Cross-category material interpolation and binocular material fusion

Hua-Chun Sun[1,2,*], Roland W. Fleming[1,2]

[1]Justus Liebig University Giessen, Otto-Behaghel-Str. 10, 35394 Giessen, Germany
[2]Center for Mind, Brain and Behavior (CMBB), University of Marburg, Justus Liebig University Giessen and TU Darmstadt, Germany

Abstract

Recognizing and identifying materials is essential for navigating the visual world. In this study, we investigate perceptual scaling, material interpolation, and binocular combination of materials across three experiments using cross-category image morphs derived from deep neural network (DNN)-based interpolation algorithm [1]. Twenty-four real-world material images from the STUFF dataset [2] were selected to create 12 cross-category morph pairs (e.g., moss–hair). By systematically adjusting the morph weights, each image gradually transitioned from one material to another, producing a continuum of intermediate blended materials. In Experiment 1, we examined the perceptual scaling of these synthesized blends using a rating task. Results revealed that participants' perceptual judgments generally followed a linear relationship with the interpolation weights, though the degree of compression and nonlinearity varied across different morph pairs. In the following experiments, we employed matching tasks in which participants adjusted a test stimulus along 49 morphing steps (ranging from 2% to 98%) to achieve perceptual equivalence with a reference. In Experiment 2, participants adjusted the test stimulus to match the perceived midpoint blend of two original materials presented one the two sides of the test. In Experiment 3, the adjustment aimed to match the perceptual outcome of dichoptic viewing—where each eye was presented with a different weighted combination of a material pair (e.g., 30% sand + 70% grass in the left eye and 70% sand + 30% grass in the right eye). We found that participants' adjustments deviated systematically from the 50% interpolation midpoint across different material pairs in both experiments. Image statistics revealed that RMS contrast was the primary predictor, accounting for a substantial portion of the variance in both tasks. These findings suggest that perceptual interpolation across materials and binocular material integration may rely on a shared scaling mechanism within a common representational space.

Keywords

material perception, perceptual scaling, image statistic, binocular integration

1. Introduction

Material properties provide essential information about the nature and status of objects, enabling us to identify what we see and guiding how we physically interact with our surroundings. This perceptual information supports object recognition and informs motor planning. Despite its importance in daily life, the mechanisms underlying material perception are still not fully understood.

Current theories suggest that visual material information is encoded in continuous, multidimensional feature spaces in the brain [2, 3]. These spaces are thought to represent both familiar and novel materials and this organization facilitates efficient recognition and discrimination. However, a key challenge remains: how do we perceive and interpret materials that fall between established categories—those that lie in intermediate regions of the material space (e.g., between wood and metal)? While past studies have often concentrated on isolated material properties, our understanding of how full materials are represented in a multidimensional perceptual space is still limited.

Another knowledge gap concerns how material information is integrated binocularly. Binocular vision provides crucial cues for perceiving material properties, yet much of the existing research has

MANER Conference Mainz/Darmstadt 2025 (MANER2025), August 29, 2025, Darmstadt, Germany
*Corresponding author.
hua-chun.sun@psychol.uni-giessen.de (H. Sun); roland.w.fleming@psychol.uni-giessen.de (R. W. Fleming)
https://sites.google.com/view/huachunsun/home (H. Sun); https://www.allpsych.uni-giessen.de/fleminglab/ (R. W. Fleming)
0000-0002-8724-7844 (H. Sun); /0000-0001-5033-5069 (R. W. Fleming)

