

Visual and Haptic Representations of Material Properties

Elisabeth Baumgartner^{*,1}, Christiane B. Wiebel¹ and Karl R. Gegenfurtner

Abteilung Allgemeine Psychologie, Universität Gießen, 35394 Gießen, Germany

Received 22 March 2013; accepted 17 September 2013

Abstract

Research on material perception has received an increasing amount of attention recently. Clearly, both the visual and the haptic sense play important roles in the perception of materials, yet it is still unclear how both senses compare in material perception tasks. Here, we set out to investigate the degree of correspondence between the visual and the haptic representations of different materials. We asked participants to both categorize and rate 84 different materials for several material properties. In the haptic case, participants were blindfolded and asked to assess the materials based on haptic exploration. In the visual condition, participants assessed the stimuli based on their visual impressions only. While categorization performance was less consistent in the haptic condition than in the visual one, ratings correlated highly between the visual and the haptic modality. PCA revealed that all material samples were similarly organized within the perceptual space in both modalities. Moreover, in both senses the first two principal components were dominated by hardness and roughness. These are two material features that are fundamental for the haptic sense. We conclude that although the haptic sense seems to be crucial for material perception, the information it can gather alone might not be quite fine-grained and rich enough for perfect material recognition.

Keywords

Material perception, vision and haptics, perceptual space

1. Introduction

Material perception is inherently multimodal in nature. Touching objects in order to discover what they feel like is a very natural thing to do. When we want to judge the ripeness of a fruit we not only look at it but also feel very

* To whom correspondence should be addressed. E-mail: elisabeth.baumgartner@psychol.uni-giessen.de

¹ These authors contributed equally to this work.

tempted to touch it. Humans seem to be rather good at making judgments about material classes (Sharan *et al.*, 2009; Wiebel *et al.*, 2013) or material properties (Fleming *et al.*, 2013). But what kind of information do the senses provide? How is each modality's information processed and stored, and how is information exchanged and integrated between senses to form a common percept?

Whereas a great deal of research has been conducted to investigate both uni- and bimodal shape perception, the perception of materials and material qualities has only recently received more attention in vision research (Buckingham *et al.*, 2009; Cant and Goodale, 2011; Fleming *et al.*, 2003; Ged *et al.*, 2010; Giesel and Gegenfurtner, 2010; Hiramatsu *et al.*, 2011; Kim and Anderson, 2010; Liu *et al.*, 2010; Motoyoshi, 2010; Motoyoshi *et al.*, 2007; Olkkonen and Brainard, 2010; Olkkonen *et al.*, 2010). In this study we want to shed light on the sensory interplay between the visual and the haptic spaces underlying material perception.

In a recent study investigating visuo-haptic shape perception, Gaissert *et al.* (2010) showed that the visual and haptic perceptual representations for three-dimensional shapes are highly similar. They used artificial shell-like stimuli, which they modulated systematically along three shape dimensions: distance between aperture and tip, number of convolutions, and symmetry. Participants' similarity judgments reflected the underlying parameter space both via visual and haptic exploration of the stimuli. This indicates a close link between visual and haptic shape processing. In further experiments, they strengthened these findings by showing that their results generalize to natural stimuli (shells) (Gaissert & Wallraven, 2012). Furthermore, they could show that these highly similar perceptual spaces give rise to very similar categorization behavior in both modalities (Gaissert *et al.*, 2011).

These results indicate that shapes are represented at a multimodal level in the brain. This is strengthened by the brain imaging literature. Several studies could show that there are brain regions which get activated both by visual and haptic exploration of objects (for a review, see Amedi *et al.*, 2005). Especially LOTv, a subregion of LOC, seems to respond in a very similar fashion to both visually and haptically presented shapes and might have bimodal geometric representations (Amedi *et al.*, 2002).

So far, no comparable convergence of visual and haptic representations has been described for materials and surfaces of objects. A study investigating bimodal perception of surface texture with fMRI failed to find clear evidence for a region where both visual and haptic material information might be encoded (Stilla and Sathian, 2008). There are no psychophysical studies that have compared the general dimensionality of both visual and haptic perception of objects' surfaces and materials in a systematic and comprehensive fashion. Most studies have focused either on a single modality or considered a lim-

ited range of material qualities. Here we set out to examine both the visual and haptic representations with a large and identical stimulus set and across a wide range of material qualities.

The majority of studies on bimodal surface perception have investigated roughness. For example, a comparison between the visual and the haptic sense was reported by Bergmann Tiest and Kappers (2007). They used a large set of different material samples characterized through different physical roughness measures (industrial roughness standards and spectral densities at different spatial scales). In separate haptic and visual conditions, participants were asked to rank the samples according to perceived roughness. Participants' ordering matched physical roughness rather well, and importantly, both haptic and visual orderings corresponded about equally to the physical measures.

