University of Toronto ethics board.

Stimuli were shapes sampled from VCS space (*Figure 3b*). The set-up of the apparatus, including the size of visual stimuli, monitor, and computer, was identical to *Experiment 1*.

Procedure

We adapted a continuous retrieval task based on a previous experiment involving the color of objects (Sun, Fidalgo, et al., 2017). At the beginning of every trial (n = 330), participants saw a screen with the current trial number. When the space key was pressed, a central fixation cross appeared with "*Remember*" on screen for 1000 ms. Each trial then contained three phases: study, interference (1-back task involving distracting shapes, one of which was repeated), and recall (*Figure 12*), described in detail below.

Prior to the start of the experiment, participants were given detailed instructions and a set of 12 practice trials. Participants were explicitly asked: 1) to make their best guess if they could not remember the target shape, 2) to avoid verbally naming the shapes, 3) to be as accurate as possible, and 4) to press the space key upon identification of the repeating shape, which they were told could occur more than once during a given trial. In reality, only one repeat occurred during the n-back task. This instruction was provided to ensure that participants paid attention throughout the trial so that all interfering shapes were processed.

Study phase. Immediately following the "*Remember*" screen, a target shape was sampled randomly from the 360 shapes corresponding to each degree on VCS space, displayed for 1000 ms on the center of the screen.

Interference phase. After a 250 ms ISI, a grey-scale mask appeared for 280 ms to prevent visual afterimages. After another 250 ms ISI where nothing was presented on the screen, five distractor shapes (350 ms each, ISI 250 ms) were presented sequentially in three possible conditions: distal, proximal, and baseline. As a measure of attention and to ensure that interfering material was processed, one of the distractor shapes repeated consecutively (i.e. n-back task). The position of this repeat was random every trial.

The nature of the interference was varied across three conditions (*Figure 12*). In the *distal* ("subjectively dissimilar") condition, distractor shapes were sampled from a cut of 180 degrees on the opposite side of circular space relative to the target (i.e. 90 degrees to the left and right of a shape 180 degrees away from the target). Distal distractors ranged from approximately 30% to 0% match in subjective similarity with respect to the target (estimated using the VCS distance function, *Figure 10c*). In the *proximal* ("subjectively similar") condition, distractor shapes were sampled from a cut of 180 degrees on the same side of circular space relative to the target (i.e. 90 degrees to the left and right of the target). Proximal distractors ranged from approximately 85% to 30% match in subjective similarity with respect to the target (estimated using the VCS distance function, *Figure 10c*). The 180 degree slices in the distal and proximal conditions were further subdivided into four equal size subdivisions of 30 degrees, separated by "buffer zones" of 15 degrees' distance from each other (buffer zones shown by the white segments in between the darker shaded regions in *Figure 12c*). The four subdivisions were each independently sampled in a random order to produce the overall set of distractor stimuli during each trial. These subdivisions were included to ensure that distractors in the proximal condition were never identical to the target shape and that distractors in both the proximal and distal conditions were always separated by a minimum of 15 degrees' distance. Performance on the distal and proximal conditions was compared to a baseline "control" condition with distractor stimuli that were scrambled lines. These lines were created by scrambling the shape stimuli. Each of the distal, proximal, and baseline conditions contained 100 trials, and condition order was randomly determined.

Thirty trials with "random" interference were also randomly interspersed to provide variability in the task so that participants could not guess the trial type throughout the

experiment. For these trials, distractor shapes were sampled from each quadrant of VCS space relative to the target item. These trials were not included in the statistical analysis because any potential impact of random interference could be due to a combination of effects caused by dissimilar and similar distractors.

Recall phase. The recall phase began with the mouse cursor centered on the screen. When the mouse cursor was moved by the participant, the specific shape that corresponded to the angle on VCS space occupied by the mouse cursor was displayed at the center of the screen (*Figure 12a*). All other shapes were masked. The space was fully continuous, as the shapes were generated by morphing (see a visual example on the Open Science Framework: https://osf.io/d9gyf/). Once a shape was reported through a mouse click, the next trial began.

Across participants, the shapes occupied different locations on the response circle to ensure that there were no systematic relationships between particular shapes and locations. For example, if shape 1 was positioned on the response circle at 1 degree, shape 2 was positioned at 2 degrees, and shape 3 was positioned at 3 degrees for participant A, all the shapes were shifted by a random jitter for participant B (e.g. shape 1 may now be at 41 degrees, shape 2 at 42 degrees, shape 3 at 43 degrees, and so forth). These shape-location mappings were not jittered trial-by-trial because we reasoned that doing so would serve as an additional source of interference because the participant would have to extensively explore the space (and be exposed to uncontrolled interference) prior to being able to select a shape.

Statistical Analysis

Mixture model analysis. Memory detail was characterized using a mixture model (Zhang & Luck, 2008). For every trial, error was first calculated as the signed angular distance

between the target and reported shape. These response errors were bounded from -180 to 180 degrees, with 0 degrees representing an exact memory of the target. The histogram of response errors for each participant was then fit with two parameters using maximum likelihood estimation (*Figure 2*).

If the participant could not remember either fine- or coarse-grained information about the target shape, every shape on circular space would have an equal likelihood of being reported. Thus, a parameter representing the probability of random guesses, g, was modelled as a uniform distribution across response errors. P_{mem} was subsequently defined as $1 - g$, or the probability that some fine- or coarse-grained memory about the target was available. The second parameter, a von Mises distribution, was fit to the response errors centered around 0 degrees' error. The inverse of circular standard deviation ($1/sd$) was defined as the representational quality of a memory, termed precision. In this way, high precision implied a memory with more detail, as participants reported shapes closely overlapping in angular distance to the target shape.

All model fitting analyses were conducted using MemToolbox (Suchow et al., 2013; the 2-component mixture model was implemented using the *StandardMixtureModel* function). Overall, two measures of memory detail were obtained from the mixture model analyses: P_{mem}, reflecting the probability of remembering some memory about the target (i.e. some fine- or coarse-grained information), and *precision*, reflecting the fidelity of the memory trace (i.e. memories with high precision contain more fine-grained information than memories with low precision).

We removed trials with incorrect or missing n-back responses so that only trials in which we could verify that participants were paying attention to distractors were included in the analysis (see *Figure 13a*). Next, the mixture model was fit to the data of every participant

separately for the baseline, distal, and proximal conditions. Each participant had a P_{mem} and precision value for each interference condition. Separate linear mixed models were then used to statistically analyze P_{mem} and precision as a function of interference condition (baseline, distal, proximal). To account for the within-subjects design, a random intercept was modelled for each participant (Pinheiro & Bates, 2000). The linear mixed models were estimated with an unstructured covariance matrix using the *lme* function from the *nlme* package (Pinheiro et al., 2018) in R version 3.4.1 (R Core Team, 2017). Interference conditions were dummy coded in each linear mixed model (West, Aiken, & Krull, 1996), so that we compared the distal to the baseline condition as well as the proximal to the baseline condition. Effect sizes for individual comparisons were calculated using Cohen's d, measured by the difference between condition means divided by the pooled standard deviation across conditions (Lakens, 2013).