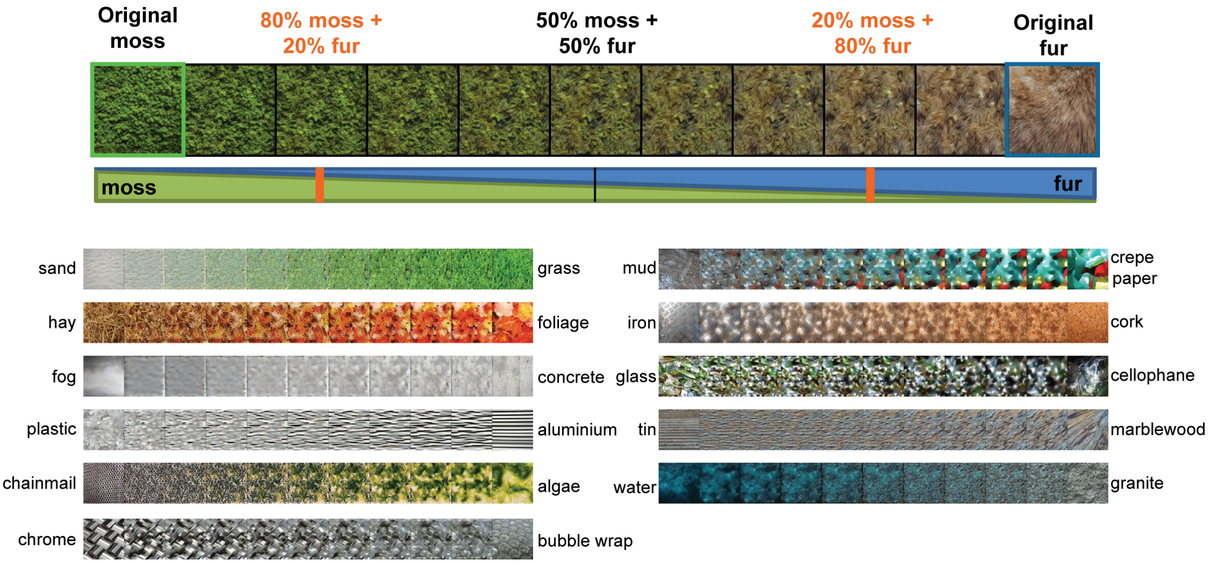

Figure 1: Example material morph stimuli. The top row illustrates a gradual material interpolation from moss to fur in 10% increments from left to right. Eleven additional morph pairs are shown in the bottom row (sand–grass, hay–foliage, mud–crepe paper, iron–cork etc). In total, 12 morph pairs were used in this study.

focused primarily on gloss perception—showing that the visual system can use binocular disparities in specular reflections to infer surface gloss. However, it remains unclear how binocular processing contributes to the perception of other material dimensions. Specifically, how does the brain integrate conflicting material information presented separately to each eye? Addressing this question may not only clarify the mechanisms underlying binocular material perception but also provide insight into broader processes such as the integration of naturalistic, chromatic signals across the eyes [4] and the neural balance between binocular fusion and rivalry [5].

In this study, we address these knowledge gaps by examining how the visual system integrates material information that spans across established material category boundaries. Using deep learning based image interpolation, we investigate both the internal perceptual scaling and the processes by which different material information presented separately to each eye is integrated to support coherent material perception.

2. General Methods

We employ image interpolation based on the deep convolutional neural network (CNN) [1] to examine the perceptual scaling of synthesized material mixtures derived from images of real-world materials. We selected 24 natural images from the STUFF dataset [2] , each representing a distinct material category, to generate 12 cross-category morph pairs (e.g. moss-fur, iron-cork, see Figure 1). All digital images depict fronto-parallel views of material textures under unknown lighting conditions. In total 49 morphing steps (ranging from 2% to 98%) were generated for each pair of morphs. The interpolation was obtained with the Wasserstein loss using VGG19 pretrained weights [1].

The experimental procedures were approved by the Ethics Committee of Justus Liebig University Giessen and conducted in accordance with institutional guidelines and the Declaration of Helsinki. Participants were recruited through the SONA human subject management system of the Department of Psychology and Sport Science at Justus Liebig University Giessen. The recruitment and data handling procedures comply with European Union regulations on research ethics and data protection. All participants provided written informed consent prior to participation.

In addition to behavior measurements, here we also evaluate the interpolation algorithm through analysis of image statistics—specifically Root mean square (RMS) contrast and hue, saturation, and value (HSV) values—to determine whether these features varied systematically with the generated