A few studies have looked at the general dimensionality of the spaces underlying visual material (texture) perception in order to identify a generic texture space. Rao and Lohse (1996) asked participants to rate images from the Brodatz database (Brodatz, 1966) on several Likert scales. By means of multidimensional scaling (MDS) they found three underlying dimensions of the perceptual space: repetitiveness, contrast/directionality, and complexity/coarseness. In a subsequent study, Bhushan *et al.* (1997) sought to establish a representation of words describing textures. They confirmed the three-dimensional structure identified in their previous study, establishing a close correspondence between the texture word space and the visual texture space.

More studies have been conducted in the haptic than in the visual domain. Hollins *et al.* (1993) had their participants sort different haptically presented materials into groups according to similarity. Using MDS, they identified roughness and hardness as the dimensions underlying haptic judgments. In a further experiment, Hollins *et al.* (2000) identified slipperiness as a third dimension, present only in a subgroup of participants. Using a set of car seat fabrics to investigate haptic dimensions of material perception, Picard *et al.* (2003) found their data to be represented best by four dimensions, namely harshness, thickness, relief, and hardness.

A comprehensive account of studies on the haptic dimensionality of textures/materials was recently given by Okamoto *et al.* (2013). They summarized a variety of studies and tentatively suggest five dimensions of haptic texture perception: macro roughness, fine roughness, coldness, hardness, and friction. However, not all studies report the same dimensions, as these strongly depend on experimental and methodological factors. This illustrates the problem of comparing results from different studies using different methods and different stimuli. Accordingly, it is difficult to gain knowledge about commonalities between the visual and the haptic perceptual spaces for materials by comparing separate visual and haptic studies.

In this experiment, we seek to overcome these limitations of previous studies and determine the degree of correspondence between both senses in a systematic fashion using the same set of stimuli and the same procedure for both modalities. Importantly, we also decided to investigate a wide variety of material qualities in order to form a differentiated overview of visual and haptic material perception. We pose the question if the perceptual space that underlies material perception is the same for haptic and visual judgments. In addition, we want to gain insight into the aptitude of the two senses for the perception of various material qualities. The roles of the visual modality and the haptic modality in material perception have previously been characterized as both overlapping and complementary (e.g. Whitaker *et al.*, 2008). Therefore, we do not assume that there exists (like it has been suggested for shape perception (Gaissert *et al.*, 2010)) a bisensory representational level of both modalities for the perception of materials. The question we ask is rather: what is the degree of correspondence between both senses?

2. General Methods

2.1. Stimuli

Our stimuli consisted of 84 different material samples (14×14 cm in size) that were mounted onto pieces of 12 mm thick medium density fiberboard (MDF) of the same size by means of glue or double-sided adhesive tape. Flexible materials were only in the corners glued to the MDF in order to retain their flexibility.

We used stimuli from seven general material categories for the experiment: plastic, paper, fabric, fur and leather, stone, metal and wood. The materials spanned a wide range of different samples in order to represent the large variety of material appearances we encounter in everyday life. Examples of our stimuli can be seen in Fig. 1.

2.2. Material Properties

The material properties that we asked our participants to assess were chosen to represent a wide variety of material qualities as well as both visual and haptic features of materials. Participants were asked to rate each material property on a 7-point Likert scale. The following ten qualities were assessed, with the following definitions.

2.2.1. Glossiness

How glossy or matte does the material appear to you? Low values indicate the material is matte and shows no or very little reflections; high values indicate a shiny, reflective appearance.

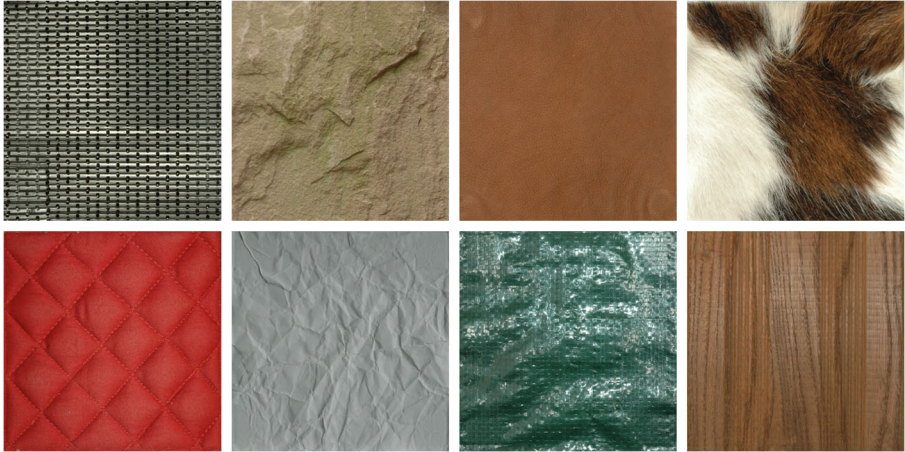


Figure 1. Example images taken from our seven material categories. Top row, from left to right: metal, stone, leather, fur. Bottom row, from left to right: fabric, paper, plastic, and wood. This figure is published in colour in the online version.

