Title: The eyes grasp, the hands see: metric category knowledge transfers between vision and touch

Authors: Christian Wallraven[1], Heinrich H. Bülthoff[1,2], Steffen Waterkamp[2], Loes van Dam[3], Nina Gaißert[2]

[1] Department of Brain and Cognitive Engineering, Korea University, Seoul, Korea

[2] Max Planck Institute for Biological Cybernetics, Tübingen, Germany

[3] University of Bielefeld, Bielefeld, Germany

Correspondence concerning this article should be addressed to Christian Wallraven, Department of Brain and Cognitive Engineering, Korea University Anam-Dong 5ga, Seongbuk-gu, Seoul 136-713, Korea. E-mail: wallraven@korea.ac.kr

The version of record of this article, first published in Psychonomic Bulletin & Review is available online at Publisher's website: http://dx.doi.org/10.3758/s13423-013-0563-4

Abstract: Categorization of seen objects is often determined by the shape of objects. However, shape is not exclusive to the visual modality: the haptic system also is expert at identifying shapes. Hence an important question for understanding shape processing is whether humans store separate modality-dependent shape representations or whether information is integrated into one multi-sensory representation. To answer this question, we created a metric space of computer-generated, novel objects varying in shape. These objects were then printed using a 3D printer to generate tangible stimuli. In a categorization experiment, participants first explored the objects visually and haptically. We found that both modalities led to highly similar categorization behavior. Next, participants were trained either visually or haptically on shape categories within the metric space. As expected, visual training increased visual performance and haptic training increased haptic performance. Importantly, however, we found that visual training also improved haptic performance and vice versa. Two additional experiments showed that the location of the categorical boundary in the metric space also transferred across modalities, as did heightened discriminability of objects adjacent to the boundary. This observed transfer of metric category knowledge across modalities indicates that visual and haptic shape information is integrated into a shared multi-sensory representation.

Keywords: shape, object categorization, vision, haptics, categorization, multi-sensory representations

1. Introduction

Humans learn about the world by interacting with it, and much of this interaction is mitigated through the sense of touch of our hands. Starting with active exploration of objects in infants, up to the precise manipulation skills required by surgeons, the haptic modality allows us to gather information about an object's shape, texture, softness, weight, temperature, and other material properties. Many of these fundamental object properties, such as temperature and weight, are not or only indirectly accessible to the other modalities, e.g. vision, highlighting the importance of interacting with the world through haptics for developing perceptual and precision skills. In contrast, other object properties, such as shape and texture are readily accessible to both the visual and haptic modality. Of these two properties, shape has been found to play a crucial role in object identification and categorization for the visual (e.g., Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) and haptic (e.g., Klatzky, Lederman, Metzger, 1985) modalities. Since shape can be perceived visually and haptically the question arises whether shape representations integrate sensory input from both modalities, or whether the human brain stores two separate, modality-dependent shape representations.

The idea of common representations is supported by cross-modal priming observed between vision and haptics (Bushnell & Baxt, 1999; Reales & Ballesteros, 1999). Similarly, several recent studies investigating perceptual spaces of complex, parametrically-defined objects (Cooke, Jäkel, Wallraven, & Bülthoff, 2007; Gaissert, Wallraven, & Bülthoff, 2010) and natural objects (Gaissert & Wallraven, 2011) revealed high similarity between the visual and haptic perceptual spaces, suggesting similar processing of shape. If one multisensory representation is formed, however, one might expect cross-modal shape comparisons to be equivalent in performance to uni-modal shape comparisons. In this context, Norman et al. (Norman, Norman, Clayton, Lianekhammy, & Zielke, 2004) reported high accuracy but also significant performance differences between cross-modal and uni-modal shape comparisons concluding functionally overlapping but distinguishable representations. Likewise, several studies suggest that haptic performance might be limited by shape complexity (Phillips, Egan, & Perry, 2009; Dopjans, Wallraven, & Bülthoff, 2009). Lacey et al. (Lacey, Campbell, & Sathian, 2007) reviewed further previous studies and suggested that evidence from both behavioral and imaging studies were consistent with shared shape representations that enable efficient cross-modal transfer of object information.

In other words, there is ample evidence that shape identities are shared at least to some extent between the modalities and that shape metrics are at least similar across vision and haptics. However, this leaves open the question whether the metric representations of the separate modalities are merely similar, but independent, or if the modalities share a shape representation. To address this question, we trained participants on a specific metric shape categorization task using only one modality, either vision or haptics, and tested for transfer effects to the other modality. If the shape metrics are shared, the trained categorization knowledge should transfer also to the other modality. However, if vision and haptics do not share the same metric shape space the untrained modality should remain largely unaffected.

In this study we additionally investigated two aspects of categorization learning. First, we investigated categorization performance itself, before and after training, for both the haptic and visual modalities. Second, we looked for “categorical perception”, a feature of category learning in which stimuli straddling the categorical boundary are easier to discriminate than stimuli within each category (Pastore, 1987; Harnad 1987).