Model-free analysis. Next, we examined memory detail as a function of interference condition without fitting the mixture model. We developed a method of analyzing continuous retrieval data without assuming any parameters in the data, in contrast to the mixture model which assumes two kinds of parameters (i.e. P_{mem} and precision, see *Figure 2*). All n-back incorrect responses were first excluded. Across all participants, responses were then binned three ways within each interference condition: 1) fine-grained responses, 2) coarse-grained responses, and 3) responses with minimal fine- and coarse-grained information (*Figure 12c*). Fine-grained responses include regions on circular space 14 degrees on either side of the target, reflecting high fidelity responses of the target shape where no distractors could have been sampled from (approximately 100% to 85% match in subjective similarity; *Figure 10c*). Coarse-grained responses include 15 – 89 degrees' distance with respect to either side of the target on circular space, reflecting responses that overlapped with the positions of distractors sampled from the

proximal condition (approximately 85% to 30% match in subjective similarity; *Figure 10c*). Responses with minimal fine- and coarse-grained information include 91 – 180 degrees' distance with respect to either side of the target on circular space, reflecting responses that overlapped with the positions of distractors sampled from the distal condition (approximately 30% to 0% match in subjective similarity; *Figure 10c*). Responses corresponding to exactly 90 degrees' error overlapped with the distractors sampled from both the distal and proximal conditions and were excluded (7 excluded trials out of a possible 7720 trials in the baseline, distal, and proximal conditions across all participants). Importantly, these three bins enabled us to investigate whether there were systematic differences between interference conditions without fitting any model parameters. The implications of the mixture model analyses compared to our model-free analyses will be discussed in the *Discussion* section.

The proportion of responses made to each bin type was calculated separately for the baseline, distal, and proximal conditions, so that each participant had three values for each interference condition. Critically, the total proportion of responses made to each interference condition was 100% (e.g. the proportion of responses made to fine-grained + coarse-grained + minimal fine- and coarse-grained information added up to 100%), though the proportion of responses made to each bin type *within* an interference condition could vary (e.g. proportionally more responses could be made to the fine-grained compared to the coarse-grained bin). To statistically analyze this data, we used a linear mixed model to examine the proportion of responses made as a function of bin type and interference condition. Note that there could be no main effect between interference conditions at the omnibus level of the model, as the proportion of responses made to each interference condition totalled to 100%. To account for the within-subjects design, a random intercept was modelled for each participant (Pinheiro & Bates, 2000).

Next, we tested a series of planned comparisons to examine the influence of distractor similarity on memory fidelity. Interference conditions were dummy coded (West, Aiken, & Krull, 1996) to compare the distal to the baseline condition as well as the proximal to the baseline condition for each bin type (fine-grained, coarse-grained, and minimal fine- and coarse-grained).

Results

Of the thirty participants recruited, two participants performed below two standard deviations on the n-back task and were thus excluded. Descriptive results for the remaining twenty-eight participants are shown in *Figure 13a*.

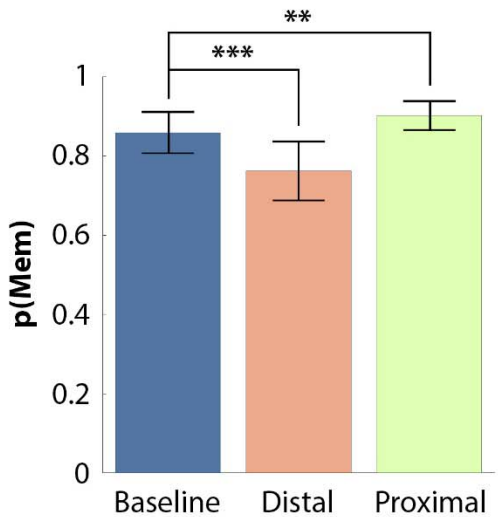
Mixture Model Analysis

We examined both the probability that participants remembered some information about the target shape, P_{mem}, as well as the level of memory detail, precision, as a function of interference condition (baseline, distal, and proximal; *Figure 12*). A linear mixed model revealed that P_{mem} (*Figure 13b, c*) was significantly influenced by interference condition ($F_{2,54}$ = 19.20, p < 0.0001, $partial\ R^2$ = 0.42). Relative to the baseline ("control") condition, P_{mem} was significantly lower in the distal ("subjectively dissimilar") condition (t_{54} = 5.52, p < 0.0001, d = 0.54), but significantly greater in the proximal ("subjective similar") condition (t_{54} = 2.82, p = 0.0067, d = 0.35). The second linear mixed model revealed that precision (*Figure 13d, e*) was also significantly influenced by interference condition ($F_{2,54}$ = 13.50, p < 0.0001, $partial\ R^2$ = 0.33). Notably, however, the nature of the interference affected precision differently than it did P_{mem}. Precision in the distal condition was not significantly different compared to the baseline condition (t_{54} = 0.31, p = 0.76). In contrast, precision was significantly lower in the proximal condition compared to the baseline condition (t_{54} = 4.22, p = 0.0001, d = 0.54).

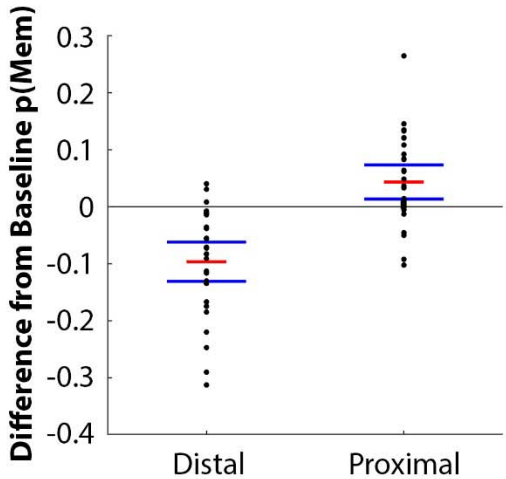
a. Mixture Model Descriptive Results

	P_{mem}			*Precision*			*1-back*		
Condition	Baseline	Distal	Proximal	Baseline	Distal	Proximal	Baseline	Distal	Proximal
Mean	0.86	0.76	0.90	0.065	0.066	0.055	94%	91%	91%
SD	0.14	0.20	0.10	0.020	0.019	0.019	4.6%	8.5%	10.5%