2.2.2. *Colorfulness*

How colorful or pale does the material appear to you? Low values indicate a greyish, monochrome appearance; high values indicate a colorful appearance, which could be either a saturated single color, or several colors.

2.2.3. *Roughness*

How rough or smooth does the material appear to you? Low values indicate that the surface feels smooth; high values indicate that it feels rough.

2.2.4. *Orderliness*

How ordered or chaotic does the material appear to you? Low values indicate that the material's surface shows no regularities but rather is random or chaotic. High values mean that the surface has an ordered, regular structure.

2.2.5. *Hardness*

How hard or soft does the material appear to you? How much force would be required to change the shape of the material? Low values indicate that the surface feels soft; little force is required to change the shape of the material. High values indicate that it feels hard and cannot easily be deformed.

2.2.6. *Warmth*

How warm or cold does the material appear to you? Low values indicate that the material feels warm or body temperature; high values indicate that the material feels cold to the touch.

2.2.7. *Elasticity*

How elastic or stiff does the material appear to you? Low values indicate that the material is not elastic. It is either stiff or its form remains changed after deformation. High values indicate the material is very elastic. After deformation, it will return to its original form.

2.2.8. *Friction*

How high does the friction of the material appear to you? Low values indicate that the material has low friction and is slippery. High values indicated that you feel a lot of friction when touching the material.

2.2.9. *Three-Dimensionality*

How three-dimensional does the material's surface appear to you? Low values indicate that the surface is flat, high values indicate that the surface has a three-dimensional structure.

2.2.10. *Texture*

How textured/patterned or homogeneous/uniform is the material's surface? Low values indicate that the surface is uniform, high values indicate that the material has a pattern or texture.

2.3. *Participants*

12 participants (all university students, mean age 27.4 years, eight female, four male) performed the rating task. 10 other participants (all university students or employees, mean age 24.2 years, six female, four male) participated in the categorization task. All participants had normal or corrected to normal visual acuity, and all but one were right-handed. They received financial compensation for their participation in the experiment. In the rating experiment, half of the participants started with the haptic condition of the rating, the other half started with the visual condition. In the categorization task, five participants performed the visual task and five the haptic task.

2.4. *Procedure*

Participants were seated in front of a table with a pedestal on which the material samples were mounted. At the beginning of each block the participant was given a written description of the material quality and the rating scale. The participants were encouraged to ask questions to ensure their understanding of the material property and rating scale. The order of blocks, i.e. material qualities, was randomized for each participant. The order of stimuli was randomized within one modality (visual or haptic) and within a single participant and remained the same in each block of one modality. We decided against presenting a stimulus and asking for all 10 material properties one after each other because we wanted our participants to concentrate on a single material property at a time, thus increasing the consistency of the judgments.

At the end of each trial, the participant would verbally give his/her response to the experimenters; this was entered into a Matlab GUI.

2.4.1. Haptic Condition

Participants were blindfolded and wore earphones over which we presented auditory broadband noise during the trials. The auditory noise served two purposes: first, to prevent the participant from hearing the sound of their hand exploring the material, and second, to signal the start and the end of the exploration. During each trial, the participant's task was to assess the stimulus for the perceptual quality of the current block. As soon as the correct material stimulus had been placed onto the pedestal by one of the experimenters, the other experimenter started the auditory noise, which signaled the beginning of the manual exploration to the participant. Participants were allowed to freely explore the materials with their right hand but we asked them to neither scratch nor knock onto the materials and not to explore the edges of the material samples. One of the experimenters was always nearby and made sure that participants would follow our instructions. After 13 s the mean frequency of the auditory broadband noise became higher, which signaled the participant to withdraw his/her hand from the stimulus, and stopped completely after another two seconds. The participant would verbally give a response, which was entered into a Matlab GUI by one of the experimenters.