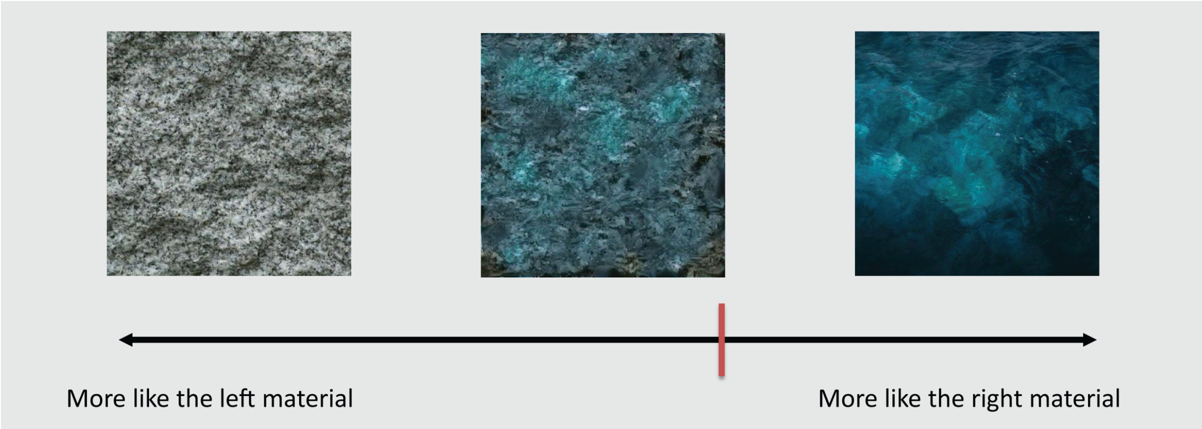

Figure 2: Example trial from Experiment 1. Participants viewed three images arranged horizontally—two original material images on the left (granite) and right (water), and a morphed image in the center (50-50 morph in this example). They rated the central image based on its material similarity to the originals using a continuous scale.

material morph images. RMS contrast is calculated as the standard deviation of pixel intensities divided by their mean. HSV values are computed as the mean of pixel values within each channel.

3. Experiment 1

3.1. Experiment 1 design and task

To quantify the perceptual scaling of CNN-generated interpolation weights, 29 participants completed a rating task in which they evaluated 9 material blends ranging from 10% to 90% of two original materials in each pair. In each trial, participants used a red slider—controlled via mouse or keyboard—to indicate whether the centrally presented morph image resembled the original material shown on the left and right (see Figure 2). Each participant completed 108 trials (12 morph pairs × 9 morph levels), presented in randomized order.

3.2. Experiment 1 Results

Our findings demonstrate a linear relationship in general between the perceived material scale and the interpolation weights (Figure 3, main panel). Perceptual scaling closely tracked the physical morph levels, as indicated by a strong correlation between the two ($r = 0.95$, $p < .001$). However, the regression slope revealed perceptual compression of approximately 77%, suggesting that internal scaling was more compact than the external morphing scale.

Notably, we observed that perceptual scaling varied across the 12 morph pairs, warranting a more nuanced interpretation. For example, morph pair 1 (moss–fur; see Figure 3, subpanel 1) exhibited an almost perfect linear correspondence between physical and perceptual scales (r=0.9), with minimal compression (98%, as indicated by slope). In contrast, morph pairs such as glass–cellophane (pair 5), plastic–aluminium (pair 6), and chrome–bubble wrap (pair 7) showed weaker linear relationships (r 0.7) and large compression effects, with slopes indicating approximately 60% of the physical scale range.

Interestingly, the slope values showed a strong correlation with the saturation difference between the original material images (r = 0.76, $p < .01$). Likewise, the correlation between physical and perceptual scales was also significantly associated with saturation differences (r = 0.63, $p < .05$). These results suggest that chromatic properties of image statistics may play a key role in estimating perceptual midpoint between materials.