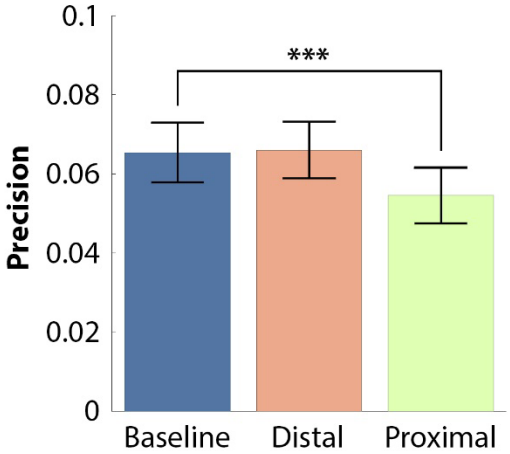
b. P_{mem} by Interference Condition



c. P_{mem} Differences from Baseline



d. Precision by Interference Condition



e. Precision Differences from Baseline

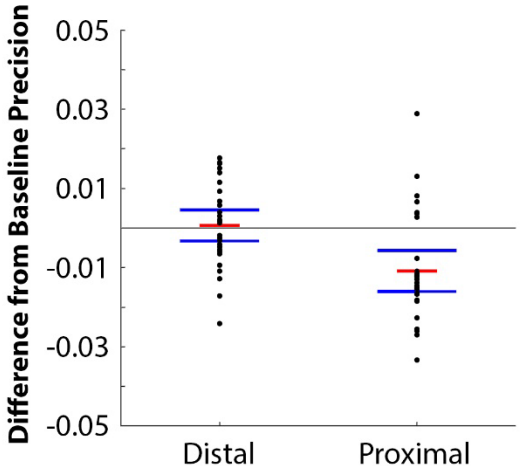


Figure 13. Mixture model analyses. *** denotes $p < 0.001$, whereas ** denotes $p < 0.01$. (*a*) Descriptive results. (*b, c*) P_{mem} was significantly reduced in the distal relative to the baseline condition, and significantly increased in the proximal relative to the baseline condition. (*d, e*) In contrast, precision was not different between the distal relative to the baseline condition, whereas precision was significantly lower in the proximal relative to the baseline condition. Error bars in (*b, d*) reflect the 95% CI for the condition means. Error bars in (*c, e*) reflect the 95% CI for the differences between condition means.

The results from the mixture model suggest that distractor similarity differentially impacted memory detail. Relative to baseline, dissimilar interference decreased P_{mem} with no influence on precision. In other words, dissimilar distractors disrupted both fine- and coarse-grained information, rendering the memory inaccessible. In contrast, similar interference increased P_{mem} but simultaneously decreased precision relative to baseline. That is, similar distractors disrupted fine-grained information but increased the reliance on coarse-grained information, rendering the memory blurred.

Model-free Analysis

The proportions of responses were binned three ways for each condition (*Figure 14*) and a linear mixed model was conducted to examine the influence of distractor similarity on each bin type: 1) fine-grained responses corresponding to regions on circular space 14 degrees on either side of the target (approximately 100% to 85% match in subjective similarity; *Figure 10c*); 2) coarse-grained responses corresponding to 15 – 89 degrees' distance with respect to either side of the target on circular space (approximately 85% to 30% match in subjective similarity; *Figure 10c*); and 3) responses with minimal fine- and coarse-grained information, corresponding to 91 – 180 degrees' distance with respect to either side of the target on circular space (approximately 30% to 0% match in subjective similarity; *Figure 10c*).

The linear mixed model found a significant main effect of bin type ($F_{1,248}$ = 101.89, p < 0.0001, $partial\ R^2$ = 0.32), suggesting that collapsed across interference conditions, participants made responses at different proportions to the bins with fine-grained, coarse-grained, or minimal fine- and coarse-grained information (*Figure 14b*). In contrast, there was no main effect of interference condition collapsed across bin type, as the proportion of responses totalled to 100%

within each condition (see *Statistical Analysis. Model-free analysis*). Next, we conducted planned comparisons to examine the influence of distractor similarity on memory fidelity.

The proportion of fine-grained responses was differentially influenced by interference condition ($F_{2,54}$ = 12.27, p < 0.0001, $partial\ R^2$ = 0.31). The proportion of fine-grained responses was lower in both the distal condition (t_{54} = 3.89, p = 0.0003, d = 0.34) and the proximal condition (t_{54} = 4.47, p < 0.0001, d = 0.34) compared to the baseline condition. Thus, consistent with the mixture model analyses, both distal and proximal interference disrupted fine-grained information.

The proportion of coarse-grained responses was differentially influenced by interference condition ($F_{2,54}$ = 12.90, p < 0.0001, $partial\ R^2$ = 0.32). The proportion of coarse-grained responses was not significantly different between the distal and baseline conditions (t_{54} = 0.37, p = 0.71). In contrast, the proportion of coarse-grained responses were significantly higher in the proximal compared to the baseline condition (t_{54} = 4.81, p < 0.0001, d = 0.59). Thus, consistent with the mixture model analyses, proximal interference increased the reliance on coarse-grained information.

The proportion of responses with minimal fine- and coarse-grained information was differentially influenced by interference condition ($F_{2,54}$ = 15.03, p < 0.0001, $partial\ R^2$ = 0.36). The proportion of responses containing minimal fine- and coarse-grained information was significantly higher in the distal compared to the baseline condition (t_{54} = 4.88, p < 0.0001, d = 0.52). In contrast, the proportion of responses lacking both fine- and coarse-grained information was significantly lower in the proximal compared to the baseline condition (t_{54} = 2.50, p = 0.016, d = 0.20).

a. Model-free Descriptive Results

	Fine-grained			*Coarse-grained*			*Minimal*		
Condition	Baseline	Distal	Proximal	Baseline	Distal	Proximal	Baseline	Distal	Proximal
Mean	0.60	0.54	0.55	0.34	0.34	0.41	0.059	0.11	0.046
SD	0.16	0.18	0.16	0.10	0.094	0.12	0.073	0.12	0.060

b. Fine- and Coarse-grained Responses by Interference Condition

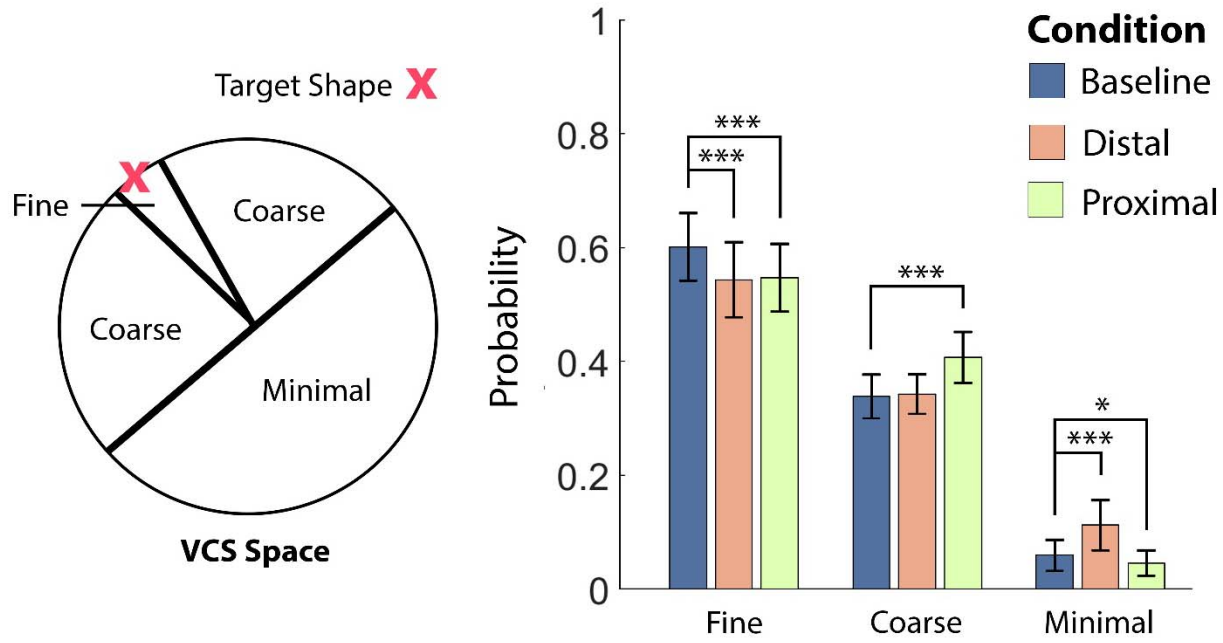


Figure 14. Model-free analysis. *** denotes $p < 0.001$, whereas * denotes $p < 0.05$. (*a*). Descriptive results. Mean refers to the proportion of each type of response (fine-grained, coarse-grained, or minimal fine- and coarse-grained). (*b*). Inferential results. Responses containing fine-grained, coarse-grained, and minimal fine- and coarse-grained information was determined by binning regions of shape space. The probability of reporting a fine-grained response was significantly lower in both the distal and proximal relative to the baseline condition. The probability of reporting a coarse-grained response was significantly higher in the proximal relative to the baseline condition. The probability of reporting a response with minimal fine- and coarse-grained information was significantly higher in the distal relative to the baseline condition, and significantly lower in the proximal relative to the baseline condition. Error bars reflect the 95% CI for the condition means.

