

# Categorizing natural objects: a comparison of the visual and the haptic modalities

Nina Gaissert · Christian Wallraven

Received: 22 August 2011 / Accepted: 18 October 2011 / Published online: 3 November 2011  
© Springer-Verlag 2011

**Abstract** Although the hands are the most important tool for humans to manipulate objects, only little is known about haptic processing of natural objects. Here, we selected a unique set of natural objects, namely seashells, which vary along a variety of object features, while others are shared across all stimuli. To correctly interact with objects, they have to be identified or categorized. For both processes, measuring similarities between objects is crucial. Our goal is to better understand the haptic similarity percept by comparing it to the visual similarity percept. First, direct similarity measures were analyzed using multidimensional scaling techniques to visualize the perceptual spaces of both modalities. We find that the visual and the haptic modality form almost identical perceptual spaces. Next, we performed three different categorization tasks. All tasks exhibit a highly accurate processing of complex shapes of the haptic modality. Moreover, we find that objects grouped into the same category form regions within the perceptual space. Hence, in both modalities, perceived similarity constitutes the basis for categorizing objects. Moreover, both modalities focus on shape to form categories. Taken together, our results lead to the assumption that the same cognitive processes link haptic and visual similarity perception and the resulting categorization behavior.

**Keywords** Vision and haptics · Similarity measures · Categorization · Perceptual spaces

## Introduction

For the first month after being born, babies are fully dependent on the mother, for being protected, for being fed, but also for exploring their surrounding. They can only see objects presented to them or passively watch their surrounding while being carried around. But soon, babies learn to actively manipulate objects by using their hands. Rotating the object enables the baby to view the object under different angles, and thus not only to form a three-dimensional representation but also to gather additional information about texture, weight, material, and absolute size of an object. If later on the baby wants to correctly manipulate an object, e.g., drink from a bottle, the object has to be recognized or at least categorized first. This is done by comparing the bottle to a mental representation of bottles. Importantly, we posit that this representation is multimodal, integrating several different object features like shape, color, weight, texture, and so on. This multimodality allows for better action planning, e.g., when the bottle should be lifted to the mouth for drinking and also for robust inference of object properties from visual input.

So far, research on mental representations mostly focused on the visual modality. Furthermore, research concerning the haptic modality mostly focused on low-level object features, such as curvature, edges, or texture granularity (Plaisier et al. 2009; van de Horst et al. 2008). In their seminal research, Klatzky et al. (1985) demonstrated that haptic categorization of every-day household objects was surprisingly efficient to the degree that they talked about an “expert system” for touch. In addition,

---

N. Gaissert  
Max Planck Institute for Biological Cybernetics, Tübingen,  
Germany

C. Wallraven (✉)  
Korea University, Seoul, Republic of Korea  
e-mail: wallraven@korea.ac.kr

Norman et al. (2004, 2008) have compared visual and haptic processing of natural object shapes (bell peppers) and found comparable performance for the two modalities. Here, we try to take a further step into understanding haptic perception of every-day, natural (as opposed to every-day, human-made) objects. For this, we selected a set of natural, complex objects—in our case, seashells from different families—to generate a set of objects that differ in a variety of object features like shape, color, and texture, while sharing other object features like material.

For identifying and categorizing objects, it is crucial to measure similarities between objects (Palmeri and Gauthier 2004). The most detailed analysis of similarity data is possible using multidimensional scaling (MDS) techniques. MDS takes distances between pairs of objects as input and returns coordinates of the objects and their relative positions in a multidimensional space. Using human similarity ratings, the MDS output map can be understood as a perceptual space (Borg and Groenen 1997; Shepard 1987). This perceptual space provides information about how many dimensions are apparent to the participants, about the weighting of these dimensions, and furthermore, makes it possible to infer whether these dimensions correspond to certain object features. Here, we are interested in determining the structure of this perceptual space of complex seashells as analyzed by similarity ratings and multidimensional scaling.

Assessing similarities between objects is a crucial component of categorization (Mervis and Rosch 1981; Goldstone 1994; Shepard 2001). As our second line of inquiry, we will, therefore, investigate categorization behavior of natural objects. More specifically, we will compare haptic categorization with visual categorization to gain insights into how the two modalities form categories of complex objects. Interestingly, by choosing natural seashells, we can also compare visual and haptic categorization behavior to categorization predicted by biological reasoning, i.e., to the biologically defined taxonomy of the seashells.

Finally, we will link the two studies in that we will compare the categorization behavior of both modalities to the underlying perceptual spaces that were analyzed in the first experiment. With this comparison, we want to test two hypotheses: first, we want to test whether Shepard's hypothesis (Shepard 2001), that objects of the same basic kind (that is, objects that share important features) generally form local regions in perceptual space, can hold on a small scale. Secondly, we want to test whether Edelman's hypothesis (Edelman 1999), that categories form clusters within a veridical perceptual space, will not only hold for *visual* object exploration as it was previously shown (Cutzu and Edelman 1998) but whether this hypothesis also holds for *haptic* object exploration.

Previous studies by Cutzu and Edelman (1998) showed that the visual modality can form a veridical perceptual space of a multidimensional physical object space, that is, that the visual modality correctly recovers the dimensionality of the parameter space and that the topology within the space is conserved. Furthermore, they showed that within such a veridical perceptual space, categories form clusters, i.e., that perceived similarity is higher for pairs of objects in the same category than for pairs in two different categories.

Although the visual and the haptic sensory systems rely on different types of input information, our previous study showed that the haptic modality is as capable of forming a veridical perceptual space (Gaissert et al. 2010). In this previous study, we generated a complex object space of parametrically defined shell-shaped objects spanning a three-dimensional object space. These objects were generated combining computer graphics modeling with 3D printing techniques. The shell-shaped form resulted from the fact that a biologically plausible model exists that allowed us to generate naturalistic yet well-defined objects, which varied along visually and haptically perceivable dimensions. All of these variations, however, were restricted to shape—which previous research has also shown to be the most informative feature for the visual (e.g., Mervis and Rosch 1981) and the haptic (Lederman and Klatzky 1990) modality for categorization. Here, we extend this research to natural objects that not only vary in shape but also in other features like color, patterning, etc. and that are much richer in detail than our computer-generated objects. With our data, we will analyze whether shape is also the dominant feature in categorizing our *natural* seashells, as well as for forming a perceptual space.