2. General Method

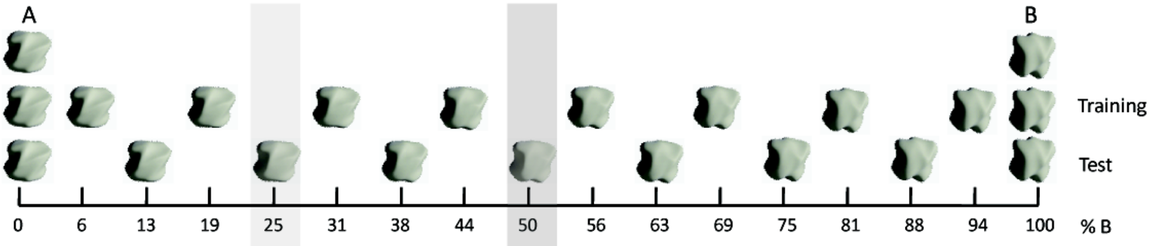
2.1. Stimuli

To generate tangible objects on a metric scale, computer-graphics modeling was combined with rapid 3D-prototyping. Two prototype objects A and B were generated using the software Autodesk 3ds Max (Autodesk, Canada) by taking a sphere of 7 cm diameter and modifying its shape using two orthogonally positioned "wave modifiers", resulting in smoothly deformed, undulated objects. Two sets of parameters for the wave modifiers were chosen as to provide two distinct, but not too dissimilar prototype objects. To obtain the metric scale, the two objects were linearly morphed into each other in 15 intermediate steps. Objects were then printed using a 3D printer (ZPrinter 650, ZCorporation, Germany) and were mounted on little stands for easier haptic exploration. The final stimulus set consisted of 17 different shapes (see Figure 1a) equal in weight and volume. Since the experiments consisted of training and testing conditions, we split the stimulus set into a training set and a test set to ensure that participants did not simply learn single objects but attended to shape features that would generalize.

2.2. Experiment Setup

For the visual experiments, participants were seated in front of a sliding door (see Figure 1b) for the setup), which could be opened and closed automatically. The experimenter placed one object in one of six possible orientations (0°, 60°, 120°, 180°, 240°, and 300°) behind the door. Upon a signal given by the experimental computer, the door opened and participants were able to visually explore the objects for exactly two seconds before the door closed again.

In the haptic experiments, the same setup was used. The objects were placed in one of six possible orientations into the stand behind the sliding door. For all haptic trials, however, the door remained closed and participants were only able to touch the objects using their dominant hand by reaching around the door. Participants were allowed to freely explore the objects for four seconds. A tone signaled the start of the exploration time and a second tone signaled the end. The experimenter took care that participants did not exceed the exploration time.

a) 

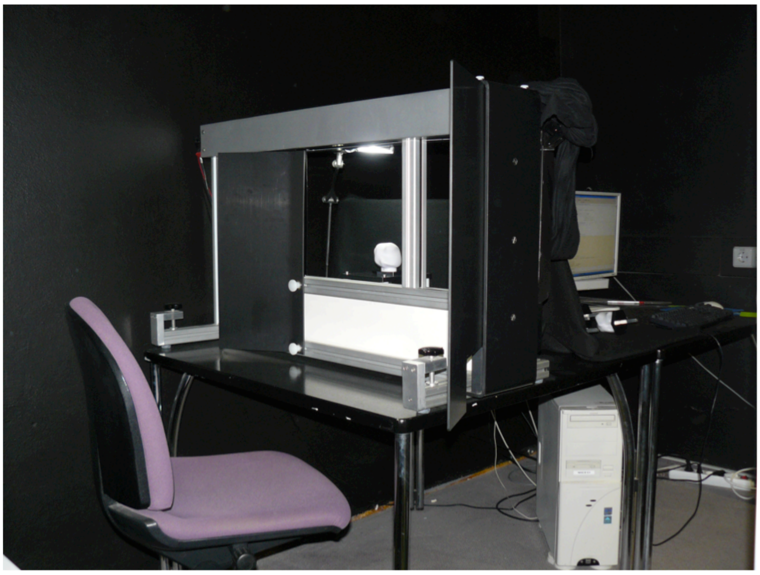
b) 

Figure 1. a) Stimuli. The figure shows stimulus A and B and the 15 intermediate morph-steps (The x-axis displays the amount of B-features in percent). The stimulus set was divided into a training set and a test set. In Experiment 1, participants were trained on the categorical boundary at 50% (dark gray), in Experiment 2, participants were trained with a categorical boundary at 25% (light gray). b) Setup. View of the setup from the participant’s side. A computer-controlled sliding door (shown open here as in the visual conditions in the experiments) was used to hide or reveal the objects. In the haptic conditions, the door was closed and participants reached around the door to touch the objects.

3. Experiment 1: Categorization Experiment

3.1. Method

The experiment consisted of a pre-training test (testing visual and haptic performance separately), a unimodal training phase (training for one modality only), and a post-training test (testing again visual and haptic performance separately). Both pre- and post-training test phases used the seven objects of the test set, whereas the training phase used the eight objects of the training set. Since the whole experiment took about four hours, it was split into two sessions that took place on two consecutive days.

In the pre-training test, participants had to categorize the objects of the *test set* visually and haptically in separate trials. To avoid order effects, half of the participants started exploring the objects visually while the other half started exploring the objects haptically. The test started by introducing the participants to object A and B in the chosen modality, and participants were informed that these two stimuli represented the prototypes of category A and B, respectively. For this, object A was presented in orientation 0°, 60°, 120°, 180°, 240°, and 300°. Then object B was presented in the same orientations. Next, the seven objects of the test set were presented in randomized order and in one of six random orientations (the randomization ensured that every orientation of every object occurred at least once in all blocks). Participants were asked to indicate whether the object belonged to category A or to category B. No feedback was provided. The testing was repeated ten times in total, resulting in 70 trials. After half of the test trials A and B were presented again from all six orientations as a reminder.

Next, participants had to complete the training. Ten participants were trained visually; ten other participants were trained haptically. During this phase, the training set was used to ensure that participants attended to category features and did not simply learn the objects themselves. Similarly to the previous pre-training blocks, the prototypes for each category (objects A and B) were presented from all six orientations, then the morphed objects were presented randomized by order and orientation. Again, participants had to indicate if the presented object belonged to category A or B. This time feedback was provided in such a way that all objects with less than 50% B features were assigned to category A and all objects with more than 50% B features were assigned to category B. The training ended when participants reached the performance criterion for which at least 7 out of 8 objects needed to be categorized correctly, over three consecutive runs with a run consisting of the presentation of the 8 objects of the training set. On the next day, participants repeated another training in the same fashion. After reaching the performance criterion again, participants went on to the post-training test, in which categorization performance was then tested for both the trained and the untrained modalities. Testing after training was identical to that before training.