Taken together, the model-free analyses were consistent with the mixture model analyses: dissimilar interference disrupted both fine- and coarse-grained information about the target, whereas similar interference disrupted fine-grained information but increased the reliance on coarse-grained information about the target.

Discussion

Here we report the first application of the *Validated Circular Shape Space* (VCS space), whereby the properties of VCS space were applied to precisely measure memory fidelity for shape. Across both mixture model and model-free analyses, we found that interference differentially influenced memory fidelity. Dissimilar distractors disrupted both fine- and coarse-grained information about the target, rendering the memory inaccessible. In contrast, similar distractors disrupted fine-grained but increased the reliance on coarse-grained information about the target, rendering the memory blurred. Overall, these results provide a striking replication of what has been observed in the domain of color (Sun, Fidalgo, et al., 2017), providing evidence for a set of general principles regarding how the nature of interference impacts memory.

Importantly, our results were consistent across both the mixture model analyses which assume two parameters (*Figure 2*) and the model-free analyses which do not assume any parameters. One paper recently suggested that the parameters of the mixture model may not be directly interpretable because of the nonlinear relationship between angular distance and subjective similarity on circular space (Schurgin, Wixted, & Brady, 2018). Indeed, angular distance was not identical to subjective similarity on VCS space (*Figure 10c*). VCS space was designed and validated with two properties in mind (*Figure 1c*): validated so that the space was perceptually uniform (*Figure 7, 8*), and validated so that the magnitude of subjective similarity was approximated by angular distance (*Figure 9, 10*). To account for the nonlinearity on VCS

space, a set of model-free analyses was conducted to directly examine how memory fidelity changed with distractor similarity. The mixture model (*Figure 13*) and model-free analyses (*Figure 14*) were consistent with each other, suggesting that our findings were not driven by the nonlinear relationship between angular distance and subjective similarity on VCS space. Rather, memory fidelity was systematically and differentially altered by the type of interference.

Continuous measures of fidelity like the one used in the present experiment (*Figure 12*) may be more informative than dichotomous yes-no metrics of accuracy, and may be one way to reconcile the seemingly contradictory findings within the interference literature. More specifically, we believe that similarity can either help or hinder memory depending on the kinds of representations needed to perform a task. In tasks when preserved coarse-grained information from similar distractors is sufficient to serve as a reminder of the target item, similarity can benefit memory (Jiang et al., 2016; Lin & Luck, 2009; Mate & Baques, 2009; Morey, 2018; Sanocki & Sulman, 2011). For example, researchers have shown that auditory distraction can help memory when the coarse-grained category information of sounds are cued (Hanczakowski, Beaman, & Jones, 2017). However, in tasks when coarse-grained information cannot be used to disambiguate between the fine-grained representations of distractors and targets, similar distractors can instead impair memory performance (Baddeley, 1964; Baddeley & Dale, 1966; Barense et al., 2012; Cann, McRae, & Katz, 2011; Gordon, Hendrick, & Levine, 2002; Neely & LeCompte, 1999; Roediger & McDermott, 1995; Yeung et al., 2013; Watson & Lee, 2013; Wickelgren, 1965; Wickelgren, 1966; Wickens, 1970). Indeed, when the category is equated across distractors and targets so that coarse-grained information could no longer serve as a categorical reminder of the target, similar auditory distraction caused memory impairment (Hanczakowski, Beaman, & Jones, 2017).

In behavioral experiments, continuous measures of memory have driven a wealth of research findings over the past decade (Ma, Husain, & Bays, 2014). Simultaneously, there has been increasing interest in the neuroanatomical signatures of memory fidelity. Recent neuroimaging experiments suggest that different aspects of complex scenes can be represented concurrently, such as the fine-grained information about the color of objects (Richter et al., 2016) as well as the coarse-grained information about the general environment (Brunec, Moscovitch, & Barense, 2018). Moreover, the distinction between fine- and coarse-grained information for complex scenes are dissociable within regions of the medial temporal lobes and prefrontal cortex (Brunec, Moscovitch, & Barense, 2018; Nilakantan, Bridge, VanHaerents, & Voss, 2018; Richter et al., 2016; Stevenson et al., 2018; Wais et al., 2017; Wais, Montgomery, Stark, & Gazzaley, 2018; Yonelinas, 2013). In some tasks, structures such as the angular gyrus continuously track the amount of fine-grained information represented in memory, whereas other structures such as the hippocampus track memories containing some fine- or coarse-grained detail as opposed to memories with minimal detail (Richter et al., 2016). Supporting these results, other researchers examining navigation and spatial memory have highlighted the especial role of the hippocampus in the representation of different aspects of fine- and coarse-grained information (Brunec et al., 2018; Nilakantan, Bridge, VanHaerents, & Voss, 2018). These experiments have all made fruitful advances to the neuroanatomical understanding of memory, though researchers are currently limited by the kinds of stimuli and tasks that can capture memory fidelity. To this end, we demonstrate that VCS space is not only perceptually uniform, but can also continuously track the subjective similarity of shape (see *Figure 1c* and *Figure 11* for the two kinds of similarity on circular space). These properties enable VCS space to precisely quantify memory fidelity, offering experimenters access to an underexplored stimulus domain

for future work. As VCS space can be used to create complex objects (*Figure 11a*), one direction of novel investigation may involve accessing the neural correlates of high-fidelity objects created from conjunctions of shapes and colors.

More broadly, measuring memory detail can inform understanding of the changes in memory experienced during aging. Interference typically hinders memory in older adults more so than younger adults (Biss, Campbell, & Hasher, 2013; Dalla Barba, Attali, & La Corte, 2010; Ly, Murray, & Yassa, 2013; Norman & Schacter, 1997; Schacter, Koutstaal, & Norman, 1997). This highly replicated pattern of results is thought to reflect the inability to suppress irrelevant information (Campbell, Hasher, & Thomas, 2010), rendering older adults especially vulnerable to interference (Weeks & Hasher, 2014). Supporting this interpretation, the *hyperbinding* effect occurs when older adults experience preserved memory for irrelevant distraction even though memory about the target is disrupted (Biss et al., 2013; Campbell & Hasher, 2018; Campbell, Hasher, & Thomas, 2010; Weeks, Biss, Murphy, & Hasher, 2016). An interesting future direction may be to examine whether there are certain kinds of interference that may be most difficult to suppress in memory (Campbell, Hasher, & Thomas, 2010), such that perhaps irrelevant coarse- as opposed to irrelevant fine-grained information may be most likely to be retained in older adults.