2.4.2. Visual Condition

In this condition, the stimulus pedestal was placed in a box which was open in the front, in the back, and on top (from the participants' point of view). The front and the back of the box could be closed by means of curtains. At the beginning of each trial, a stimulus was placed on the pedestal within the box, and both curtains were closed. Then a gong signaled the beginning of a trial. After the gong, the participant was allowed to open the curtain in front of him/her and look at the stimulus. We paid attention that the stimuli were always at approximately the same distance from the participant. We also told the participant to sit still during the trial and minimize head movements. After 10 s, a higher pitched gong sounded, and the participant had to close the curtain and give his/her response verbally.

Data were collected during daylight conditions. Light entered the room through a window directly behind the participant.

2.5. Categorization Tasks

2.5.1. Categories

As mentioned before, we used stimuli from seven general material categories for the experiment: plastic, paper, fabric, fur and leather, stone, metal and wood. Here we examined how participants would perform in assigning the stimuli to these distinct material classes using only the haptic or only the visual

sense. We asked participants to choose one of these seven material categories for each stimulus. These categories reflect the everyday use of material category words. They are also based on physical and chemical properties and manufacturing processes.

2.5.2. Haptic Categorization of Materials

In the haptic categorization task, participants were presented with each stimulus in a randomized order. Participants had 13 s to explore the stimulus in the same manner as in the rating task. They were shown the list of stimulus categories beforehand and were asked to remember them. In each trial they had to verbally assign the stimulus to one of the given material classes.

2.5.3. Visual Categorization of Materials

In the visual categorization task, participants viewed each stimulus for 10 s. Again, stimuli were randomized. The exploration procedure was the same as for the rating task. In each trial they were asked to assign the stimulus to one of the material classes the participant had been told at the start of the categorization task.

3. Results

3.1. Visual and Haptic Material Property Ratings

3.1.1. Correlations Between Participants

We asked participants to rate our material stimuli on a variety of material properties in order to determine the relationship between the visual and the haptic sense in the perception of materials.

To assess the degree of correspondence between the two senses, we first analyzed to what extent participants gave similar ratings in the two tasks. We calculated correlation coefficients between participants across materials and material properties to gain a broad overview of the consistency between observers. Data for the visual and the haptic modality can be seen in Fig. 2. The range of correlations appears similar within each modality. In general, correlations were quite high, mostly between 0.6 and 0.75, indicating that participants do have similar concepts of the dimensions they had to assess. This indicates that property ratings do generalize to a large degree across observers.

3.1.2. Correlations Between Material Properties

Given that participants seemed to agree quite well on the assessment of the different material qualities, we wanted to look at how much the different qualities are related to each other across materials and samples. We were particularly interested in the correspondence between the visual and the haptic sense. Correlation matrices for both modalities are shown in Fig. 3.

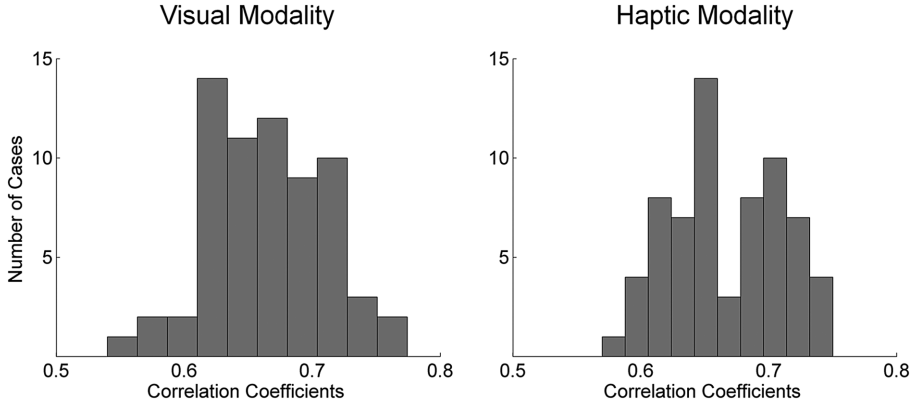


Figure 2. Histogram of the correlation coefficients between all 12 participants across all materials and dimensions tested. Data for the visual and the haptic modality are shown on the left and the right side, respectively. Each histogram consists of 66 correlation coefficients. For both modalities all of these correlations are highly significant ($p < 0.001$).