3.2. Results

In order to test the statistical significance of the results, the ratings of each participant before and after training were analyzed by fitting psychometric functions to the data. A cumulative Gaussian was fitted to these data points using the *psignifit* toolbox for Matlab, which implements the maximum-likelihood method (Wichmann & Hill, 2001). The fitted psychometric function yields estimates of the PSE (point of subjective equivalence as indicated by 50% "B" ratings) and JND (just noticeable difference, calculated as the morph-difference that would bring performance from the PSE to 75%). Perfect performance would yield a step-like function in which all objects with less than 50% B-features would be identified as category A and all objects with more than 50% B-features would be identified as category B. The object at morph-level of 50% would be arbitrarily assigned to either A or B and hence become the PSE. In other words, the PSE represents the category boundary, whereas the JND represents its sharpness (large JND: fuzzy category boundary, small JND: sharp category boundary).

The PSE and JND data were compared for pre- and post-training conditions using Wilcoxon sign-rank tests for paired data, and Mann-Whitney-U tests for comparing unpaired data across training conditions. The data for one representative participant is shown in Figure 2, and group data for JNDs and PSEs is shown in Figure 3a,b.

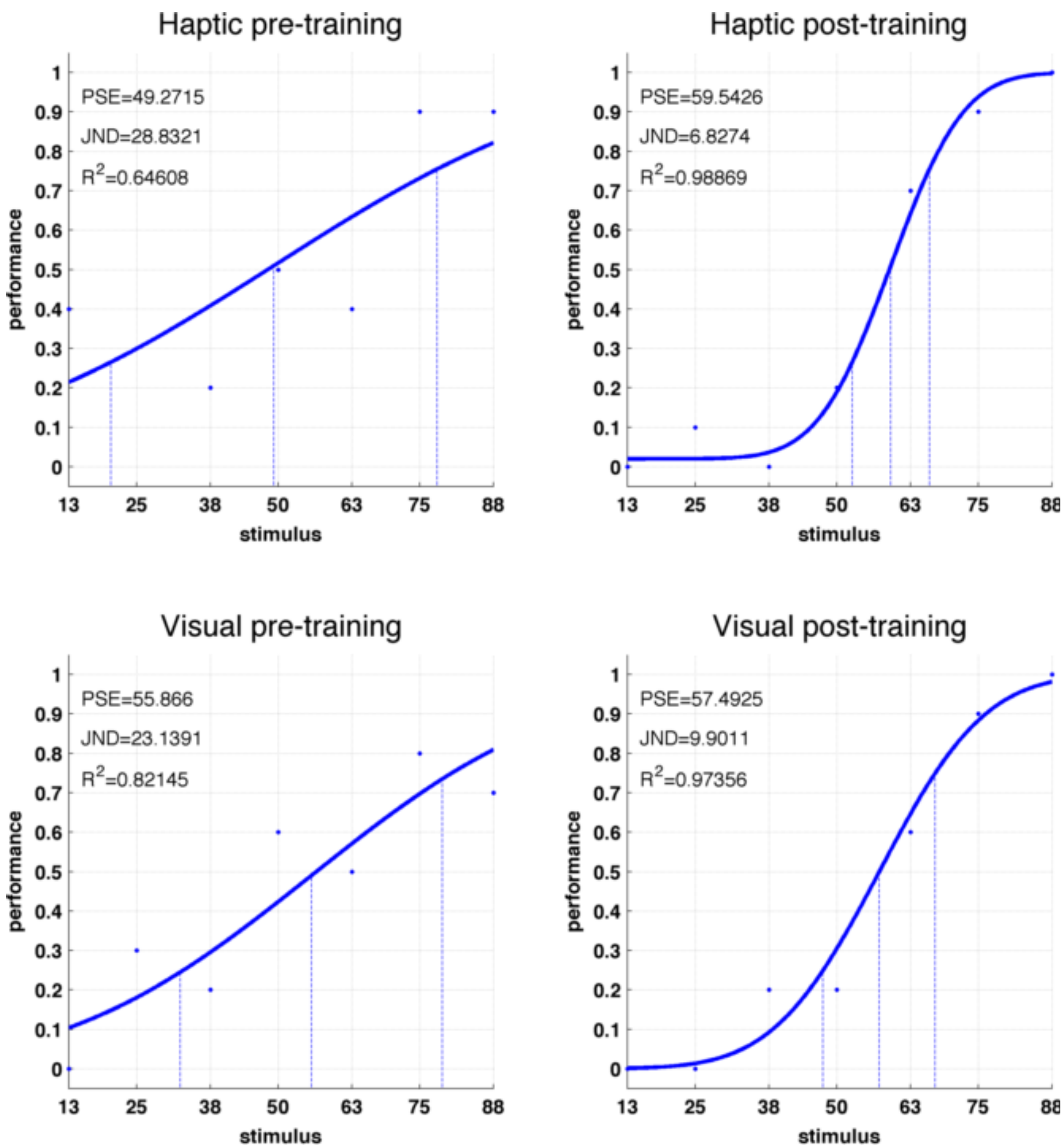


Figure 2. Experiment 1. Psychometric function fits for a representative participant who was trained in the haptic modality. The upper row shows categorization results for the haptic modality, the lower row for the visual

modality. Values for the PSE, JND, and the goodness-of-fit are given for each sub-figure. Note that performance (as indicated by the steepness of the curve, or decreased JND-values) increases for both modalities.