A wealth of research suggests that older adults are more likely to utilize gist-based representations compared to younger adults, perhaps reflecting the compensatory reliance of spared coarse-grained information in response to declining high-fidelity memories (Brainerd & Reyna, 2001; Brainerd & Reyna, 2002; Castel, 2005; Flores et al., 2016; Koutstaal & Schacter, 1997; Paige et al., 2016; Peich, Husain, & Bays, 2013; Pertzov et al., 2014; Tun, Wingfield, Rosen, & Blanchard, 1998; Wais & Gazzaley, 2014). For example, the Deese-Roediger-

McDermott (DRM) procedure involves presenting a list of study words to participants (Roediger & McDermott, 1995). A critical lure that is semantically related to the words within the study list is presented, and participants judge whether this new lure was previously displayed in the study list. In this task, older adults have more false memories compared to younger adults, such that greater proportions of false alarms are made to the semantically related but never displayed lure (Balota et al., 1999; Dehon & Bredart, 2004; Sugrue, Strange, & Hayne, 2009). Intriguingly, perhaps the coarse-grained information reflecting the semantic category of the DRM study list is preserved in older adults, resulting in higher probabilities of false memory (Wilson, Potter & Cowell, 2018). As few studies have directly examined memory fidelity in relation to aging, this avenue may be a natural direction of future investigation.

The existing neuroimaging and aging work demonstrates that memories are not a unitary process whereby observers can only ever remember or forget, but rather memories can range on a continuum of detail (Brunec et al., 2018; Yonelinas, 2013). Continuous measures of fidelity may be an increasingly important method to operationalize memory in future work. To this end, circular stimulus spaces such as VCS space offer a unique and well-controlled method to quantify detail (*Figure 1c*, *Figure 11*). This is not to say that other stimuli types such as naturalistic video sequences or images are not important. Naturalistic stimuli may be more generalizable than the often simple colors, dots, lines, and words typically tested in the laboratory. However, naturalistic stimuli also have their drawbacks, including the lack of precise experimenter control for factors such as subjective similarity. Though time consuming, the development of complex well-controlled stimulus spaces may be one method to include both the needed complexity of naturalistic stimuli and the needed experimenter control of extraneous factors. Indeed, emerging research have begun characterizing highly complex multidimensional

spaces for faces (Busey, 1998; Chang et al., 2017; Hopper et al., 2014; Nestor, Plaut, & Berhmann, 2016), objects (Hout, Goldinger, & Brady, 2014), and methods to expedite the validation of stimulus spaces (Hout, Cunningham, Robbins, & MacDonald, 2018; Kriegeskorte & Mur, 2012).

Mechanisms Underlying Distractor Similarity

Throughout this manuscript, we have suggested that dissimilar distractors disrupt both fine- and coarse-grained information, rendering the memory inaccessible. We have furthermore suggested that similar distractors disrupt fine-grained but increase the reliance on coarse-grained information, rendering the memory blurred. However, it is important to emphasize that we use these terms as mere metaphors for the complex processes governing how interference impacts memory. Though it is clear that distractor similarity differentially impacted memory detail (*Figure 13, 14*), the mechanisms by which memory representations were altered by interference are less clear.

The disrupted fine- and coarse-grained information after dissimilar distractors can not only reflect the erasure of the memory trace (Zhang & Luck, 2009), but can also reflect many types of mechanisms which can render the memory inaccessible. One possibility is that the memory trace cannot be accessed due to competition from distractors at retrieval (Cowell, Bussey, & Saksida, 2006; Sadil & Cowell, 2017; Schurgin, Wixted, & Brady, 2018; Tulving, 1974). On a continuous retrieval task, participants report an item along a response circle (*Figure 12*). During this retrieval period, observers are searching for the target and are subsequently exposed to uncontrollable interference. Some of the responses with minimal fine- and coarse-grained information may therefore reflect the inability to distinguish between distractors and targets (Diana, Peterson, & Reder, 2004; Sadil & Cowell, 2017; Schurgin, Wixted, & Brady,

2018). Moreover, although the representation of the target may be stored in memory, this veridical representation can be temporarily unavailable (Rajsic, Swan, Wilson, & Pratt, 2017; Rose et al., 2016; Tulving, 1974). Memories are not always lost when observers cannot retrieve a memory, as seemingly forgotten memories can be reactivated back into focal attention using a burst of transcranial magnetic stimulation (Rose et al., 2016). In other cases, responses with minimal fine- and coarse-grained information may reflect binding errors, such that participants incorrectly believe a previously displayed distractor was actually the target item ("swap" errors, see Bays, Catalao, & Husain, 2009).

The disrupted fine-grained but increased reliance on coarse-grained information after similar interference was said to reflect "blurring", but this blurring may also reflect many different mechanisms. One possibility is the direct distortion of the target memory by distracting information. Observers can retain the statistical regularities present in many different kinds of stimuli at a brief glance, such that individuals can recall the average color of an array, the average direction of pointed lines, or even the average emotional state of complex faces without fine-grained memory of individual items (Rademaker, Bloem, De Weerd, & Sack, 2015; Whitney & Yamanashi Leib, 2018). Thus, each displayed distractor may be incrementally averaged with the target item during interference. Other possibilities that may also look like blurring include competition from highly similar but non-veridical items due to the uncontrollable interference at retrieval (Cowell, Bussey, & Saksida, 2006; Diana, Peterson, & Reder, 2004; Sadil & Cowell, 2017; Schurgin, Wixted, & Brady, 2018), as well as swap errors made to previously displayed distractors (Bays, Catalao, & Husain, 2009). Additionally, various types of categorical "gist" responses can also appear as blurring (Bae et al., 2015, Hardman, Vergauwe, & Ricker, 2017). Rather than the representation of the target directly distorted by

similar interference, in this case a separate representation of the shared information encapsulating the distractors and targets may be what is preserved. These possibilities can all appear as blurring, and the behavioral manifestations of these mechanisms may be effectively indistinguishable. Intriguingly, perhaps even the theoretical differences between the described mechanisms may in actuality have much in common with each other, providing an interesting direction of future work. Elegant neuroimaging experiments are needed to elucidate the potential neural differences or commonalities between mechanisms.

For these reasons, our methodology was not designed to delineate between the cognitive mechanisms at play during interference. Instead, we combined a set of mixture model (*Figure 13*) and model-free analyses *(Figure 14*) to isolate the *differences* in memory detail after dissimilar and similar interference. Converging analyses revealed that dissimilar distractors increased the likelihood of reporting items very low in detail (i.e. reflecting a decrease in fine- and coarse-grained information). In contrast, similar distractors increased the likelihood of reporting similar but non-veridical items (i.e. reflecting an increased reliance on coarse-grained information). Last, both dissimilar and similar distractors reduced the likelihood of reporting highly veridical items (i.e. reflecting a decrease in fine-grained information). Across both mixture model and model-free analyses, these results confirm that distractor similarity differentially and robustly influenced memory detail.