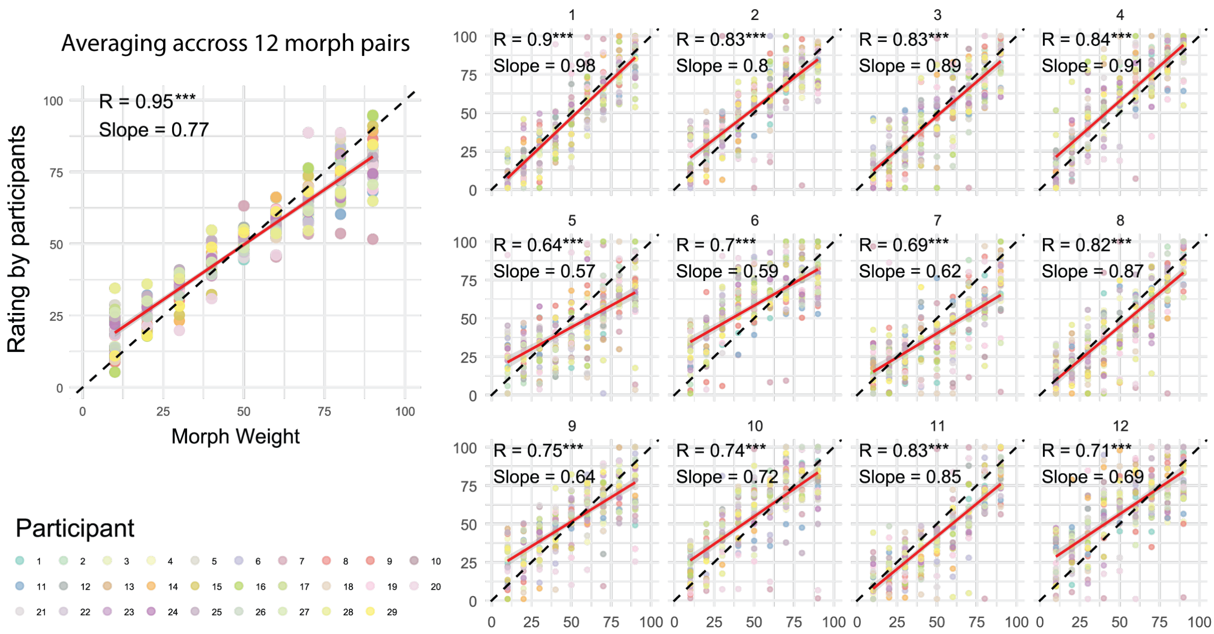

Figure 3: Results from Experiment 1 rating task. Participants rated morph images composed of two original materials, with morph weights ranging from 10% to 90%. In each trial, they judged the perceived similarity of the central morph to the two original materials shown on either side. The main panel (left) shows the average ratings across all 12 morph pairs, while the smaller panels (right) display individual results for each pair. Each colored dot represents an individual participant's response. For each morph pair, participant responses across all nine morph levels were fitted with a linear regression model. The red lines indicate the best-fitting regression lines, with shaded bands representing the standard error. Slope values were derived from these regression lines, and Pearson's r values indicate the strength of the correlation between presented morph weight (x-axis) and participant rating (y-axis). *** denotes statistical significance at $p < .001$.

4. Experiment 2

4.1. Experiment 2 design and task

In Experiment 2, we further employed matching tasks to measure the perceived midpoint blend of two original materials. 27 participants from Experiment 1 also participated in Experiment 2. The experiment setup was similar—two original material images were shown on the left and right, with a test stimulus displayed in the center (Figure 2). This time, participants adjusted the middle test stimulus along 49 morphing steps (ranging from 2% to 98%) to achieve a perceptual midpoint blend between the two original materials. One participant was excluded from further analysis due to an excessive number of extreme responses—over 98% of trials showed deviations greater than ±45% from the midpoint (the maximum possible range being ±50%).

4.2. Experiment 2 Results

Figure 4 presents the adjustment results from 26 participants for each of the 12 morph pairs. The data are shown as the percentage deviation from the 50% interpolation midpoint, ranging from 2.6% to 16.9% across morph pairs, with a mean deviation of 8.81%. Notably, these deviations were significantly correlated with the RMS contrast difference between the original material images ($r = 0.67$, $p < .05$), suggesting that low-level image statistics may play a role in perceptual material interpolation.

5. Experiment 3

Experiments 1 and 2 aim to investigate perceptual scaling and material space across different material categories under ordinary, non-stereoscopic viewing conditions. In Experiment 3, we extend this