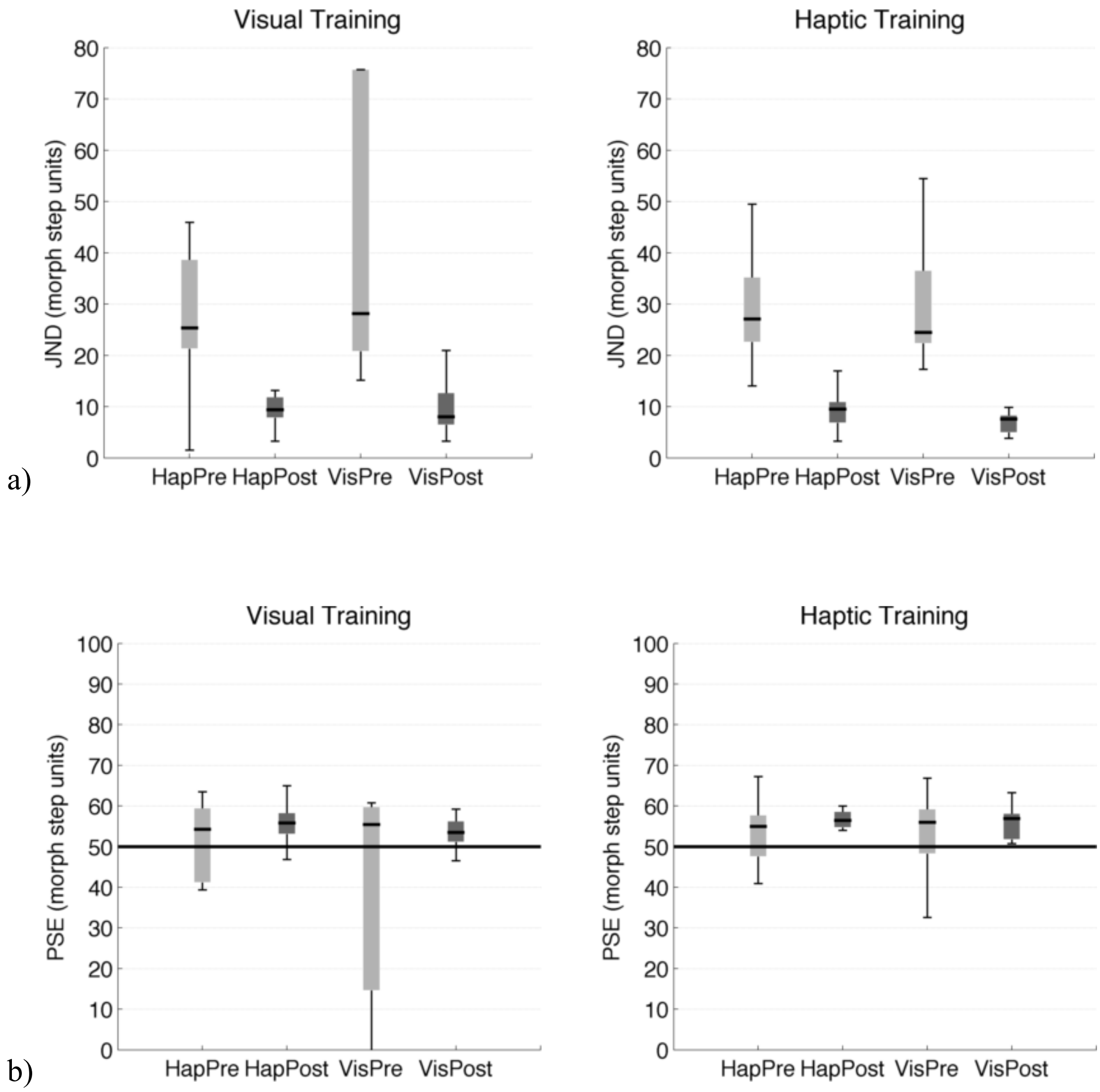


Figure 3. Experiment 1. Group results for a) JND-values and b) PSE-values separated for the two groups. Note that the JND improves for both groups regardless of whether haptic or visual training is performed. PSE-values do not change through training – the trained categorization boundary of 50% is indicated with a bold line.

Figures use boxplots with line showing median, the shaded area covering the inter-quartile range (IQR) from the 25%- to the 75%-quartile, and whiskers extending 1.5 times the IQR from the median.

As the single-participant data shows (Figure 2), the psychometric curve before training is very shallow, corresponding to a poor separation of the two categories. This is true for both the visual and the haptic modalities (JNDs are both around 25%). After training, both psychometric curves become steeper, indicating good discriminability of the categories (both JNDs around 6-9%). Note, that this result means that training is equally effective for both modalities, despite the fact that this participant was only trained in the haptic modality. Also note that the PSE for all four conditions stays at roughly the same level around 50% (see discussion below).

Next, we analyzed the group data for JND-values. As shown in Figure 3a, visual training increased visual performance significantly ($W=47.000$, $Z=-2.797$, $p=0.025$) and haptic training increased haptic performance significantly ($W=55.000$, $Z=-3.628$, $p=0.001$). More importantly, however, cross-modal transfer was also significant in that visual training increased haptic performance significantly ($W=52.000$, $Z=-2.797$, $p=0.005$) and that haptic training significantly increased visual performance ($W=55.000$, $Z=-3.780$, $p=0.002$). In addition, the post-training JNDs tested for cross-modal transfer were not significantly different in both training modalities (all $p>0.225$), showing that trained category knowledge can transfer equally well across modalities.