To summarize, in this study, we will investigate human haptic object perception of complex objects by analyzing the haptic similarity percept and comparing it to the similarity percept of the visual modality in a similarity rating task and a categorization task. Further, we will investigate whether natural objects form clusters within perceptual spaces, and if so, we will analyze whether these clusters can account for human categorization behavior.

## Stimuli

Since we were explicitly interested in the visual and haptic similarity percept of complex objects, we were searching for a set of stimuli belonging to a group of objects, but still varying in several object features like shape, color, patterning, etc. Further, we wanted to avoid man-made objects as used in other categorization studies (e.g., Haag 2011), since their design is dependent on the

purpose of the objects and the designer. Finally, we decided to gather 24 natural seashells that are rich in both visually and haptically perceivable object features (Fig. 1).

The seashells are taken from several superfamilies, which are members of two larger seashell classes (gastropods and bivalves, see also Fig. 4 for the full phylogenetic tree). Each superfamily is distinctly different due to variations in shape (for example, conical shells, turban-like shells, and cone-like shells), color (for example, darker, monochrome colors versus highly textured patterns), and texture (for examples, grooved versus smooth seashells). In choosing exemplars, we tried to make the variations within each superfamily explicit using the help of an expert on this subject matter.

## Perceptual space

### Similarity ratings

Visual and haptic similarity ratings were collected to visualize the perceptual spaces of both modalities. The task was to rate the similarity of pairs of objects on a seven-point scale from low similarity (1) to high similarity (7). Eleven participants with normal or corrected-to-normal vision performed the visual similarity ratings. Eleven other participants were blindfolded and performed the haptic similarity ratings, palpating the objects with both hands. All participants were naïve to the stimuli and were paid 8€ per hour.

The experiment was started by introducing the objects to the participants. Every object was presented to the



**Fig. 1** Twenty-four natural sea shells: We chose 20 gastropods. Four objects have a conical shell: *Patella barbara*<sup>1</sup>, *Patella longicosta*<sup>2</sup>, *Patella granularis*<sup>3</sup>, and *Patella vulgata*<sup>4</sup>. Four shells have a turban-like shell: *Turbo argyrostomus*<sup>5</sup>, *Turbo coronatus*<sup>6</sup>, *Turbo crassus*<sup>7</sup>, and *Turbo setosus*<sup>8</sup>. Four objects are extremely smooth and shiny: *Cypraea eglantine*<sup>9</sup>, *Cypraea histrio*<sup>10</sup>, *Cypraea lynx*<sup>11</sup>, and *Ovula ovum*<sup>12</sup>. Four members of the olive shells were selected: *Oliva*

*irisans*<sup>13</sup>, *Oliva miniacea*<sup>14</sup>, *Agaronia gibbosa*<sup>15</sup>, and *Olivancillaria vesica auricularia*<sup>16</sup>. Four objects have a cone-like shell: *Conus figulinus*<sup>17</sup>, *Conus malacanus*<sup>18</sup>, *Conus marmoreus*<sup>19</sup>, and *Conus textile*<sup>20</sup>. Every group of four is a group of objects belonging to the same superfamily. Further, we chose four bivalve molluscs from different superfamilies: *Mactra stultorum*<sup>21</sup>, *Pecten maximus*<sup>22</sup>, *Acanthocardia tuberculata*<sup>23</sup>, and *Glycymeris insubrica*<sup>24</sup>

participants in a randomized order. In the visual modality, one object was placed on a black plateau, a black curtain was automatically opened, and the participant was able to explore the object visually for 12 s before the curtain closed automatically. During this time, two different perspectives of the object were presented to show all features of the object. For haptic exploration, the object was placed on the same plateau. A beep gave the signal to start the haptic exploration. Fifteen seconds later, a second beep signaled the end of the exploration. More time was given in the haptic modality to allow participants to sample all potentially informative stimulus properties—as is common practice in visual–haptic experiments (Lacey et al. 2009a, b).

Participants were allowed to palpate the objects in a very natural way, with both hands and no restrictions to the exploratory procedure, knowing that different exploratory procedures can influence similarity perception (Cooke et al. 2010). Hence, participants were totally free in focusing on every object feature that might be deemed relevant for the task.

In the experimental trials, every object was paired once with itself and once with every other object. The pairs of objects were shown in randomized order. Every participant had to rate every pair just once, because previous experiments showed that the judgments did not vary over repetitions (Gaissert et al. 2010). In both modality conditions, objects were placed on the plateau successively. In the visual modality, the curtain was opened for 6 s, and the object was rotated by the experimenter to afford a fuller visual exploration of the object. The curtain was then closed, and the object was exchanged for the second member of the object pair. The curtain was opened again, and the exploration phase was repeated. After the second object, the participant was asked to say out loud the similarity rating, which was then recorded by the experimenter. In the haptic modality, the two objects were also presented successively with beeps signaling the beginning and the end of the exploration, which lasted for 8 s for each object.

After performing similarity ratings, participants had to answer a questionnaire in which they were asked to rate how strongly they relied on special object features to perform their similarity ratings. The answers were given on a scale from 0 (feature is not important at all) to 6 (feature is very important) (Fig. 3). These questionnaires were analyzed to better understand which object features participants used for forming their perceptual spaces.

## Analysis

First, we wanted to analyze whether there is a difference in task difficulty between the visual and the haptic modality. Therefore, we counted how often an object was directly paired with itself and how often participants correctly

identified this match (by rating it with a 7). Since the data is not normally distributed, we performed Mann–Whitney  $U$  tests.