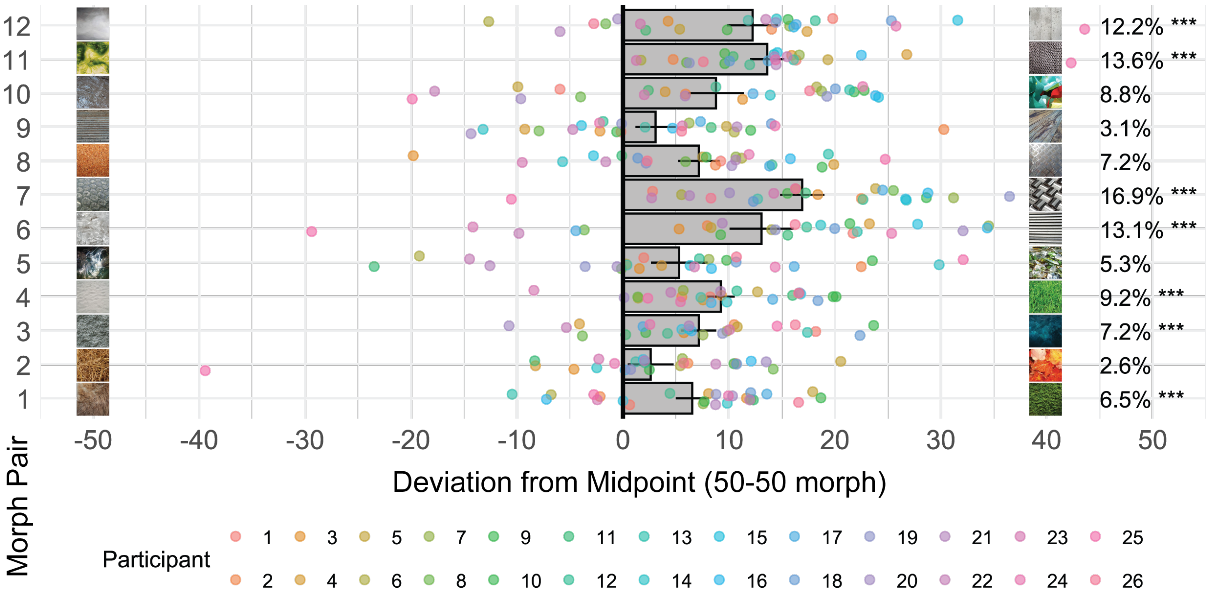

Figure 4: Experiment 2 adjustment task results. Each bar is flanked by the original material images used in that pair. Participants adjusted the test stimulus to match their perceived midpoint between the two materials. The bars show the average percentage deviation from the 50–50 morph point of the test stimulus. The direction of each bar has been aligned to reflect a bias toward the material shown on the right. Individual participant data are represented as color dots. Error bars indicate the standard error of the mean. *** indicates a statistically significancy at the $p < .001$ level (one-sample t-test against zero)

investigation to binocular viewing. The central question is: how does the visual system perceive and interpret materials when each eye receives conflicting material information? Does binocular integration of different materials rely on a similar mechanism as perceiving a single, morphed image composed of the two materials?

5.1. Experiment 3 design and task

Observers (N=21) with normal or adjusted-to-normal visual acuity took part in the experiment. Stimuli were presented dichoptically using a stereoscope and consisted of two types: a reference stimulus and a match stimulus.

The reference stimulus was displayed in the top row of the screen (Figure 5). Each eye was shown a different weighted blend of a given material morph pair (e.g., 90% moss + 10% fur to the left eye and 10% moss + 90% fur to the right eye in Figure 5 top, referred to as the 10–90 condition). Image-eye assignments were counterbalanced across trials, with each image presented equally often to the left and right eyes. Four interocular morph weight conditions were tested: 40–60, 30–70, 20–80, and 10–90. The match stimulus, presented in the bottom row of the screen, was identical in both eyes and served as the comparison for perceptual matching (Figure 5).

Each participant completed 192 trials, comprising 12 morph pairs × 4 morph weight conditions × 4 repetitions, presented in randomized order. As in Experiment 2, participants adjusted the match stimulus along 49 morph steps between the original images to achieve perceptual equality with the reference stimulus. Trials with unfusable or rivalry perception are skipped and removed from analysis.

5.2. Experiment 3 Results

The adjustment results deviated from the 50% interpolation midpoint by 0% to 17% on average (Figure 6 main panel), depending on the material pair. Greater midpoint deviations were observed in the most extreme binocular difference condition (10–90), with a mean deviation of 14.44% across the 12 morph pairs. The magnitude of deviation progressively decreased as the interocular difference diminished: 11.38% for the 20–80 condition, 6.40% for 30–70, and 4.00% for 40–60 (Figure 6, top panels).