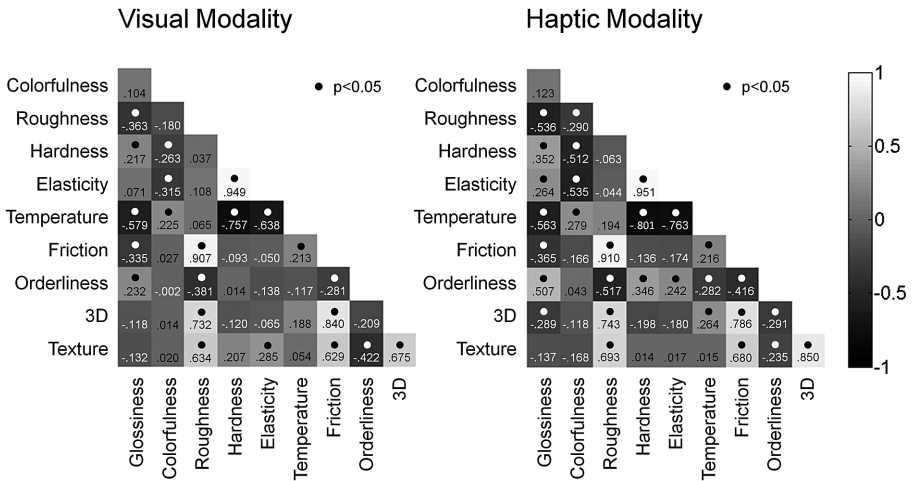


Figure 3. Correlation matrices between material properties across the different material classes and participants. Ratings on each property dimension were averaged over all 12 participants for each stimulus separately. The left side shows data for the visual modality, the right side shows data for the haptic modality. Significant correlations are indicated by a dot. White numbers indicate negative correlation coefficients, black numbers indicate positive correlation coefficients.

At first sight, the two patterns of results look very similar. Correlation coefficients ranged between $r = -0.76$ and $r = 0.95$ in the visual modality and between $r = -0.80$ and $r = 0.95$ in the haptic modality. Overall, the correlation patterns seem intuitively plausible. The highest correlations were observed between hardness and elasticity in both modalities. Both hardness and

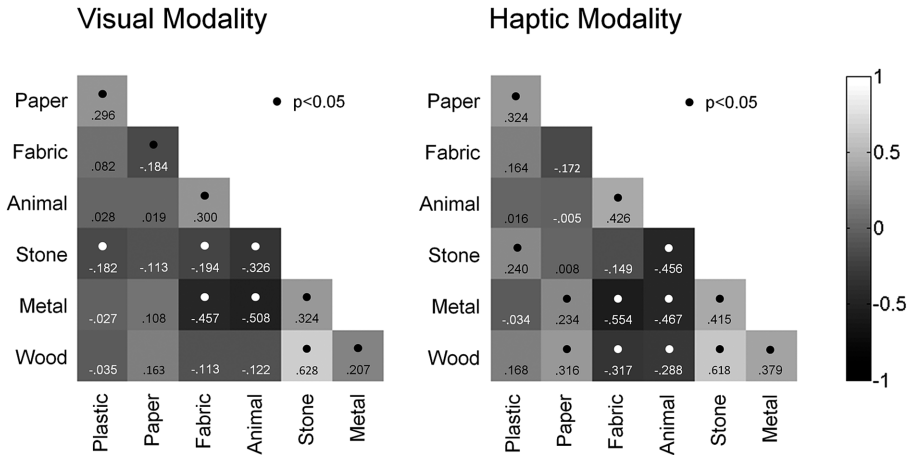


Figure 4. Correlation matrices for the different material classes, calculated across material properties and participants. Ratings for each stimulus on each property dimension were averaged across observers. The left side shows data for the visual modality, the right side shows data for the haptic modality. Significant correlations are indicated by a dot. White numbers indicate negative correlation coefficients, black numbers indicate positive correlation coefficients.

elasticity also correlated highly with temperature. Other very high correlations we found were between roughness, friction, texture and three-dimensionality.

3.1.3. Correlations Between Materials

We were also interested in how the different material categories were associated with each other (see Fig. 4). Again, results were highly intuitive in both modalities. Correlations ranged between $r = -0.51$ and $r = 0.63$ in the visual modality and between $r = -0.55$ and $r = 0.62$ in the haptic modality. The correlation matrices look rather similar in both modalities. Stone, metal and wood are in general negatively correlated with fabric and leather and fur, while being positively correlated with each other. Paper and plastic seem to be less correlated with other materials; they mainly correlate with each other. This could be due to the fact that both of these material categories consisted of rather heterogeneous samples.

3.1.4. Visual and Haptic Representations

We performed a principal component analysis over the mean property ratings (z-scored) for all material samples in both modalities (Fig. 5). This allowed us to see how the material samples were represented in a subjective material property space. The scree plots for both PCAs indicate that the data are represented best by two or three components (Fig. 6). For reasons of simplicity and clarity we plot only two dimensions for each modality in the following figures.