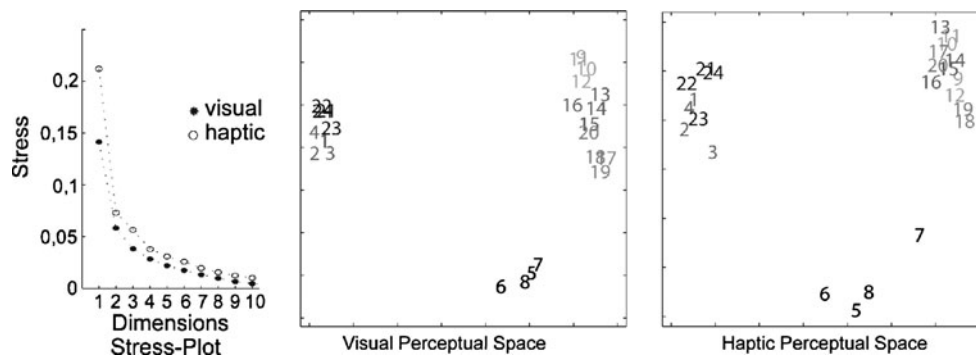
Next, the raw similarity ratings were correlated to determine the degree of variation across participants. The similarity ratings (ranging from 1 to 7) were then converted to dissimilarities (by subtracting similarities from 7) that were then averaged for both modalities over all participants and all trials. The correlation between average dissimilarity matrices of both modalities was calculated.

For the subsequent multidimensional scaling (MDS) analysis, we used the non-metric MDS algorithm (MDSSCALE) in MATLAB. Non-metric MDS takes the rank order of the pairwise proximity values into account, and thus fits the human similarity data better than classical, metric MDS (Cooke et al. 2007). To determine how many dimensions were necessary to explain the data, the stress values from one to ten dimensions were plotted. An “elbow” in the plot indicates how many dimensions are sufficient to explain the data. The elbow is visible at two dimensions, both for the visual and the haptic data. Hence, we plotted the perceptual spaces for two dimensions (Fig. 2). A goodness-of-fit-criterion between the perceptual spaces of both modalities was calculated using the *procrustes* function of MATLAB. This function fits two sets of points by performing linear transformations (translation, reflection, and orthogonal rotation; note that these are valid operations, since MDS does not yield any absolute positions, but relative positions in space). Its output represents the distance between the points of both sets as the sum of squared errors. Low values, therefore, indicate a better fit than high values, with a value of 0 for a perfect fit.

Correlations of the dimensions of the visual and the haptic perceptual spaces were used to look for differences in the saliency of the first two dimensions. We then tried to find shape features determining these dimensions by investigating the objects and how potentially salient features change along the dimensions. In addition, we also analyzed the questionnaires filled in by the participants.

## Results

Two groups of eleven participants rated the similarity between pairs of objects in a visual and haptic condition, respectively. To test for the task difficulty, the amount of correctly identified match trials was calculated. Visually participants were more often correct than haptically (median<sub>visual</sub> = 96%, median<sub>haptic</sub> = 83%,  $U = 27.5$ ,  $P = 0.03$ ). Haptic performance, therefore, seems more prone to errors in this case. However, if one would be a bit more lenient in the criterion, and count values of 6 and 7 as “correct” answers for identical stimuli, then both modalities score 100% correct.



**Fig. 2** Perceptual spaces of natural sea shells. On the *left*, the stress values are displayed for one to ten dimensions. The elbow indicates that two dimensions explain the data sufficiently for the visual and the haptic modality. In the *middle*, the two-dimensional visual perceptual

space is displayed; on the *right*, the haptic perceptual space is displayed. Objects are numbered according to Fig. 1. Shells within one column of Fig. 1 are closely related and are marked with the same shade

The raw similarity ratings were then correlated across participants, yielding an  $11 \times 11$  matrix for each condition. The mean correlation values for this raw data were  $r_{\text{visual}} = 0.823$  ( $r_{\text{visual,min}} = 0.690$ ;  $r_{\text{visual,max}} = 0.917$ ) and  $r_{\text{haptic}} = 0.822$  ( $r_{\text{haptic,min}} = 0.750$ ;  $r_{\text{haptic,max}} = 0.900$ ). The high mean correlation value in both conditions indicates a high degree of inter-participant consistency—in addition, the values are virtually identical for both conditions, suggesting a similar degree of consistency across the two modalities.

To further analyze the similarity ratings, the ratings were averaged across participants. The correlation between the visual and the haptic similarity ratings is very high ( $r = 0.967$ ,  $P = 0.000$ ) and shows that humans perceive similarities between those complex objects visually and haptically in a very similar fashion.

Using the average matrices for both modalities, we performed MDS for one to ten dimensions and plotted the stress values (Fig. 2). Interestingly, the elbow in the stress plot indicates that visually and haptically, participants mostly relied on two dimensions. Given this data, we decided to plot the MDS output maps for two dimensions (Fig. 2). For the haptic data, one might also go up to four dimensions. These other dimensions might then also play a (minor) role in judging similarities. We will have a closer look at these details in further analyses.

Looking at the two-dimensional MDS maps already reveals that the visual and the haptic perceptual spaces are astonishingly similar. To better compare the perceptual spaces of both modalities, we calculated the goodness-of-fit criterion. The perceptual spaces were found to be highly similar (goodness-of-fit of only  $d = 0.057$ , where 0 would indicate perfect alignment). This result again highlights the astonishing fact that our haptic modality is able to precisely recover the same perceptual space as the visual modality, although humans have rather little experience in touching shell-shaped objects.

As can be seen, the perceptual spaces of both visual and haptic exploration are not only highly congruent but also exhibit a very consistent clustering of the different stimulus groups. Figure 2 shows that visually, as well as haptically, participants form three clusters of object shapes. The first cluster contains objects 1–4 and 21–24. Although objects 1–4 are gastropods, whereas objects 21–24 are bivalves, the proximity within the perceptual space can be explained by the fact that all of these shells show no convolutions, whereas all other shells have a distinct convolution.

Objects 5–8 form their own cluster in the perceptual spaces, while objects 9–20 form a large cluster within the visual and the haptic perceptual spaces. There are several features that can explain this clustering pattern. The first feature is the form of the aperture: objects 5–8 have a circular aperture, while the aperture of objects 9–20 resembles a groove. This property is also closely related to the tip of the shell. All shells with a circular aperture have a pronounced tip, while the tip is less pronounced for shells 13–20, having a groove-like aperture. Objects 9–12 do not even have a tip, but have a very pronounced groove as aperture. The second feature is the shape of the convolutions, which for objects 5–8 results in their distinct “turban” shape—a feature, which the other shells do not possess.