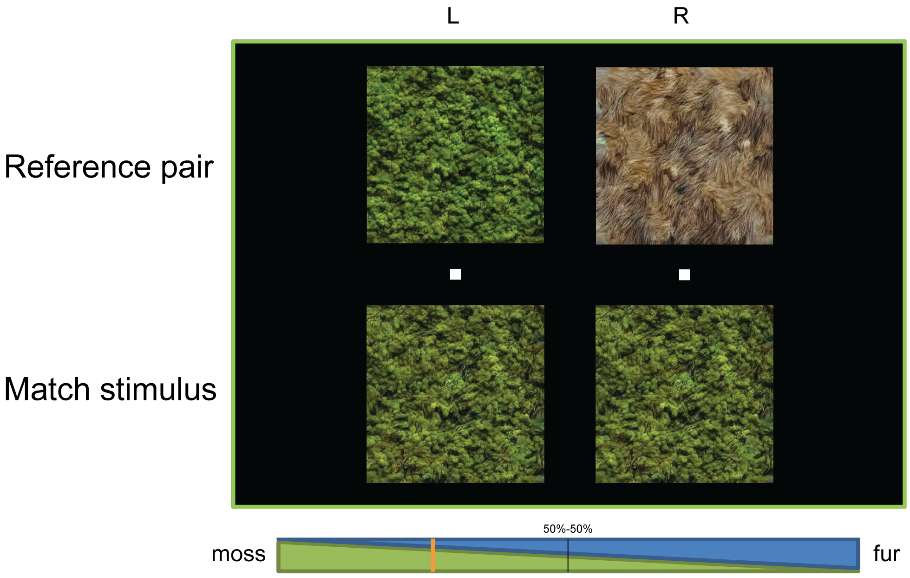

Figure 5: Example trial from Experiment 3. Participants viewed two pairs of images binocularly using a stereoscope to enable binocular fusion. The reference stimulus, presented in the top row, consisted of a moss–fur morph pair with a 10–90 weight condition (i.e., a 90% moss–10% fur morph presented to one eye and a 10% moss–90% fur morph to the other eye). As in Experiment 2, participants adjusted the match stimulus displayed in the bottom row, selecting from 49 morph levels ranging from 2% to 98%. In this example, perceptual equality was achieved at a 72% moss–28% fur morph showing in the bottom row.

Using RMS contrast and saturation, hue, and value (from the HSV color space) as predictors, we conducted a stepwise regression analysis to assess their contribution to the adjustment results. For the 10–90 morph weight condition, the analysis revealed that RMS contrast and saturation were the most influential predictors, jointly explaining 50% of the variance in perceived material dominance ($R^2 = 0.59$; Adjusted $R^2 = 0.50$). Notably, RMS contrast alone accounted for a substantial portion of this variance ($R^2 = 0.40$; Adjusted $R^2 = 0.34$). Similar analyses for other morph weight conditions showed that RMS contrast remained a robust predictor:

- 20–80 morph: $R^2 = 0.49$; Adjusted $R^2 = 0.44$
- 30–70 morph: $R^2 = 0.78$; Adjusted $R^2 = 0.75$
- 40–60 morph: $R^2 = 0.50$; Adjusted $R^2 = 0.45$

These findings suggest that RMS contrast is a primary driver of perceptual bias in binocular material integration, particularly when the difference between the two eyes is moderate (i.e., 20–80 to 40–60 conditions). In contrast, saturation appears to play a larger role only when the binocular difference is more pronounced, as in the 10–90 morph condition.

6. Discussion

Our findings demonstrate a generally linear relationship in the perceptual scaling of morphed materials, which mirrors the linear changes observed in image statistics across interpolation weights. Moreover, we show that perceptual interpolation and binocular material fusion are closely linked to low-level image features. Materials with higher contrast and greater color saturation are weighted more heavily in the resulting perceptual blend. Together, these results suggest that material scaling, interpolation, and interocular summation may rely on a shared representational framework within the visual system.

The way our visual system integrates conflicting visual information across eyes is not only crucial for depth perception but also influences how we interpret texture and material. Here we show that perceptually distinct materials presented separately to each eye can be integrated into a novel coherent material percept in the brain.