Similar tests on the PSEs for pre- and post-training tests failed to yield any significant differences (all $p>0.275$, see Figure 3b) – PSEs on average were around 55 both before and after training showing that learning of the category boundary was stable. Note that a shift in PSE would also not be expected as the untrained category boundary would most likely also occur somewhere in the middle of the metric scale (see Experiment 2 below).

In summary, the analysis of the individual data clearly demonstrates that participants were able to benefit from within-modal training, and that the acquired knowledge about the metric visual or haptic category structure can easily transfer to the other modality.

4. Experiment 2: Shifting the Categorical Boundary

As a next step, one needs to verify that the observed training effects actually resulted from the training phase (supervised learning) and were not due to the repeated exposure to the stimuli in the testing phase (unsupervised learning). In order to confirm this training effect, the same experiment was repeated, but this time participants were trained on a categorical boundary shifted to the left compared to the previous experiment. This experiment also tests the degree of malleability of the category boundary – if the set of objects that were chosen afforded a "natural" category boundary at morph levels of 50%, perhaps it would be harder to train participants with another categorical boundary.

4.1. Method

The procedure was the same as for Experiment 1 with the following changes: the feedback during the training phase was adjusted such that all objects with less than 25% B features were labeled as category A and all objects with more than 25% B features were labeled as category B. This shift in categorical boundary resulted in fewer A objects and more B objects. As the experiment was already very time consuming, we decided not to increase the amount of trials with more objects for A, opting for imbalanced categories. Another group of twenty participants took part in this experiment; again ten participants were trained visually, another ten participants were trained haptically.

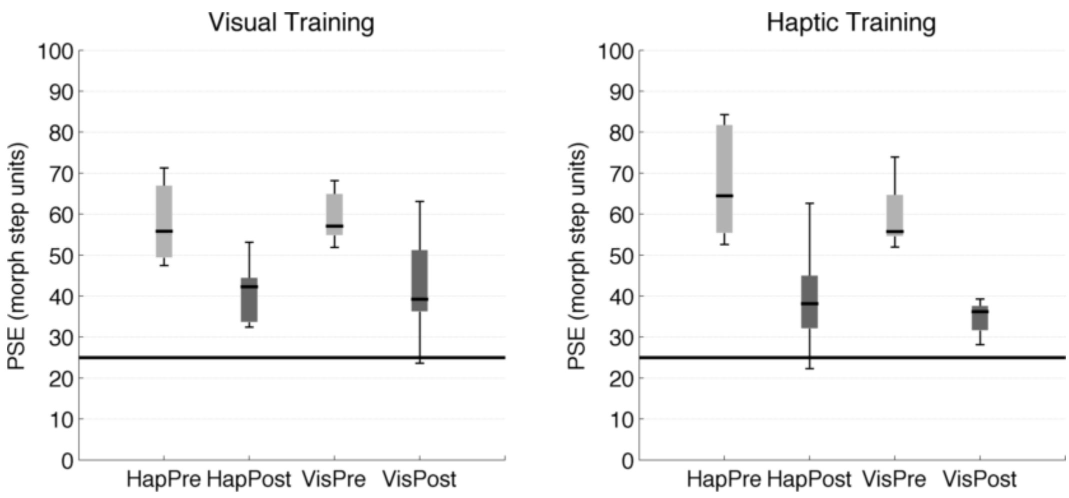


Figure 4. Experiment 2. Group results for PSE-values. Note, that for training of the 25% category boundary, the PSE- values shift from around 50% (the "naïve boundary") to significantly lower values. The 25% category boundary is shown in bold in the figure. Figures use boxplots with line showing median, the shaded area covering the inter-quartile range (IQR) from the 25%- to the 75%-quartile, and whiskers extending 1.5 times the IQR from the median.

4.2. Results

As before, ratings of every single participant before and after training were analyzed by fitting a cumulative Gaussian to the participants' data and retrieving individual JNDs and PSEs, which were compared using Wilcoxon sign rank tests. The data for PSEs is shown in Figure 4.

As in the previous experiment, training of participants resulted in significant decreases of the individual JNDs for both within- and across-modality testing (all $p<0.005$) confirming the effectiveness of the training, as well as the cross-modal transfer of category knowledge. For this experiment, however, the main interest lay in evaluating the effect of shifting the categorical boundary and hence in comparing the pre- and post-training PSEs (see Figure 4). As should be expected, both categorical boundaries were shifted in the within-modal conditions (visual: $W=53.000$, $Z=-2.948$, $p=0.006$; haptic: $W=55.000$, $Z=-3.250$, $p=0.002$). Most importantly, haptic training affected visual performance significantly ($W=54.000$, Z=-

3.099, p=0.004), as did visual training for haptic performance (W=55.000, Z=-3.553, p=0.002). Again, the two shifted PSEs after training did not differ significantly for the two groups (all p>0.325). Note that the shifted PSE after training did not quite reach the target of 25% (the average was around 38% for all modalities). This effect is most likely due to the smaller number of exemplars for category A participants were exposed to in the training runs. Overall, these results demonstrate a clear influence of the training phase on performance in being able to shift the categorical boundary and hence show that the training was effective. In addition, not only does sensitivity transfer across modalities, but also knowledge about the location of the categorical boundary.

5. Experiment 3: Discrimination Experiment

5.1. Method

Categorical perception effects are indicated by the fact that object pairs straddling the category boundary are easier to discriminate than object pairs located within the same category (Bornstein, 1987). The standard procedure for testing this effect is to run a same-different discrimination experiment on pairs of equidistant stimuli and to use the results to calculate d' – a measure of sensitivity taking into account hits and false alarms.