Conclusion

In *Experiment 1*, we provide the first perceptually uniform shape space whereby angular distance on a 2D circle is a proxy for subjective similarity. Relative to any position on the

Validated Circular Shape Space (VCS space), shapes closer in angular distance tend to be more subjectively similar compared to shapes further in angular distance. These properties of VCS space can be used to not only quantify shape similarity, but to also quantify memory detail along a continuum. Demonstrating the validity of VCS space in *Experiment 2*, VCS space was applied for the first time to quantify both the subjective similarity of distractors and the fidelity of shape memory. Memories were systematically and differentially altered by the similarity of distracting information. Dissimilar distractors disrupted both fine- and coarse-grained information, rendering the memory inaccessible. In contrast, similar distractors disrupted fine-grained but increased the reliance on coarse-grained information, rendering the memory blurred. These findings suggest that in some cases, similar distraction can help memory when preserved coarse-grained information is sufficient to act as a reminder about the target. In other cases, this benefit from similar distraction can revert to a memory impairment when precise fine-grained representations are needed for the task. As these effects were consistent across domains for previous color and current shape findings, we highlight that these results may provide a general set of principles regarding how the nature of interference impacts memories.

References

Baddeley, A. D. (1964). Semantic and acoustic similarity in short-term memory. *Nature, 204*, 1116–1117.

Baddeley, A. D., & Dale, H. C. A. (1966). The effect of semantic similarity on retroactive interference in long- and short-term memory. *Journal of Verbal Learning & Verbal Behavior, 5*, 417–420.

Bae, G.-Y., Olkkonen, M., Allred, S. R., Wilson, C., & Flombaum, J. I. (2014). Stimulus-specific variability in color working memory with delayed estimation. *Journal of Vision, 14*(4), 7.

Bae, G.-Y., Olkkonen, M., Allred, S. R., & Flombaum, J. I. (2015). Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *Journal of Experimental Psychology: General, 144*(4), 744-763.

Balota, D. A., Cortese, M. J., Duchek, J. M., Adams, D., Roediger, H. L., McDermott, K. B., & Yerys, B. E. (1999). Veridical and false memories in healthy older adults and in dementia of the Alzheimer's type. *Cognitive Neuropsychology, 16*(3/4/5), 361-384.

Barense, M. D., Groen, I. I. A., Lee, A. C. H., Yeung, L.-K., Brady, S. M., Gregori, M., … Henson, R. N. A. (2012). Intact memory for irrelevant information impairs perception in amnesia. *Neuron, 75*(1), 157–167.

Bays, P. M., & Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *J Vis, 9*(10), 7.1-711.

Bays, P. M., Wu, E. Y., & Husain, M. (2011). Storage and binding of object features in visual working memory. *Neuropsychologia, 49*(6), 1622-31. doi:10.1016/j.neuropsychologia.2010.12.023

Bell, J., Hancock, S., Kingdom, F. A. A., & Peirce, J. W. (2010). Global shape processing: Which parts form the whole? *J Vis, 10*(6), 16.

Biss, R. K., Campbell, K. L., & Hasher, L. (2013). Interference from previous distraction disrupts older adults' memory. *J Gerontol B Psychol Sco Sci, 68*(4), 558-61.

Biss, R. K., Joan Ngo, K. W., Hasher, L., Campbell, K. L., & Rowe, G. (2013). Distraction can reduce age-related forgetting. *Psychological Science, 24*(4), 448-455.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2011). A review of visual memory capacity: Beyond individual items and towards structured representations. *Journal of Vision, 11*(5), 4. http://doi.org/10.1167/11.5.4

Brainerd, C. J., & Reyna, V. F. (2001). Fuzzy-trace theory: Dual processes in memory, reasoning, and cognitive neuroscience. *Advances in Child Development and Behavior, 28*, 41-100.

Brainerd C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11*(5), 164–169.

Brouwer, G. J., & Heeger, D. J. (2009). Decoding and Reconstructing Color from Responses in Human Visual Cortex. *Journal of Neuroscience, 29*(44).

Brunec, I. K., Bellana, B., Ozubko, J. D., Man, V., Robin, J., Liu, Z-X, …, & Moscovitch, M. (2018). Multiple scales of representation along the hippocampal anteroposterior axis in humans. *Current Biology, 28*(13), P2129-2135.

Brunec, I. K., Moscovitch, M. M., & Barense, M. D. (2018). Boundaries shape cognitive representations of spaces and events. *Trends in Cognitive Sciences, 22*(7), 637-650.

Busey, T. A. (1998). Physical and psychological representations of faces: Evidence from morphing. *Psychological Science, 9*(6), 476-483.

Campbell, K. L., & Hasher, L. (2018). Hyper-binding only apparent under fully implicit test conditions. *Psychology and Aging, 33*(1), 176-181.

Campbell, K. L., Hasher, L., & Thomas, R. C. (2010). Hyper-binding: A unique age effect. *Psychological Science, 21*(3), 399-405.

Cann, D. R., McRae, K., & Katz, A. N. (2011). False recall in the Deese-Roediger-McDermott paradigm: The roles of gist and associative strength.

Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *J Neuro, 32*(22), 7685-7700.

Castel, A. D. (2005). Memory for grocery prices in younger and older adults: The role of schematic support. *Psychology and Aging, 20*(4), 718-721.

Chang, C-H, Nemrodov, D., Lee, A. C. H., & Nestor, A. (2017). Memory and perception-based facial image reconstruction. *Scientific Reports, 7*, 6499.

Charest, I., Kievit, R. A., Schmitz, T. W., Deca, D., & Kriegeskorte, N. (2014). Unique semantic space in the brain of each beholder predicts perceived similarity. *Proc Natl Acad Sci, 111*(4), 14565-70.

Cheung, V. (2016). Uniform color spaces. *Handbook of Visual Display Technology*. Springer International Publishing. doi:10.1007/978-3-319-14346-0_14

Corey, D. M., Dunlap, W. P., & Burke, M. J. (1998). Averaging correlations: Expected values and bias in combined Pearson rs and Fisher's z Transformations. *The Journal of General Psychology, 125*(3), 245-261.

Cowell, R. A., Bussey, T. J., & Saksida, L. M. (2006). Why does brain damage impair memory? A connectionist model of object recognition memory in perirhinal cortex. *The Journal of Neuroscience, 26*(47), 12186 – 12197.

Dalla Barba, G., Attali, E., & La Corte, V. (2010). Confabulation in healthy aging is related to interference of overlearned, semantically similar information on episodic memory recall. *Journal of Clinical and Experimental Neuropsychology, 32*(6), 655-660.

Dehon, H., & Bredart, S. (2004). False memories: Young and Older adults think of semantic associates at the same rate but young adults are more successful at source monitoring. *Psychol Aging, 19*(1), 191-7.

Drucker, D. M., & Aguirre, G. K. (2009). Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cereb Cortex, 19*(10), 2269-80.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*, 1149-1160.

Flores, C. C., Hargis, M. B., McGillivray, S., Friedman, M. C., & Castel, A. D. (2017). Gist-based memory for prices and "better buys" in younger and older adults, *Memory*, *25*(4), 565-573, DOI: 10.1080/09658211.2016.1197944

Fougnie, D., & Alvarez, G. A. (2012). Object features fail independently in visual working memory: Evidence for a probabilistic feature-store model. *J Vis*, *11*(12), 10.1167/11.12.3 3.

Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. *Current Opinion in Neurobiology*, *20*(2), 177–182. https://doi.org/10.1016/j.conb.2010.03.005

Gordon, P. C., Hendrick, R., & Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, *13*(5), 425-30.

Hanczakowski, M., Beaman, C. P., & Jones, D. M. (2017). When distraction benefits memory through semantic similarity. *Journal of Memory & Language*, *94*, 61-74.

Hardman, K. O., Vergauwe, E., & Ricker, T. J. (2017). Categorical working memory representations are used in delayed estimation of continuous colors. *Journal of Experimental Psychology: Human, Perception, and Performance*, *43*(1), 30-54.

Haushofer, J., Livingstone, M. S., & Kanwisher, N. (2008). Multivariate patterns in object-selective cortex dissociate perceptual and physical shape similarity. *PLoS Biol.*, *6*(7): e187.

Hopper, W. J., Finklea, K. M., Winkielman, P., & Huber, D. E. (2014). Measuring sexual dimorphism with a race-gender face space. *Journal of Experimental Psychology, Human Perception, and Performance, 40*(5), 1779-1788.

Hout, M. C., Cunningham, C. A., Robbins, A., & MacDonald, J. (2018). Simulating the fidelity of data for large stimulus set sizes and variable dimension estimation in multidimensional scaling. *SAGE Open*, 1 – 5. https://doi.org/10.1177/2158244018773143

Hout, M. C., Godwin, H. J., Fitzsimmons, G., Robbins, A., Menneer, T., & Goldinger, S. D. (2016). Using multidimensional scaling to quantify similarity in visual search and beyond. *Atten Percept Psychophys, 78*(1), 3-20.

Hout, M. C., Goldinger, S. D., & Ferguson, R. W. (2013). The versatility of SpAM: a fast, efficient, spatial method of data collection for multidimensional scaling. *J Exp Psychol Gen, 142*(1), 256-281.

Hout, M. C, Papesh, M. H., Goldinger, S. D. (2013). Multidimensional scaling. *WIREs Cognitive Science, 4*(1), 93-103.

Jozwik, K. M., Kriegeskorte, N., & Mur, M. (2016). Visual features as stepping stones toward semantics: Explaining object similarity in IT and perception with non-negative least squares. *Neuropsychologia, 83*, 201-226.

Kahana, M. J., Zhou, F., Geller, A. S., & Sekuler, R. (2007). Lure-similarity affects visual episodic recognition: Detailed tests of a noisy exemplar model. *Mem Cognit, 35*(6), 1222-1232.

Kleiner, M., Brainard, D., Pelli, D. (2007). “What’s new in Psychtoolbox-3?”. *Perception, 36*. ECVP Abstract Supplement.

Kriegeskorte, N., & Mur, M. (2012). Inverse MDS: Inferring dissimilarity structure from multiple item arrangements. *Front Psychol, 3*: 245.

Koutstaal, W., & Schacter, D. L. (1997). Gist-based false recognition of pictures in older and younger adults. *Journal of Memory and Language, 37*(4), 555-583.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for *t*-tests and ANOVAs. *Front Psychol, 4*: 863.

Larkey, L. B., & Markman, A. B. (2005). Processes of similarity judgment. *Cognitive Science, 29*(6), 1061-1076.

Lin, P.-H., & Luck, S. J. (2009). The influence of similarity on visual working memory representations. *Visual Cognition, 17*(3), 356–372.

Loffler, G. (2008). Perception of contours and shapes: Low and intermediate stage mechanisms. *Vision Res, 48*(20), 2106-27. doi:10.1016/j.visres.2008.03.006

Logie, R. H., Della, S. S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *Q J Exp Psychol A, 53*(3), 626-46.

Ly, M., Murray, E., & Yassa, M. A. (2013). Perceptual versus conceptual interference and pattern separation of verbal stimuli in young and older adults. *Hippocampus, 23*(6), 425-30.

Ma, W. J., Husain, H., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience, 17*(3), 347-356.

Martin, C. B., Douglas, D., Newsome, R. N., Man, L. L. Y., & Barense, M. D. (2018). Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *eLIFE, 7*:e31873.

Mate, J., & Baques, J. (2009). Visual similarity at encoding and retrieval in an item recognition task. *The Quarterly Journal of Experimental Psychology, 62*(7), 1277-1284.

McDermott, K. C., & Webster, M. A. (2012). Uniform color spaces and natural image statistics. *J Opt Soc Am A Opt Image Sci Vis, 29*(2), A182-A187.

Morey, C. C. (2018). Perceptual grouping boosts visual working memory capacity and reduces effort during retention. *Br J Psychol*. doi:10.1111/bjop.12355

Neely, C. B., & LeCompte, D. C. (1999). The importance of semantic similarity to the irrelevant speech effect. *Memory & Cognition, 27*(1), 37-44.

Nestor, A., Plaut, D. C., & Behrmann, M. (2016). Feature-based face representations and image reconstruction from behavioral and neural data. *PNAS, 113*(2), 416-421.

Nilakantan, A. S., Bridge, D. J., VanHaerents, S., & Voss, J. L. (2018). Distinguishing the precision of spatial recollection from its success: Evidence from healthy aging and unilateral mesial temporal lobe resection. *Neuropsychologia, 119*, 101-106.

Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition, 25*(6), 838-848.

Op de Beeck, H. P., Torfs, K., & Wagemans, J. (2008). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *J Neurosci, 28*(40), 10111-10123.

Osserman, R. (1978). The isoperimetric inequality. *Bulletin of the American Mathematical Society, 84*(6), 1182 – 1238.

Paige, L. E., Cassidy, B. S., Schacter, D. L., & Gutchess, A. H. (2016). Age differences in hippocampal activation during gist-based false recognition. Neurobiology of aging, 46, 76-83.

Pasupathy, A. (2006). Neural basis of shape representation in the primate brain. *Progress in Brain Research, 154*(A), 293-313.

Peich, M. C., Husain, M., & Bays, P. M. (2013). Age-related decline of precision and binding in visual working memory. *Psychol Aging, 28*(3), 729-43.

Pertzov, Y., Heider, M., Liang, Y., & Husain, M. (2014). Effects of healthy ageing on precision and binding of object location in visual short term memory. *Psychology and aging, 30*(1), 26-35.

Pertzov, Y., & Husain, M. (2014). The privileged role of location in visual working memory. *Atten Percept Psychophys, 76*(7), 1914-24. doi:10.3758/s13414-013-0541-y.

Pinheiro, J., & Bates, D. (2000). Mixed-effects models in S and S-PLUS. Springer-Verlag New Tork.

Pinheiro J., Bates D., DebRoy S., Sarkar, D., & R Core Team (2018). nlme: Linear and nonlinear mixed effects models. R package version 3.1-137, https://CRAN.R-project.org/package=nlme.

Pointer, M. R. (1981). A comparison of the CIE 1976 colour spaces. *Color Research and Application,* 108-118.

R Core Team. (2017). R: A language and environment for statistical computing. https://www.R-project.org/

Rademaker, R. L., Bloem, I. M., De Weerd, P., & Sack, A. T. (2015). The impact of interference on short-term memory for visual orientation. *Journal of Experimental Psychology: Human Perception and Performance, 41*(6), 1650–1665.