Given these observations, we tried to relate the dimensions of the perceptual spaces to object features. Visually as well as haptically, the first dimension divides shells with convolutions from shells without convolutions (flat shells). The second dimension, again visually as well as haptically, may be related to the form of the aperture or the form of the convolutions, which splits off the turban-like shells from the other families. So far, it was not possible to correlate further dimensions to obvious object features.

Interestingly, the first two dimensions of the visual and the haptic perceptual spaces both correlate to shape features of the seashells. Thus, in both modalities, participants analyzed the shape of objects to rate similarities, while color,

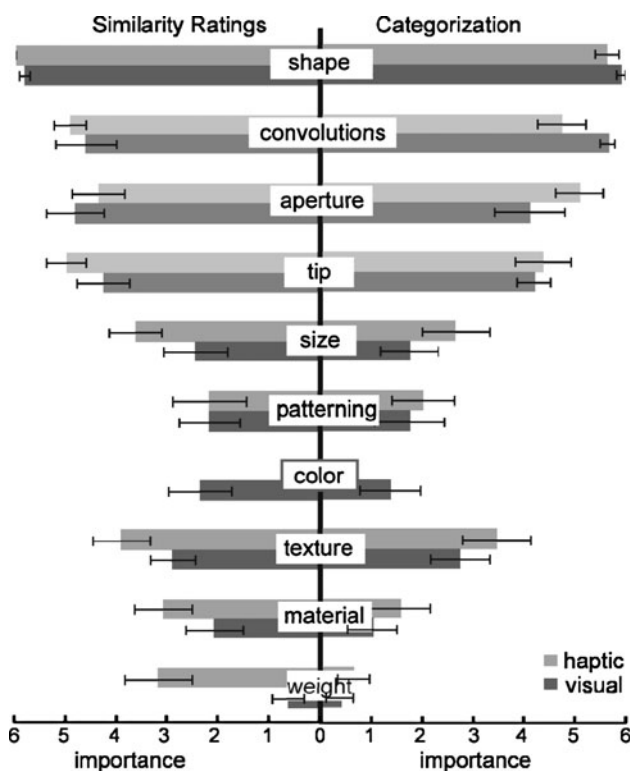
patterning, and texture only played a minor role. This is confirmed by the questionnaires participants had to fill in after the experiments (Fig. 3). Participants rated shape as more important than size, patterning, color, texture, material, and weight, while the different shape features (convolutions, aperture, and tip) were rated as equally important.

In summary, we showed that participants formed highly congruent perceptual spaces when they explored the natural seashells visually or haptically. Moreover, to judge similarities, participants mostly focused on shape features. In the next section, we will analyze whether participants still focus on shape when they are asked to categorize objects.

## Categorization

### Categorization tasks

In the previous section, we compared the visual and haptic similarity percept of natural seashells by looking at similarity ratings and by visualizing the perceptual spaces of



**Fig. 3** Questionnaires. Participants were asked to rate the importance of the following object features for performing similarity ratings (*left side*) and for categorizing the objects (*right side*): shape, size, patterning, color, texture, material, and weight (*dark colors*). Since we expected shape to play a major role, we asked for more details concerning shape: convolutions, aperture, and tip (*bright colors*). *Upper bars* represent haptic data; *lower bars* represent visual data. *Bars* represent mean ratings across participants (0 = not important, 6 = very important); *Error bars* represent SEM

both modalities. We found that both spaces show a clear clustering of the objects within the space. This raises the question whether participants would actually create the categories that correspond to the clusters in the perceptual space. As participants' similarity ratings mostly focused on shape, we were also interested in the question whether shape would determine the categorization behavior. We therefore performed three categorization tasks using the 24 seashells displayed in Fig. 1.

About 2 weeks after performing the similarity ratings on natural seashells, the same participants were asked to perform the categorization tasks. Again, they were paid 8€ per hour.

Participants had to form 2, 3, and 6 groups of objects. We asked for two groups, since the stimulus set consisted of bivalve molluscs and gastropods, two different classes of molluscs. Since the stimulus set consisted of six different superfamilies of molluscs, we also asked participants to form six groups of objects. Moreover, we asked for three groups of objects, since the perceptual spaces clearly showed three clusters.

Participants were seated in front of a black table with sound-absorbing surface. In the visual condition, participants saw the objects, but were not allowed to touch them. During the whole experiment, participants could ask the experimenter to present the objects to them from different angles of view. By pointing on the objects, participants instructed the experimenter which object to present and where to place the objects on the table. In the haptic condition, participants were blindfolded and were allowed to use both hands for exploring the objects. Here, the experimenter took care that participants did not miss any objects or drop any objects from the table.

All 24 objects were spread on the table. Participants were then instructed to explore the objects. When they reported that they had explored all objects sufficiently, they were instructed to categorize the objects by forming  $n$  ( $n = 2, 3$ , or 6 in randomized order) groups on the table. In the visual condition, participants pointed to the objects and the experimenter moved the objects around, while in the haptic condition, participants formed the groups on their own by moving the objects appropriately. When the participants were finished with forming the groups, the experimenter recorded the grouping before shuffling the objects. Participants were then instructed to form  $n$  groups again. This procedure was repeated until participants formed the same groups twice in a row and guaranteed that the final groups were stable. After performing the categorization experiments, participants were again asked to rate the importance of the different object features (Fig. 3).

### Analysis

The first analysis was to look at the task difficulty by counting how many repetitions participants needed to form

the same groups twice in a row. Next, the correlation between visual and haptic performance was determined. For this, we took the categorization of the final block specifying which object belonged into which category. This vector was then re-encoded, so that category numbers were consistent across participants. The resulting vectors were then correlated. To analyze the actual categorization behavior, we counted how often an object was paired with another object for all participants. Based on this matrix, we then calculated how an average person would categorize the objects. For this prediction, the *single linkage* algorithm of MATLAB was used with default parameters. The resulting categories are visualized in Fig. 4.

In addition, we used the same *linkage* algorithm to predict how participants would categorize the objects if they would purely rely on similarity measures; that is, we asked whether the similarity measures collected in the previous experiment would be able to predict human categorization behavior. For this, the *similarity ratings* of the natural seashells were averaged across participants, and this matrix was used for the prediction. This grouping is also displayed in Fig. 4. In addition, Fig. 4 also contains a grouping of the objects based on the biologically defined taxonomy to see whether biological reasoning might be able to account for human object categorization.