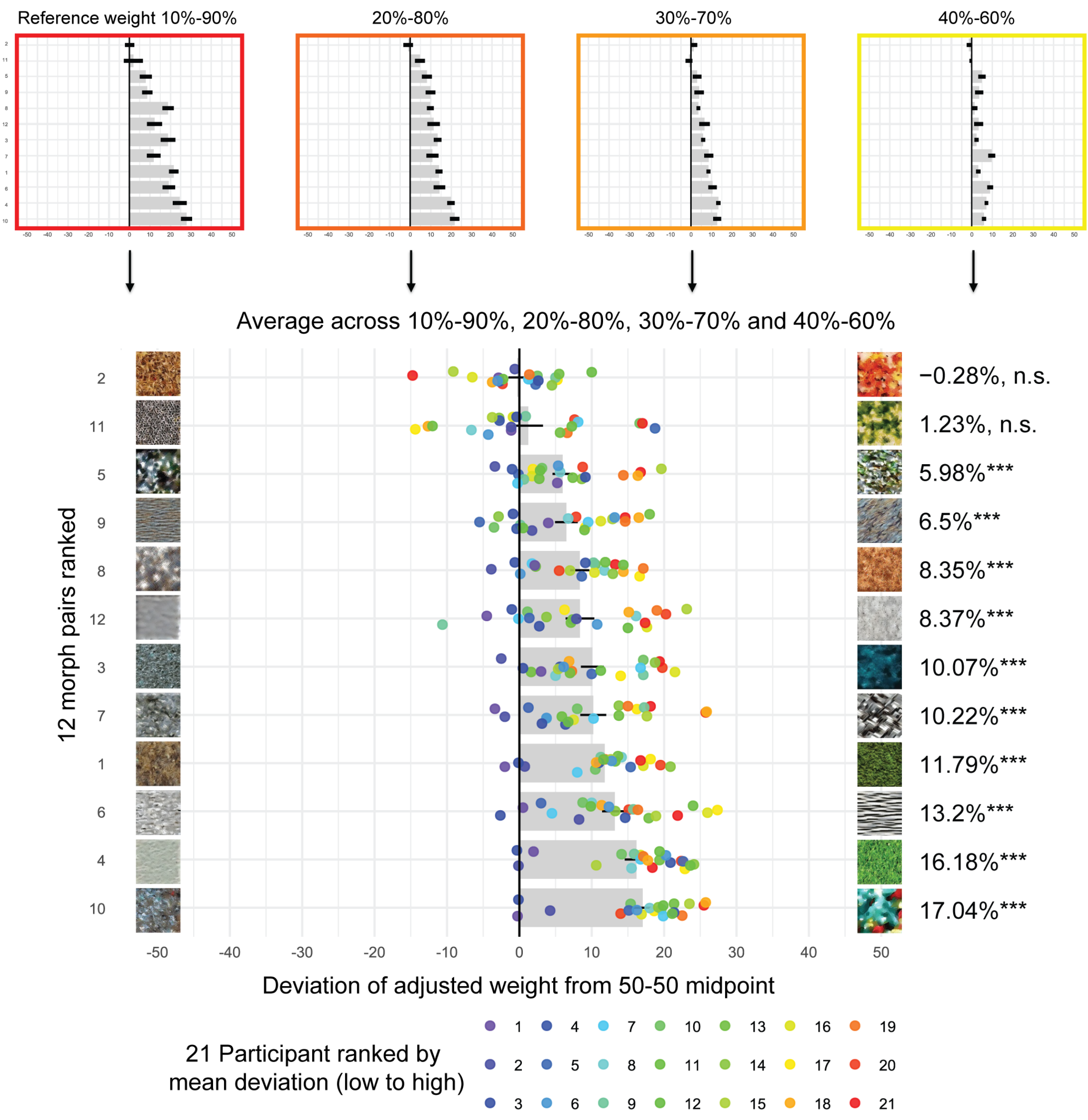

Figure 6: Perceptual equality of the 12 binocularly integrated material pairs in Experiment 3. Top panels display the deviation of participants' adjusted match stimuli from the objective 50-50 morph midpoint, averaged across 21 participants for each morph weight condition: 10-90, 20-80, 30-70, and 40-60 (left to right). The main panel summarizes the mean deviation across all four morph weight conditions for the 12 morph pairs. The example images displayed on either side of each bar were taken from the 10–90 morph weight condition for demonstration purposes. Individual participant data are represented as dots, color-coded by effect size (blue to red = small to large mean deviation). Error bars indicate the standard error of the mean. *** indicates a statistically significancy at the p < .001 level (one-sample t-test against zero)

6.1. Perceptual Scaling of Materials: Linear or Nonlinear?

In Experiment 1, we observed a consistent correspondence between perceptual scaling and physical morph levels, indicating an approximately linear relationship overall. This contrasts with findings by Vacher et al. (2020) [1], who reported nonlinear perceptual scaling for most participants. Several methodological differences may account for the discrepancy. Vacher et al. [1] employed Maximum Likelihood Difference Scaling (MLDS) [6] with a smaller sample (N = 8), whereas our study used a continuous rating task with a larger participant group (N = 29). In addition, the image sets they used are also different [7, 8]. These differences underscore the need for further research to evaluate whether perceptual scaling functions generalize across image sets, methods, and participant samples. Replicating findings under varied conditions will be critical to establishing the robustness of perceptual scaling behavior in material perception.

Note that the 0% and 100% blend conditions were not included from Experiment 1, as these test images would be identical to the original material images shown on either side. In such cases, participants' responses would be expected to align closely with the physical scale, potentially resulting in an S-shaped (sigmoidal) response curve for morph pairs with shallower slopes.

6.2. Perceptual interpolation v.s. binocular integration of materials

We observed a significant correlation between the results of Experiment 2 and Experiment 3 across the 12 morph pairs (r = 0.58, $p < .05$; see Figure 4 and Figure 6). Notably, the strength of this correlation increased systematically as the binocular difference in Experiment 3 decreased. Specifically, the correlation coefficients rose from r = 0.40 in the 10–90 morph weight condition, to r = 0.48 (20–80), r = 0.72 (30–70), and r = 0.77 (40–60), suggesting greater consistency between perceptual interpolation and binocular material integration when the visual inputs to the two eyes were more similar.