Rajsic, J., Swan, G., Wilson, D. E., & Pratt, J. (2017). Accessibility limits recall from visual working memory. *J Exp Psychol Learn Mem Cogn, 43*(9), 1415-1431.

Richter, F. R., Cooper, R. A., Bays, P. M., & Simons, J. S. (2016). Distinct neural mechanisms underlie the success, precision, and vividness of episodic memory. *eLife, 5*: e18260.

Robertson, A. R. (1977). The CIE 1976 color-difference formulae. *Color Research and Application, 2*(1), 7-11.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(4), 803-814.

Rose, N. S., LaRocque, J. J., Riggal, A. C., Gosseries, O., Starrett, M. J., Meyering, E. E., & Postle, B. R. (2016). Reactivation of latent working memories with transcranial magnetic stimulation. *Science, 354*(6316), 1136-1139.

Sadil, P. S., & Cowell, R. A. (2017). A computational model of perceptual and mnemonic deficits in medial temporal lobe amnesia. *Journal of Cognitive Neuroscience, 29*(6), 1075 – 1088.

Salmela, V. R., Henriksson, L, & Vanni, S. (2016). Radial frequency analysis of contour shapes in the visual cortex. *PLoS Comput Biol, 12*(2), e1004719.

Sanocki, T., & Sulman, N. (2011). Color relations increase the capacity of visual short-term memory. *Perception, 40*(6), 635-648.

Schacter, D. L., Koutstaal, W., & Norman, K. A. (1997). False memories and aging. *Trends in Cognitive Sciences, 1*(6), 229-236.

Schmidtmann, G., & Fruend, I. (2019). Radial frequency patterns describe a small and perceptually distinct subset of all possible planar shapes. *Vision Research, 154*, 122-130.

Schneegans, S., & Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. *The Journal of Neuroscience, 37*(14), 3913–3925. doi: https://doi.org/10.1523/JNEUROSCI.3493-16.2017

Schurgin, M. W., Wixted, J. T., & Brady, T. F. (2018). Psychophysical scaling reveals a unified theory of visual memory strength. *bioRxiv, 325472*. doi: https://doi.org/10.1101/325472

Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science, 210*(4468), 390-398.

Sims, C. R., Jacobs, R. A., & Knill, D. C. (2012). An ideal observer analysis of visual working memory. *Psychological Review, 119*(4), 807–30.

Stevenson, R. F., Zheng, J., Mnatsakanyan, L., Vadera, S., Knight, R. T., Lin, J. J., & Yassa, M. A. (2018). Hippocampal CA1 gamma power predicts the precision of spatial memory judgments. *PNAS.* Advanced Online Publication. https://doi.org/10.1073/pnas.1805724115

Suchow, J. W., Brady, T. F., Fougnie, D., & Alvarez, G. A. (2013). Modeling visual working memory with the MemToolbox. *Journal of Vision, 13*(10), 9.

Sugrue, K., Strange, D., & Hayne, H. (2009). False memories in the DRM paradigm: age-related differences in lure activation and source monitoring. *Exp Psychol, 56*(5), 354-60.

Sun, S. Z., Fidalgo, C., Barense, M. D., Lee, A. C. H., Cant, J. S., & Ferber, S. (2017). Erasing and blurring memories: The differential impact of interference on separate aspects of forgetting. *Journal of Experimental Psychology: General, 146*(11), 1606 – 1630.

Tulving, E. (1974). Cue-dependent forgetting: When we forget something we once knew, it does not necessarily mean that the memory trace has been lost; it may only be inaccessible. American Scientist, 62, 74 – 82.

Tun P. A., Wingfield, A., Rosen, M. J., Blanchard, L. (1998). Response latencies for false memories: gist-based processes in normal aging. *Psychology and Aging, 13*(2), 230.

Verheyen, S., Voorspoels, W., Vanpaemel, W., & Storms, G. (2016). Caveats for the spatial arrangement method: Comment on Hout, Goldinger, and Ferguson (2013). *Journal of Experimental Psychology: General, 145*(3), 376-382. http://dx.doi.org/10.1037/a0039758

Wais, P. E., Jahanikia, S., Steiner, D., Stark, C. E. L., & Gazzaley, A. (2017). Retrieval of high-fidelity memory arises from distributed cortical networks. *NeuroImage, 149*(1), 178-189.

Wais, P. E., & Gazzaley, A. (2014). Distractibility during retrieval of long-term memory: domain-general interference, neural networks and increased susceptibility in normal aging. *Frontiers in psychology, 5*, 280. doi:10.3389/fpsyg.2014.00280

Wais, P. E., Montgomery, O., Stark, C. E. L., & Gazzaley, A. (2018). Evidence of a causal role for mid-ventrolateral prefrontal cortex based functional networks in retrieving high-fidelity memory. *Sci Rep, 8*, 14877.

Wardle, S. G., Kriegeskorte, N., Grootswagers, T., Khaligh-Razavi, S., & Carlson, T. A. (2016). Perceptual similarity of visual patterns predicts dynamic neural activation patterns measured with MEG. *NeuroImage, 132*(15), 59-70.

Watson, H. C., & Lee, A. C. H. (2013). The perirhinal cortex and recognition memory interference. *Journal of Neuroscience, 33*(9).

Weeks, J. C., Biss, R. K., Murphy, K. J., & Hasher, L. (2016). Face-name learning in older adults: A benefit of hyper-binding. *Psychon Bull Rev, 23*(5), 1559-1565.

Weeks, J. C., & Hasher, L. (2014). The disruptive - and beneficial - effects of distraction on older adults' cognitive performance. *Frontiers in Psychology, 5*, 133. doi:10.3389/fpsyg.2014.00133

West, S. G., Aiken, L. S., & Krull, J. L. (1996). Experimental personality designs: Analyzing categorical by continuous variable interactions. *Journal of Personality, 64*(1), 1-48.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology, 69*, 105-129.

Wickelgren, W. A. (1965). Short-term memory for phonemically similar lists. *The American Journal of Psychology, 78*(4), 567-574.

Wickelgren, W. A. (1966). Phonemic similarity and interference in short-term memory for single letters. *Journal of Experimental Psychology, 71*(3), 396.

Wickens, D. D. (1970). Encoding categories of words: Empirical approach to meaning. *Psychological Review, 77*, 1–15.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *J Vis, 4*(12), 1120-35.

Wilson, D. M., Potter, K. W., & Cowell, R. A. (2018). Recognition memory shielded from semantic but not perceptual interference in normal aging. *Neuropsychologia, 119*, 448-463.

Wilson, H. R., & Wilkinson, F. (2015). From orientations to objects: Configural processing in the ventral stream. *J Vis, 15*, 4.

Xiberpix. (2008). SqirlzMorph. Retrieved from http://www.xiberpix.net/.

Yeung, L. K., Ryan, J. D., Cowell, R. A., & Barense, M. D. (2013). Recognition memory impairments caused by false recognition of novel objects. *J Exp Psychol Gen, 142*(4), 1384-97.

Yonelinas, A. P. (2013). The hippocampus supports high-resolution binding in the service of perception, working memory, and long-term memory. *Behav Brain Res, 254*, 34-44.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature, 453*(7192), 233–235.

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychol Sci, 20*(4), 423 – 428.