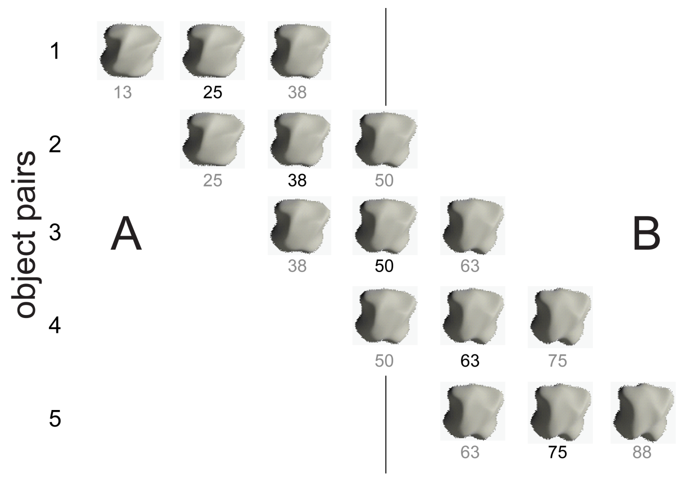


Figure 5. Experiment 3. Stimuli. The five object pairs used in the discrimination experiment. For object pair 1 stimulus 25 was shown twice in a "same" trial while object 13 and 38 were shown in a "different" trial. Object pair 3 straddles the physical categorical boundary located at 50% B features.

In order to limit the total experimental time for the haptic experiments, five object pairs were selected from the full object set. As an example, for object pair one, the "same" condition was morph-step 25 versus morph-step 25, while for "different" it was morph-step 13 versus morph-step 38, since 13 and 38 are the neighboring objects in the test set (see Figure 5 for the remaining pairs).

As for the categorization experiments, this experiment also consisted of three parts: a pre-training test and a learning phase on the first day, and a learning phase and the post-training test on the second day. Again ten participants were trained visually only, ten other participants were trained haptically only; for both groups pre-training test and post-training test were conducted visually and haptically (half of the participants started visually, while the other half started haptically).

During the pre-training phase each object pair was presented eight times. Since there were five "same" pairs and five "different" pairs, this resulted in (5+5)*8=80 pairs. The eighty object pairs were presented in random order and random orientation. First, one object was presented and explored by the participant visually or haptically. After exploration, the object was replaced by the second object of the pair (either the same or a different object) and participants had to respond "same" or different". For the first object, one of the six possible orientations was randomly selected (taking care to ensure that different orientations occurred equally often across trials). The second object was then presented in the same orientation as the first object. No feedback was provided.

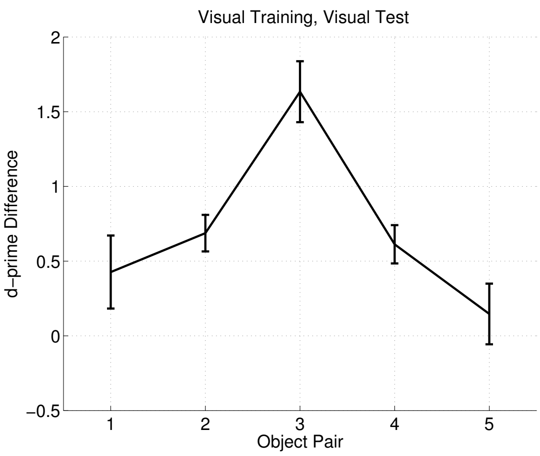
After the test phase, participants had to complete the training, which was performed in exactly the same manner as in the categorization experiment as we also wanted to train the categorization boundary. The training was again split into two sessions on two consecutive days – for both days the training criterion had to be reached. After the second training block finished, participants completed the post-training test, which was performed in exactly the same manner as the pre-training test of the discrimination experiment. Since this experiment was longer than the previous ones, we inserted additional breaks so as to prevent fatigue.

5.2. Results

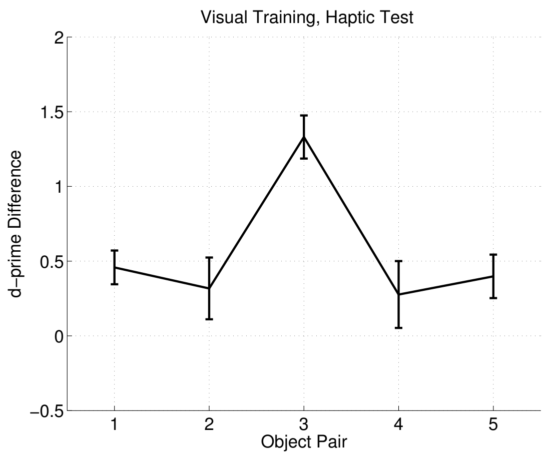
Categorical perception effects are characterized by a higher discriminability of inter-stimulus differences for stimuli straddling the categorical boundary. Standard procedure to test for this effect is to calculate d', which is the difference in z-scores between hit rate (correctly identified "same" pairs) and the false-alarm rate ("different" pairs identified as "same"). If categorical perception occurs, the d'-values should exhibit a peak for object pair 3 (see Figure 4a).

Participants' responses were converted to d' for the pre- and post-training data for each of the five different object pairs and subtracted to yield d'-differences. The average d'-differences for the two within-modality conditions were $\Delta d_{Visual/Visual}'=0.751$ and $\Delta d_{Haptic/Haptic}'=0.610$ both significantly larger than zero, showing that learning took place during training (both $p<0.001$ in Wilcoxon-sign-rank tests). Importantly, the average d'-differences for the across-modality conditions were $\Delta d_{Visual/Haptic}'=0.595$ and $\Delta d_{Haptic/Visual}'=0.515$ (both $p<0.001$), which confirms that information about training was transferred across modalities. Although these d'-differences seem to indicate a trend for a smaller gain in the across-modality conditions, this trend did not approach significance in a post-hoc Friedman test conducted across all four conditions ($\chi^2(3, N=10)=5.16$, $p=0.160$). Hence, overall training effects were similar for all four conditions.