The effect size in Experiment 2, measured as the mean deviation from the 50% midpoint (8.81%), falls between the deviation observed in the 20–80 and 30–70 morph weight conditions of Experiment 3 (11.38% and 6.40%, respectively). These comparable findings suggest that perceptual interpolation and binocular material integration may rely on a shared scaling mechanism within a common material representation space.

6.3. Individual Differences in Binocular Material Integration

In Experiment 3, we found that material information presented to each eye with differently weighted morph was integrated binocularly across all 12 tested material pairs (Figure 6). Interestingly, the degree of binocular integration varied systematically across participants. Some individuals consistently exhibited a stronger perceptual bias toward one material over the other within a pair (e.g., red data points in Figure 6), while others showed more balanced integration (blue data points). These individual tendencies were stable across different material pairs, suggesting trait-like patterns of integration. Future research should consider potential contributing factors to these individual differences, such as eye dominance [5] or interocular differences in visual acuity, to better understand the mechanisms underlying material integration in binocular vision.

6.4. The Role of Luster, Binocular Rivalry, and Stereopsis in Binocular Material Fusion

In Experiment 3, differently morphed material images were presented to each eye. Previous research has shown that binocular differences in chromatic and achromatic contrast can give rise to a percept of luster [9, 10]. This can potentially make the fused material appear glossier or more metallic. Additionally, disparities in fine-grained texture and surface patterns between the two eyes' images may evoke depth cues via stereopsis, potentially altering the perceived material structure. These effects could influence performance and bias material judgments on the adjustment task.

In conditions with large interocular differences—particularly in the 10–90 morph condition, where the interocular differences were substantial, the dissimilarity between the two images may approach a threshold that induces binocular rivalry rather than integration. Although the reported incidence of rivalry was low in our study—and those trials were excluded from analysis—future investigations should systematically examine the transition point at which binocular integration gives way to rivalry. Understanding this threshold is essential for interpreting the limits of binocular material fusion.

7. Appendices

The Material Morph Image Synthesis Algorithm [1]
https://github.com/JonathanVacher/texture-interpolation/tree/master

Acknowledgments

This research is funded by the DFG (222641018 – SFB/TRR 135 TP C1), the Excellence Cluster EXC3066 "The Adaptive Mind" and European Research Council Grant ERC-2022-AdG "STUFF" (project number 101098225). We thank Hannah Schösser, Lily Stock, and Zeynep Ceyda Demirkan for data collection.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-4 for grammar and spelling check and editing. After using the tool, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

[1] J. Vacher, A. Davila, A. Kohn, R. Coen-Cagli, Texture interpolation for probing visual perception, 2020. URL: https://arxiv.org/abs/2006.03698.

[2] F. Schmidt, M. N. Hebart, A. C. Schmid, R. W. Fleming, Core dimensions of human material perception, Proceedings of the National Academy of Sciences 122 (2025). doi:10.1073/pnas.2417202122.

[3] H.-C. Sun, F. Schmidt, A. C. Schmid, M. N. Hebart, R. W. Fleming, Cortical representations of core visual material dimensions, Journal of Vision 24 (2024) 285. doi:10.1167/jov.24.10.285.

[4] D. H. Baker, K. J. Hansford, F. G. Segala, A. Y. Morsi, R. J. Huxley, J. T. Martin, M. Rockman, A. R. Wade, Binocular integration of chromatic and luminance signals, Journal of Vision 24 (2024) 7. doi:10.1167/jov.24.12.7.

[5] R. Jiang, M. Meng, Integration and suppression interact in binocular vision, Journal of Vision 23 (2023) 17. doi:10.1167/jov.23.10.17.

[6] L. T. Maloney, J. N. Yang, Maximum likelihood difference scaling, Journal of Vision 3 (2003) 573–585. doi:10.1167/3.8.5.

[7] M. Cimpoi, S. Maji, I. Kokkinos, S. Mohamed, A. Vedaldi, Describing textures in the wild, in: 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 3606–3613. doi:10.1109/CVPR.2014.461.

[8] P. Arbeláez, M. Maire, C. Fowlkes, J. Malik, Contour detection and hierarchical image segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence 33 (2011) 898–916. doi:10.1109/TPAMI.2010.161.

[9] F. A. A. Kingdom, K. Mohammad-Ali, C. Breuil, D. Chang-Ou, A. Irgaliyev, Detection of vertical interocular phase disparities using luster as cue, Journal of Vision 23 (2023) 10. doi:10.1167/jov.23.6.10.

[10] B. J. Jennings, F. A. A. Kingdom, Detection of between-eye differences in color: Interactions with luminance, Journal of Vision 16 (2016) 23. doi:10.1167/16.3.23.