The first principal component explained 35.8% of the variance in the visual modality. In the haptic modality 41.7% of the variance was explained by the

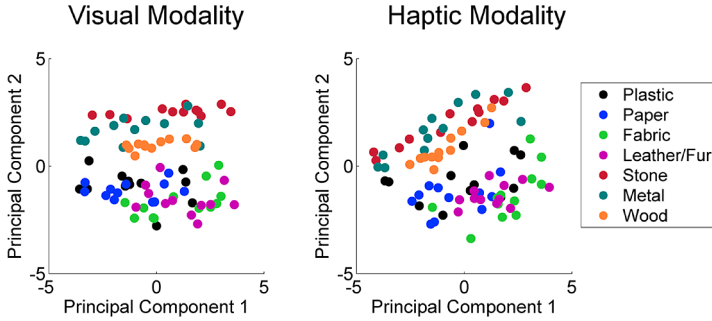


Figure 5. Representation of the different material classes based on the visual and haptic material property ratings within a two-dimensional PCA space. PCAs were performed based on the z-standardized property ratings for each stimulus averaged across participants.

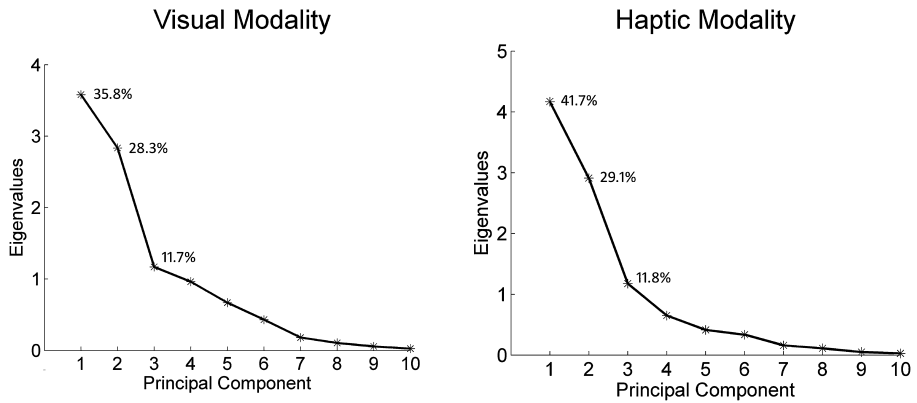


Figure 6. Scree plots of the PCAs we performed on the visual and haptic material ratings.

first factor. By looking at the factor loadings (Table 1), we can see that the first principal component is best described by the properties roughness, friction and three-dimensionality. Interestingly, this is true for both modalities, indicating a high degree of correspondence between the first principal components in the two modalities. Furthermore, in both modalities, there is no clear division of material classes along this axis. This can be attributed to the fact that we tried to cover a large variety of surface appearances in our material collection. Therefore, roughness seems to be an important dimension for organizing material surfaces but is relatively independent of material classes.

The second principal component explained 28.3% and 29.1% of the variance for the visual and the haptic modality, respectively. It separates quite nicely between hard and soft materials. In both modalities stone, metal and wood (in this order) form a cluster opposed to plastic, paper, leather and fur, and fabric. Accordingly, for the second principal component, highest factor

Table 1.

Factor loadings for the first three principal components revealed by the two PCAs in the visual and in the haptic modality

	Visual modality			Haptic modality		
	PCA 1	PCA 2	PCA 3	PCA 1	PCA 2	PCA 3
Glossiness	−0.245	0.174	0.672	−0.317	0.064	0.568
Colorfulness	−0.027	−0.239	0.499	0.006	−0.405	0.426
Roughness	0.476	0.123	0.006	0.407	0.274	0.075
Hardness	−0.079	0.561	0.056	−0.269	0.465	0.073
Elasticity	−0.021	0.548	0.195	−0.252	0.461	0.167
Temperature	0.186	−0.486	0.232	0.299	−0.375	0.122
Friction	0.491	0.026	0.140	0.403	0.227	0.131
Orderliness	−0.256	−0.063	0.236	−0.304	0.013	0.404
3D	0.450	0.030	0.325	0.393	0.205	0.331
Texture	0.404	0.202	−0.152	0.316	0.305	0.390

loadings were found for hardness and elasticity. In the visual modality temperature also showed a particularly high factor loading on this component, while in the haptic modality colorfulness loaded on this factor.

The third principal component also showed a high correspondence between the two senses with highest factor loadings for glossiness and colorfulness. It accounted for 11.8% of the explained variance in the haptic modality and for 11.7% in the visual modality. Overall, the picture in both modalities seemed very similar, leading to the conclusion that the material samples used here can be represented in congruent material property spaces.