Finally, we calculated the distance between all pairs of objects for the perceptual spaces shown in Fig. 2. These distances were averaged across pairs within the same category, as well as across pairs in different categories. We performed the calculation for the perceptual space being divided into two, three, or six categories visualized in Fig. 5 for the visual and the haptic modality. The calculation was done to test whether the hypothesis (see, e.g., Edelman 1999) that objects within one category are closer in visual perceptual space than objects from different categories is also true for our seashells; and, more interestingly, whether this hypothesis also holds for the haptic modality.

## Results

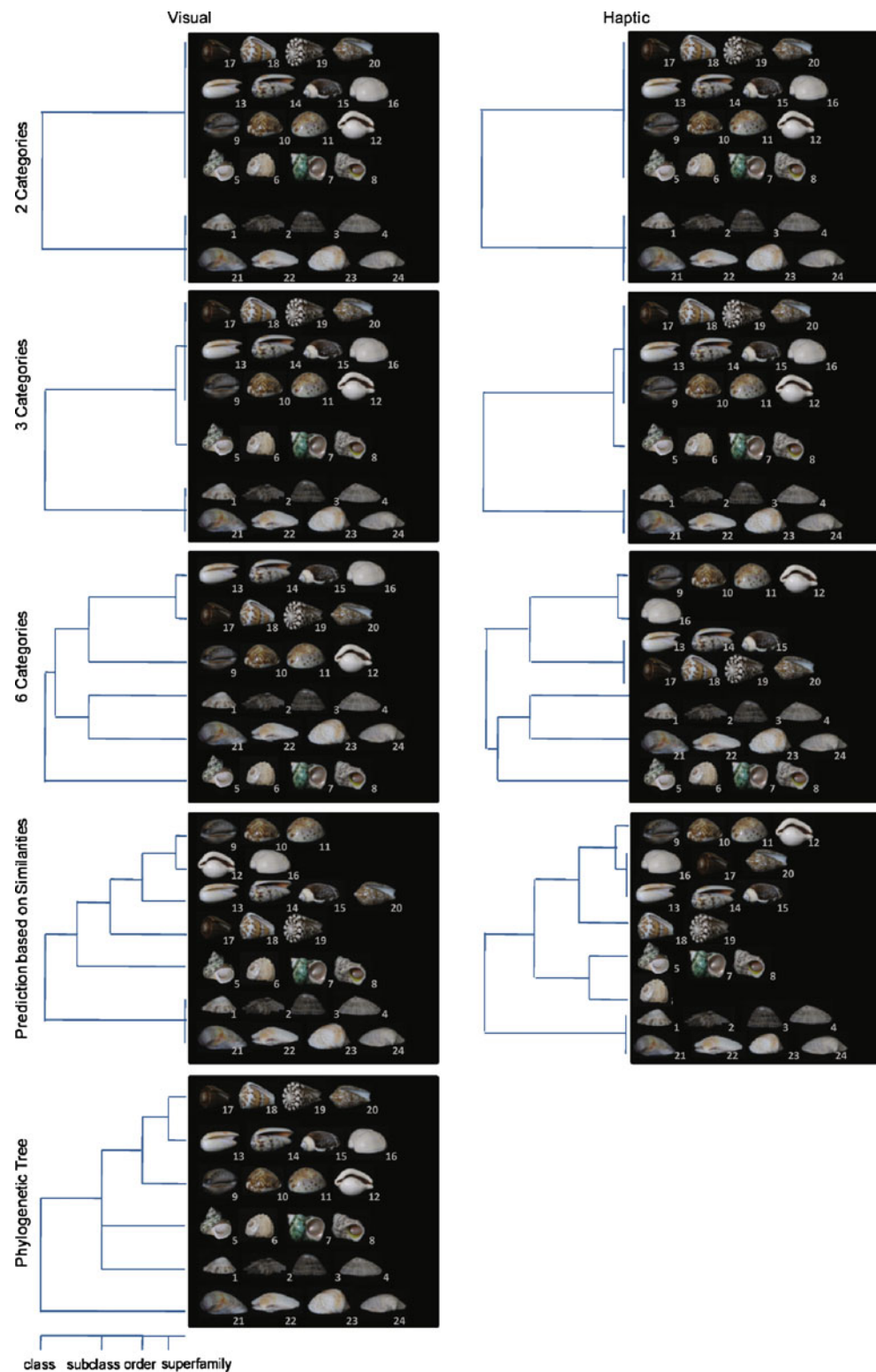
To determine whether categorizing objects is equally hard visually as haptically, we counted how many repetitions were necessary to form the same groups twice in a row. A Mann–Whitney  $U$  test showed that the task is significantly harder to perform in the haptic modality than in the visual modality ( $\text{mean}_{\text{haptic}} = 2.39$ ,  $\text{mean}_{\text{visual}} = 2.03$ ,  $\text{median}_{\text{haptic}} = 2.00$ ,  $\text{median}_{\text{visual}} = 2.00$ ,  $U = 427.00$ ,  $P = 0.012$ ). Moreover, we found that forming six categories is significantly more difficult than forming two ( $U = 177.5$ ,  $P = 0.03$ ) or three categories ( $U = 174.5$ ,  $P = 0.02$ ;  $\text{mean}_2 = 2.09$ ,  $\text{mean}_3 = 2.05$ ,  $\text{mean}_6 = 2.5$ ,  $\text{median}_2 = 2.00$ ,  $\text{median}_3 = 2.00$ ,  $\text{median}_6 = 2.00$ ), while forming two or three categories seems equally hard.

Next, the correlation between visual and haptic performance was determined. When participants were asked to form two categories, the visual and haptic performance highly correlates ( $r_2 = 0.85$ ). For forming three categories, the correlation is even higher ( $r_3 = 0.91$ ), while forming six categories results in a lower correlation ( $r_6 = 0.71$ ). All correlations are highly significant ( $P = 0.000$ ), stressing that visual and haptic object exploration result in highly similar performance.