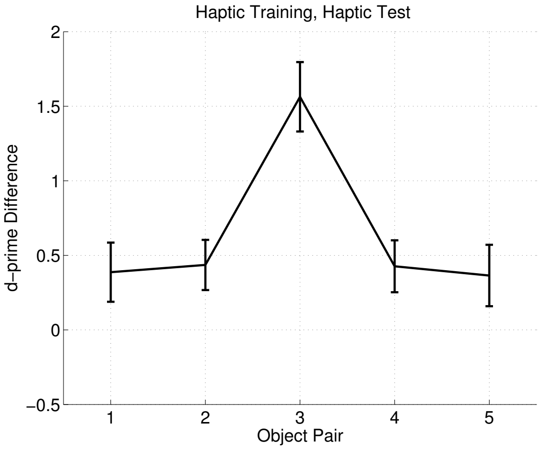
In the following analysis, we were interested in comparing the data for the two within-modality conditions and the two across-modality conditions. This was done to check whether we would find categorical perception for training within one modality, but also whether there would be evidence for transfer of increased sensitivity at the category boundary across modalities. For this, we ran Wilcoxon-sign-rank tests for the center-pair stimuli (“Object Pair 3” in Figures 5 and 6) compared to the other four pairs – all tests were Bonferroni-corrected for multiple comparisons within each condition. Except for one test (for the Haptic/Visual condition comparing Object Pair 3 with Object Pair 1 – see Figure 6d, W=48.000, Z=-3.780, p=0.019 at alpha-level of 0.0125), all other tests were significant (all p<0.010) across all four conditions. Taken together, we therefore demonstrated clear peaks in discriminability in both within- and across-modality conditions for the central object pair. Hence, we find evidence that in addition to transfer of training across modalities, the heightened discriminability of objects around the categorical boundary also seems to transfer well.



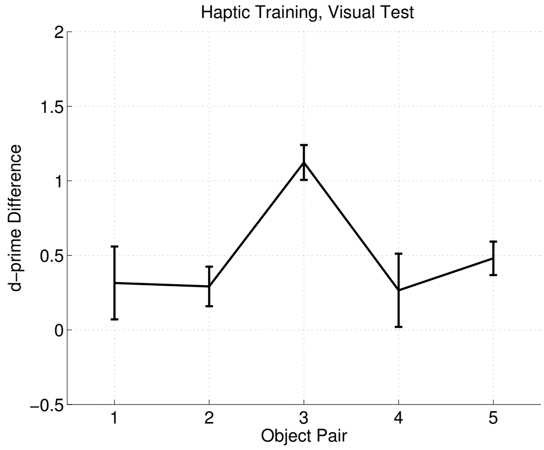
a)



b)



c)



d)

Figure 6. Experiment 3. Results in d’-differences before and after training for within-modal conditions (a & c) and across-modal conditions. Enhanced sensitivity (that is, a visible peak) for the center object pair as predicted by categorical perception is visible in all four conditions. Data show means +/-1SEM.

6. Discussion

In the present study, we investigated how metric shape information can transfer across modalities. For this, we trained participants on shape categories either visually or haptically and analyzed how this training affected visual and haptic performance in two categorization tasks and in a discrimination task. In Experiment 1, the categorization task showed a strong transfer of learning across the senses with visual learning increasing haptic performance and vice versa. This effect was verified in Experiment 2 with a categorization task in which participants were trained on a shifted categorical boundary. Here we found that training of one modality affected the percept of the other modality by shifting the categorical boundary within the untrained modality. Finally, we performed a discrimination task in Experiment 3. Following categorical perception theory (Pastore, 1987; Harnad, 1987) the formation of a categorical boundary should increase the discriminability for object pairs straddling the categorical boundary. This effect was found for within-modality training and across-modality training. Since typical experiments with unfamiliar objects require a large amount of training to obtain categorical perception effects (see Kitania, Roberson, & Hanley, 2010 for an example with unfamiliar face recognition), this experiment used a larger number of test-trials to obtain more reliable estimates of discriminability. We found evidence for an increased sensitivity at the boundary for all four conditions (regardless of training or testing modality) with a still relatively modest number of trials. In sum all three experiments therefore revealed robust and clear transfer of category learning from vision to touch and vice versa.

Note that in our experimental design, participants were exposed to the full metric space in the training phase. The training effects we observed for the JNDs may therefore be simply due to mere exposure to the full space rather than to the category training itself. In a pilot experiment using only visual training and testing, in which the training phase did not include feedback, however, JNDs were not significantly altered (N=16 participants split between a mere-exposure condition and a feedback-condition). This shows that mere exposure cannot explain the training effect on JNDs observed in our experiments.