3.1.5. Comparison of the Visual and Haptic Material Representations

In a next step we compared the visual and haptic representations of our material samples in a more direct manner. We correlated the property ratings for each material sample between the visual and the haptic modalities, see Fig. 7. All correlations were highly significant, ranging between $r = 0.40$ (texture) and $r = 0.83$ (hardness). Lowest correlations were found for texture and colorfulness, probably because those properties are better accessible to the visual modality, while in the haptic condition participants had to rely mainly on (comparatively unreliable) learned associations. The highest correlations were found for hardness and elasticity. Even though these seem to be primarily haptic qualities, the associations between certain material categories and their inherent properties of being rather hard or soft must have been as efficient in the visual modality as the haptic assessment. In sum, the overall high correlations between material quality ratings in the two modalities confirm the notion of a very tight correspondence between the visual and the haptic material space.

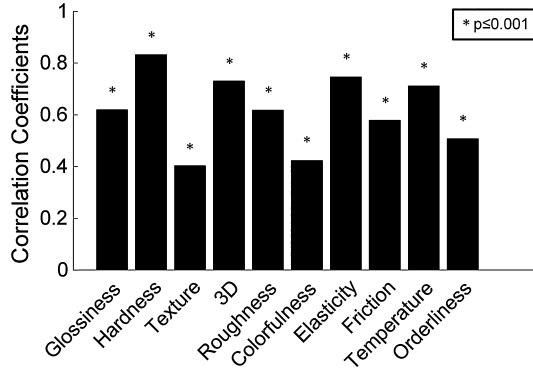


Figure 7. Correlations between the property ratings of each material sample in the visual and the haptic condition for each material quality. Correlations for each material property were calculated on the ratings given to each stimulus by each participant in both modalities. Significant correlations are indicated by an asterisk ($p \leq 0.001$).

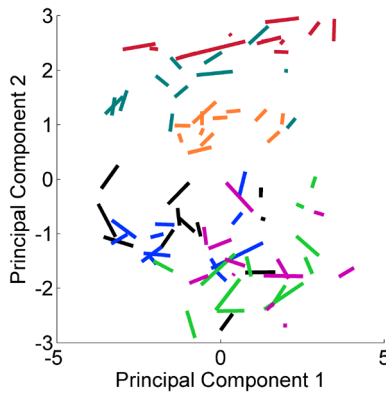


Figure 8. Procrustes analysis between the visual and the remapped haptic data. Lines indicate the distance between the stimulus locations in the visual space and the remapped haptic space.

We performed procrustes analyses on the visual and haptic spaces revealed by PCA to directly compare the two sensory representations. A procrustes analysis aims at mapping two representations onto each other as closely as possible by using linear transformations only. In Fig. 8 the best solution for our two datasets (visual and haptic) is shown on a per-stimulus basis. Figure 9 depicts the procrustes solution for the centers of the seven material categories.

The procrustes analyses show very similar representations of the visual and haptic data in the two dimensional PCA space. In addition, these illustrations again show that in both modalities similar material classes group together. The three hard materials wood, stone and metal form a cluster, while the soft ones such as fabric, leather and fur, paper, and plastic form another group.

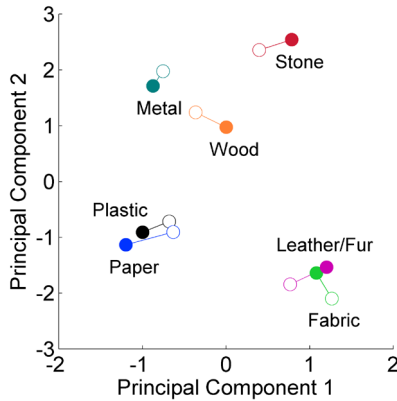


Figure 9. Procrustes analysis between the material category cluster centers of the PCA solutions for the visual (filled symbols) and the remapped haptic (open symbols) data. This figure is published in colour in the online version.

3.1.6. Order Effects

The high agreement between the visual and haptic assessment of our material samples brings up the question whether some of the correspondences can be explained by memory effects. After all, the same set of participants completed both conditions. Half of the participants started with the visual condition, while the other half of the participants started with the haptic condition. For example, participants who started visually might have been able to rely on memory strategies during the haptic assessment of the different material samples. To control for this issue, we split our dataset according to the order in which participants performed our rating. We then mapped data from the two halves of the rating procedure onto each other, comparing half a haptic dataset with the respective visual dataset (Fig. 10).