Finally, we examined the categorization behavior in more detail. Figure 4 visualizes how an average participant would categorize the 24 objects. For forming two categories, visual and haptic exploration results in exactly the same two groups: one group of flat shells without convolution (objects 1–4 + 21–24) and one group with convoluted shells (objects 5–20), stressing how similar the visual and haptic shape percept is. The same consistency is found when forming three groups with visual and haptic exploration resulting in the same categories: one group of flat, unconvoluted shells (objects 1–4 + 21–24), one group of convoluted shells with pronounced tip and circular aperture (objects 5–8), and one group of convoluted shells with almost no tip and a groove-like aperture (objects 9–20). When looking at six categories, there are some differences between visual and haptic categorization: visually, the group of convoluted shells with circular aperture (objects 5–8) is distinct from all other shells, while haptically, these objects are associated with the flat shells. For both modalities, the flat shells are split up into two groups: the bivalves (21–24) and the gastropods (1–4). Visually, the group of convoluted shells with groove-like aperture is split up into three groups of four objects (9–12, 17–20, and 13–16) with the two last groups being closely related (13–20). Haptically, the convoluted shells with groove-like aperture form one larger group (1–15 + 17–20). Object 16 forms its own group, but is associated to objects 9–12. Overall, however, visual categorization behavior is still similar to haptic categorization, with the difference between object 16 to objects 13–15 being perceived as more important in the haptic modality than in the visual modality.

The categorization pattern suggests that participants mainly focused on shape to categorize the objects. This is especially prominent when looking at the two-category condition. Object 6 should be grouped with object 1–4 + 21–24 based on *material* properties (a matte rough surface), but actually was categorized with objects 5–20 with which it shares the same shape. Another example for the shape dominance is object 7, which based on color should form its own group, and based on shininess should be categorized with objects 9–11 + 13–15. The data, however, show that it is grouped together with objects 5, 6, and 8 that again share the same form. Note again that this reliance on shape is present not only for the haptic

**Fig. 4** Categorization behavior. The figure visualizes how an average participant would categorize the 24 objects when asked to form 2, 3, or 6 categories. Moreover, the figure shows how categorization is predicted by the similarity ratings. The left side shows the visual data; the right side shows the haptic data. Finally, the figure also shows how categorization should be performed based on biological reasoning

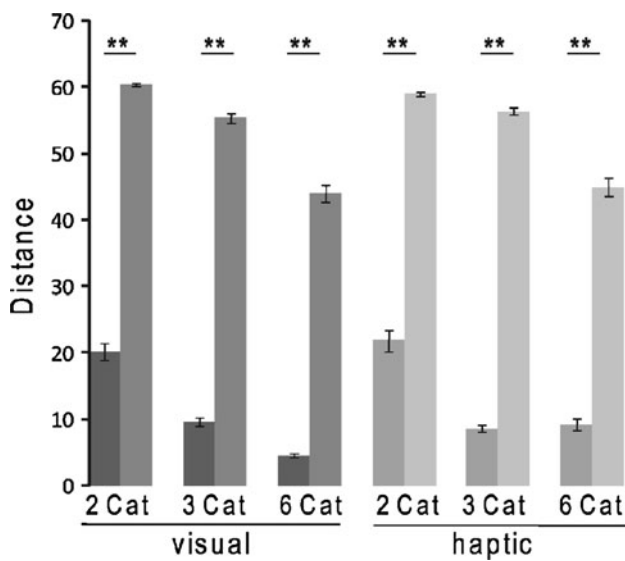


condition (where it might be expected) but also for the visual condition.

This finding about the importance of shape is supported by the questionnaires. Participants had to fill in the same questionnaires after the categorization experiment as for

the similarity rating experiment (Fig. 3). In both cases, they rated shape as more important than size, patterning, color, texture, material, or weight. The fact that participants rated the importance of the object features in a very similar fashion after both experiments suggests that they followed





**Fig. 5** Distances within pairs of objects. The distance between pairs of objects was measured based on the two-dimensional perceptual spaces. These distances were averaged across pairs lying within one category (*dark bars*) and across pairs lying in different categories (*bright bars*). In every case, the difference between within (*dark bars*) and across (*bright bars*) is highly significant even for Bonferroni-corrected *P*-values (\*\* represents  $P < 0.0001$ ). Error bars represent SEM

similar strategies in the two tasks, which were mainly driven by shape judgments.

We now turn to category prediction based on similarity ratings (Fig. 4). Both for two and for three categories, the similarity ratings exactly predict how participants categorize the objects visually and haptically (two groups: 1–4 + 21–24, and 5–20; three groups: 1–4 + 21–24, 5–8, and 9–20; valid for both modalities).

For forming six categories, however, the prediction does not match the categorization behavior anymore, both for the visual and the haptic modality (see Fig. 4). In both modalities, the similarity ratings would predict one big group of flat shells (1–4 + 21–24) and split up the convoluted shells into several small groups, which results in some differences between the predicted small groups and the actually formed groups. This may raise the question whether participants really based their categorization behavior on exactly the same object features they based their similarity judgments on, or whether they followed different strategies in the two tasks. The prediction does seem to work extremely well for two and three categories, which speaks against the use of different strategies. In contrast, for six categories, the cognitive load might increase such that participants are more likely to turn to rule-based categorization behavior, which is known to deviate from the more lower-level similarity judgments (Cooke et al. 2007). Finally, the differences between the objects might be simply too small in comparison to the

noise inherent in the human data in order for the algorithm to correctly predict the six categories. Overall, further experiments are needed to address this question in more detail—from our data, we tentatively conclude that the similarity data, indeed, can be used to predict the categorization behavior successfully.

At the beginning, we also raised the question whether biological reasoning can explain human categorization behavior. The phylogenetic tree is shown in Fig. 4 for comparison. It is especially striking that both visual and haptic categorizations do not distinguish between classes of molluscs. Taxonomically, there is the class of gastropods (1–20) and the class of bivalves (21–24), while perceptually, participants distinguish between flat (1–4 + 21–24) and convoluted shells (5–20). Not considering the relationships between the different groups, visual categorization correctly identifies the groups on the subclass (1–4, 5–8, 9–20, 21–24), the order (1–4, 5–8, 9–12, 13–20, 21–24), and the superfamily level (1–2, 5–8, 9–12, 13–16, 17–20, 21–24). The haptic modality still recovers the subclass level correctly, while it does not recover the order and the superfamily, which mainly results from the “biologically incorrect” grouping of one object (object 16).