Previous studies on information transfer between vision and haptics had so far demonstrated that metric shape spaces are similar between vision and haptics (e.g., Cooke et al., 2007, Gaissert et al., 2010) and that information necessary for recognition of individual objects can be shared (e.g., Dopjans et al., 2009, Reales & Ballesteros, 1999, Norman et al., 2004). Here, we show that general category knowledge about complex shape changes is readily shared across modalities after only little training. A recent study by Yildirim and Jacobs (2013) also demonstrated that category knowledge can transfer across the senses. In their study, a set of 40 computer-generated objects ("Fribbles") split into 4 categories was investigated in a cross-modal categorization task. After 7 training blocks in one modality, participants were able to transfer categorization results to the other modality with either full or partial transfer (the latter in case of short visual presentation). Our results go further in that they demonstrate that changes to the *metric space* in one modality readily transfer to the other modality. That is, whereas the study by Yildirim and Jacobs (2013) showed transfer for families of arbitrarily-defined objects, here we show that people are able to transfer information not only about category membership, but also about the detailed *structure* of the category across modalities. We were able to do this as we used a parametrically-defined shape space for training and testing. Our results show that - for the shape spaces used here - this cross-modal transfer is also symmetric such that information about categories *and* categorical boundaries is transferred fully between vision and haptics. As Experiment 2 demonstrated, categorical boundaries were easily shifted by training at a different location, whereas Experiment 3 provided evidence for heightened discriminability at the boundary locations as required for categorical perception. Hence, our results add to those of Yildirim and Jacobs (2013) by highlighting an even closer integration of shape processing in vision and haptics, showing that despite considerable differences in exploration strategy, visual and haptic exploration of novel shape categories gives rise to similarly structured category knowledge. The experiments reported here shed light on the properties of multisensory representational space in the context of shape processing. It has been suggested that mental representations in general (Shepard, 1987) and object representations in particular (Edelman and Shahbazi, 2012) consist of a structured space in which inter-object similarity determines distances between objects. To avoid confusion between objects, the perceptual system creates category boundaries between these objects and, in addition, for highly similar objects and categories heightens the distinctiveness of exemplars in each category through categorical perception (Harnad (1987) - see also the study by Newell and Bülthoff (2002) for morphed familiar objects in the visual domain). Our results hence can be interpreted in favor of a joint, visuo-haptic representational space encoding fine-grained shape knowledge (see also Gaissert et al., 2010, Lacey, Campbell, & Sathian, 2007). Future research will need to determine the degree to which transfer of category knowledge may be limited by object complexity and/or number of categories.

Acknowledgments

This research was supported by a PhD stipend from the Max Planck Society. Part of this research was also supported by the WCU (World Class University) program through the

National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10008-0).

Commercial relationships: none.

Corresponding author: Christian Wallraven

Email: wallraven@korea.ac.kr.

Address: Anam-dong, Seongbuk-gu, Seoul 136-713, Korea

References

Bornstein, M. H. (1987). Perceptual categories in vision and audition. In S. Harnad (Ed.), *Categorical perception: The groundwork of cognition* (pp. 535-565). New York: Cambridge University Press.

Bushnell, E. W., & Baxt, C. (1999). Children's haptic and cross-modal recognition with familiar and unfamiliar objects. *Journal of Experimental Psychology-Human Perception and Performance, 25*(6), 1867-1881.

Cooke, T., Jäkel, F., Wallraven, C., & Bülthoff, H. H. (2007). Multimodal similarity and categorization of novel, three-dimensional objects. *Neuropsychologia, 45*(3), 484-495.

Dopjans, L., Wallraven, C., & Bülthoff H.H. (2009). Cross-modal transfer in visual and haptic face recognition. *IEEE Transactions on Haptics*, 200(4):236-240.

Edelman, S., & Shahbazi, R. (2012). Renewing the respect for similarity. *Frontiers in Computational Neuroscience* 6:45.

Gaissert, N., & Wallraven, C. (2011). Categorizing Natural Objects - A Comparison of the Visual and Haptic Modalities. *Experimental Brain Research,* 216(1),123-134.

Gaissert, N., Wallraven, C., & Bülthoff, H. H. (2010). Visual and Haptic Perceptual Spaces Show High Similarity in Humans. *Journal of Vision, 10*(11:2), 1-20.

Harnad, S. (1987). Psychophysical and cognitive aspects of categorical perception: a critical overview. In S. Harnad (Ed.), *Categorical Perception: The Groundwork of Cognition*. New York: Cambridge University Press.

Kikutani, M., Roberson, D., & Hanley, J.R. (2010). Categorical Perception for Unfamiliar Faces: The Effect of Covert and Overt Face Learning. *Psychological Science* 21, 865-872.

Klatzky, R. L., Lederman, S.J., & Metzger, V. A. (1985). Identifying objects by touch: an expert system. *Perception & Psychophysics*, 37(4), 299-302.

Lacey, S., Campbell, C., & Sathian, K. (2007). Vision and touch: multiple or multisensory representations of objects? *Perception, 36*(10), 1513-1521.

Newell, F., & Bülthoff, H.H. (2002). Categorical Perception of Familiar Objects. *Cognition, 85*, 113-143.

Norman, J. F., Norman, H. F., Clayton, A. M., Lianekhammy, J., & Zielke, G. (2004). The visual and haptic perception of natural object shape. *Perception & Psychophysics, 66*(2), 342-351.

Pastore, R. E. (1987). Categorical Perception: Some Psychophysical Models. In S. Harnad (Ed.), *Categorical perception: The groundwork of Cognition*. New York: Cambridge University Press.

Phillips, F., Egan, E.J.L., & Perry B.N.(2009). Perceptual equivalence between vision and touch is complexity dependent, *Acta Psychologica*, 132(3), 259-266.

Reales, J. M., & Ballesteros, S. (1999). Implicit and explicit memory for visual and haptic objects: Cross-modal priming depends on structural descriptions. *Journal of Experimental Psychology-Learning Memory and Cognition, 25*(3), 644-663.

Reed, C. L., Lederman, S. J., & Klatzky, R. L. (1990). Haptic Integration of Planar Size with Hardness, Texture, and Planar Contour. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 44*(4), 522-545.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382-439.

Shepard, R.N. (1987). Toward a universal law of generalization for psychological science. Science 237:1317–1323.

Wichmann, F., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics, 63*(8), 1293-1313.

Yildirim, I., & Jacobs, R. A. (2013). Transfer of object category knowledge across visual and haptic modalities: Experimental and computational studies. *Cognition*, 126, 135-148.