Finally, we analyzed the degree of cluster definition in the perceptual space. For this, we checked whether the similarity between pairs of objects within one category is higher than between pairs of objects from different categories. Our data show exactly this effect for the visual data (see Fig. 5). The distances between pairs of objects within one category are always significantly smaller than the distances of pairs in different categories. This is valid for two, three, and six categories. Interestingly, the same is also true for our haptic data. Thus, we can conclude that for both, visual and haptic object exploration, categories form clusters within the perceptual space.

## Conclusion

Our line of research starts with looking at perceptual spaces of complex objects. Previously, we were able to show that the visual and the haptic modality form highly congruent perceptual spaces when the objects vary along shape dimensions only (Gaissert et al. 2010). Here, we were able to extend this finding to a set of natural objects, namely seashells. The seashells are rich in details and vary in different shape features, but in addition to that, they vary in color, patterning, texture, and other object features. Interestingly, visual and haptic object exploration still lead to the formation of highly congruent perceptual spaces, which is only possible since participants in the visual and in the haptic condition focused mostly on shape features to rate similarities. This shape dominance also emerged in our

categorization tasks (as also described by Lederman and Klatzky 1990) and was supported further by the questionnaires filed by the participants (see Fig. 3).

This finding raises the interesting question, why humans perceive shape to be more informative than other object features. Already in 1758, when Carl von Linné published the *Systema Naturae* (Linnaeus 1758), he focused on shape as an important feature to sort plants and animals to form an encompassing taxonomy. In the questionnaires, our participants were asked why they focused on shape. Several participants reported that shape is simply more informative for categorizing objects, since it is more robust against changes both on a short-time scale, referring to the life span of one animal, as well as on a long-time scale, referring to evolution of a new species. If this is an effect of the western education system (Heinrich et al. 2010), teaching students about evolution, or if humans really perceive shape as more informative to judge biological relations, represents an interesting question for future studies. However, it seems clear that function and physical parameters result in the same shape. This becomes clear when comparing the streamlined body of a dolphin and a fish, both evolved to swim through water.

We showed here that visual and haptic object exploration lead to almost identical perceptual spaces, which raises the question whether participants form one multimodal perceptual space or whether highly congruent but unimodal spaces are formed. If one representation is formed that is accessible to both modalities, one might expect cross-modal shape comparisons to be equivalent in performance to unimodal shape comparisons. In this context, Norman et al. (2004) reported not only high accuracy but also significant performance differences, and thus speaks of overlapping but distinguishable representations. In contrast, excellent cross-modal priming behavior (Bushnell and Baxt 1999; Reales and Ballesteros 1999) supports a multimodal space. Neuroimaging research also provides evidence for common neural substrates in visual and haptic object processing. Sathian et al. (1997) showed for the first time that tactile discrimination can recruit visual areas. In Amedi et al. (2001), fMRI was used to show that the ventral visual pathway was active in multi-sensory object processing. Following Lacey et al. (Lacey et al. 2009a, b), the LOTv, a subregion of LOC, contains this modality-independent representation of geometric shape (see also James et al. (2007) for a recent review of the neural underpinnings of visuo-haptic processing). In addition, similar objects evoke similar response patterns in LOC, whereas shapes perceived as more different are associated with more different response patterns (Op de Beeck et al. 2008), which points toward a possible implementation of how the perceptual space might be represented.

Since both the visual and the haptic perceptual spaces exhibited clear clusters, we became interested in how participants would categorize the natural seashells. Forming two and three categories result in exactly the same groups of objects for both the visual and the haptic modality, showing a high degree of similarity between the two modalities. Even when forming six categories, there were only minor differences between the two modalities. The high correlation between visual and haptic performance in the similarity rating task and in the categorization tasks seems to suggest that the same processes underlie visual and haptic similarity perception.

Throughout the experiments, we found highly efficient performance of the haptic modality, which is especially noteworthy when compared to the haptic performance in recognition of 2D raised-line depictions, where touch alone performs quite poorly (Lederman et al. 1990; Loomis et al. 1991). We assume that this difference is mostly due to the fact that we used natural 3D shapes in our experiments, which the haptic modality would have evolved to perceive optimally. In addition to that participants were allowed to palpate the objects in a very natural way, with both hands and no restrictions to the exploratory procedure. The good performance for haptic exploration of natural shapes is in line with the earlier results by Norman et al. (2004, 2008), in which haptic performance for natural shapes was on par with visual performance.

Our set of natural objects allows for a comparison of human categorization behavior to categorization predicted by biological reasoning. Here, neither visual nor haptic object exploration lead to a correct identification of biological relations between objects. However, the similarity percept can almost perfectly predict categorization behavior both in the visual and in the haptic modality, which highlights that similarity is an important factor for categorization in both modalities (note that generic categorization behavior can in some cases go beyond similarity judgments, such as in rule-based categorization (Hahn and Ramscar 2001)).

Next, we compared the categorization behavior to the perceptual spaces to test two hypotheses. The first hypothesis was formulated by Shepard (2001). He stated that objects of the same basic kind generally form local regions in perceptual spaces. Here, we were able to prove this hypothesis to be true for our set of natural objects. Future experiments with an extended set of stimuli will show whether this hypothesis will hold on a larger scale.

The second hypothesis was formulated by Edelman (1999). He claimed that categories form clusters within a veridical perceptual space (i.e., that object pairs from one category are closer within this perceptual space than objects from different categories) and showed this for visual exploration only (Cutzu and Edelman 1998). In our

previous study, we showed that the haptic modality can form a veridical perceptual space and even exceeds the visual modality in recovering the topology of the underlying physical object space (Gaissert et al. 2010). Here, we showed that object pairs of the same category are significantly closer than object pairs from different categories. Taken together, we thus proved Edelman's hypothesis to be true not only for the visual modality but also for the haptic modality.

Overall, we found the same link between perceptual spaces and categorization behavior to occur in the visual and in the haptic system, leading to the assumption that the same cognitive processes underlie visual and haptic object categorization. It should be noted, of course, that these conclusions are only valid for the object classes and feature variations tested so far, which do not yet span all combinations of possible object features. However, considering the complexity of the object classes tested in the literature (irregularly shaped bell peppers in Norman et al. (2004, 2008), geometric objects containing shape and texture in Cooke et al. (2007), mathematically defined shell-like objects in Gaissert et al. (2010), and natural seashells in the present study), it seems that there is considerable evidence for similar (shape) processing of complex object classes in the visual and haptic modalities.

In this study, we investigated haptic object perception by comparing it to visual object perception. We found an astonishing congruency between the two modalities. Most interestingly, we have shown that the haptic modality is capable of precise processing of complex objects, which is based on shape features for rating similarities and for categorizing objects. Our findings are based on natural 3D objects, varying in several object features. Some of these dimensions like shape can be perceived by the visual and the haptic modality; others, like color and weight can only be perceived by one modality. Our brain integrates this information to form a multidimensional object representation. This representation is afterward used to correctly interact with the object. Only if we better understand this multidimensional representation, we will also better understand how humans interact and handle every-day objects. This is especially important for designing new technologies in the field of haptic machine interfaces. If we better understand which object features are important for correctly interacting with the objects, e.g., a tool, then it is also easier to design an intuitive haptic interface that provides the relevant haptic information to the user.

**Acknowledgments** All seashells are items of loan from the natural history museum Stuttgart, Germany (Am Löwentor, Staatliches Museum für Naturkunde Stuttgart). We thank Hans-Jörg Niederhöfer for providing the seashells and helping to select an adequate set of stimuli. This study was partially 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).

## References

- Amedi A, Malach R, Hendler T, Peled S, Zohary E (2001) Visuo-haptic object-related activation in the ventral visual pathway. *Nat Neurosci* 4(3):324–330
- Borg I, Groenen P (1997) *Modern multidimensional scaling*. Springer, Berlin
- Bushnell EW, Baxt C (1999) Children's haptic, cross-modal recognition with familiar, unfamiliar objects. *J Exp Psychol Hum Percept Perf* 25(6):1867–1881
- Cooke T, Jäkel F, Wallraven C, Bühlhoff HH (2007) Multimodal similarity and categorization of novel, three-dimensional objects. *Neuropsychologia* 45(3):484–495
- Cooke T, Wallraven C, Bühlhoff HH (2010) Multidimensional scaling analysis of haptic exploratory procedures. *ACM Trans App Percept* 4:1–22
- Cutzu F, Edelman S (1998) Representation of object similarity in human vision: psychophysics, a computational model. *Vision Res* 38(15–16):2229–2257
- Edelman S (1999) *Representation and recognition in vision*. MIT Press, Cambridge
- Gaissert N, Wallraven C, Bühlhoff HH (2010) Visual and haptic perceptual spaces show high similarity in humans. *J Vis* 10(11:2):1–20
- Goldstone RL (1994) The role of similarity in categorization: providing a groundwork. *Cognition* 52(2):125–157
- Haag S (2011) Effects of vision and haptics on categorizing common objects. *Cognit Proc* 12(1):33–39
- Hahn U, Ramscar M (2001) Similarity and categorization. In: Hahn U, Ramscar M (eds) *Similarity and categorization*. Oxford University Press, Oxford
- Heinrich J, Heine SJ, Norenzayan A (2010) Most people are not weird. *Nature* 466:29
- James TW, Kim S, Fisher JS (2007) The neural basis of haptic object processing. *Can J Exp Psychol* 61(3):219–229
- Klatzky RL, Lederman SJ, Metzger VA (1985) Identifying objects by touch: an “expert system”. *Percept Psychophys* 37:299–302
- Lacey S, Tal N, Amedi A, Sathian K (2009a) A putative model of multisensory object representation. *Brain Topogr* 21(3–4):269–274
- Lacey S, Pappas M, Kreps A, Lee K, Sathian K (2009b) Perceptual learning of view-independence in visuo-haptic object representations. *Exp Brain Res* 198(2–3):329–337
- Lederman SJ, Klatzky RL (1990) Haptic classification of common objects: knowledge-driven exploration. *Cogn Psychol* 22(4):421–459
- Lederman SJ et al (1990) Visual mediation and the haptic recognition of two-dimensional pictures of common objects. *Percept Psychophys* 47(1):54–64
- Linnaeus C (1758) *Systema Naturae per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species cum Carateribus, Differentiis, Synonymis, Locis*. Editio Decima, Reformata, vol 1 Regnum Animale. Laurentii Salvii, Stockholm
- Loomis JM, Klatzky RL, Lederman SJ (1991) Similarity of tactual and visual picture recognition with limited field of view. *Perception* 20(2):167–177
- Mervis CB, Rosch E (1981) Categorization of natural objects. *Ann Rev* 32:89–115
- Norman JF, Norman HF, Clayton AM, Lianekhammy J, Zielke G (2004) The visual and haptic perception of natural object shape. *Percept Psychophys* 66(2):342–351

- Norman JF, Clayton AM, Norman HF, Crabtree CE (2008) Learning to perceive differences in solid shape through vision and touch. *Perception* 37(2):185–196
- Op de Beeck HP, 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
- Palmeri TJ, Gauthier I (2004) Visual object understanding. *Nat Rev Neurosci* 5(4):291–303
- Plaisier MA, Tiest WM, Kappers AM (2009) Salient features in 3-D haptic shape perception. *Atten Percept Psychophys* 71(2):421–430
- Reales JM, Ballesteros S (1999) Implicit and explicit memory for visual and haptic objects: cross-modal priming depends on structural descriptions. *J Exp Psychol Learn Mem Cognit* 25(3):644–663
- Sathian K, Zangaladze A, Hoffman JM, Grafton ST (1997) Feeling with the mind's eye. *Neuroreport* 8(18):3877–3881
- Shepard RN (1987) Toward a universal law of generalization for psychological science. *Science* 237(4820):1317–1323
- Shepard RN (2001) Perceptual-cognitive universals as reflections of the world. *Behav Brain Sci* 24(4):581–601 (discussion 652–671)
- van der Horst BJ, Kappers AM (2008) Using curvature information in haptic shape perception of 3D objects. *Exp Brain Res* 190(3):361–367