In the first procrustes analysis we mapped the haptic data from those six participants who started with the haptic ratings onto data from those who started with the visual ratings, that is, we compared visual and haptic data from the first halves of the rating procedure with each other. In the second procrustes analysis we compared visual and haptic data from the second halves of the rating procedure with each other by mapping the haptic data from those six participants who had started visually onto the visual data from those participants who had started haptically. If the order in which participants gave their ratings had played a role, we would expect larger differences in the first procrustes analysis than in the second one.

There is a nearly perfect overlap of the two representations in both analyses, indicating that the order of tasks did not affect our results. However, to investigate whether there were order effects for single material properties, we additionally examined the mean inter-participant correlations for each prop-

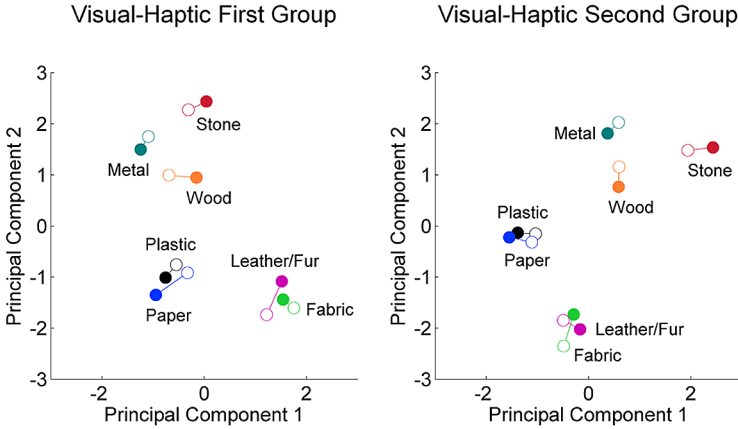


Figure 10. Procrustes analyses of the PCAs derived from all participants’ data from the first half (six participants’ visual data and six participants’ haptic data) and the second half of the rating procedure (six participants’ haptic data [open symbols] and six participants’ visual data [filled symbols]). This figure is published in colour in the online version.

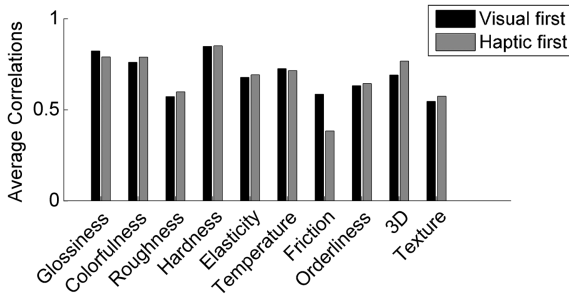


Figure 11. Mean correlation coefficients for each property rated visually. Data are aggregated across material samples and split in two groups: Participants that started with the visual modality (black) and participants that started with the haptic modality (grey).

erty for participants that started with the haptic modality and participants that started with the visual modality. It could be that for ‘unimodal’ properties such as color or temperature a learning effect is observed over the course of the rating sessions. Mean inter-participant correlation coefficients were calculated for haptic ratings and visual ratings, see Figs 11 and 12. We inspected the difference between the mean correlations for each property between the two modalities and between the two groups of participants. Mean values were similar for both types of ratings except for colorfulness where mean inter-participant correlations for haptic color ratings were substantially smaller than for visual color ratings. This only reflects the higher difficulty of the haptic colorfulness ratings compared to the visual ones. To test for memory effects the crucial comparison would be the one between the two groups of partici-

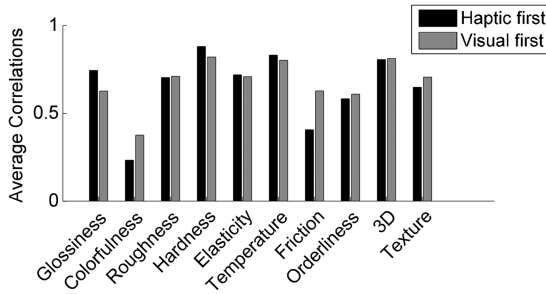


Figure 12. Mean correlation coefficients for each property rated haptically. Data are aggregated across material samples and split into two groups: Participants that started with the haptic modality (black) and participants that started with the visual modality (grey).

pants (visual first and haptic first) within one modality. In general, differences between the groups were rather small, supporting the notion that memory effects over the course of the rating cannot account for the high overlap of visual and haptic ratings. For colorfulness, a small effect of the experimental order could be observed in the haptic ratings. In this case, participants might have benefitted from having visually assessed the stimuli before doing the haptic ratings. For friction, visual inter-participant correlations were higher when participants started visually whereas haptic correlations were higher when participants started haptically. Here, having already completed the task in another modality seemed to lower observers' consistency.

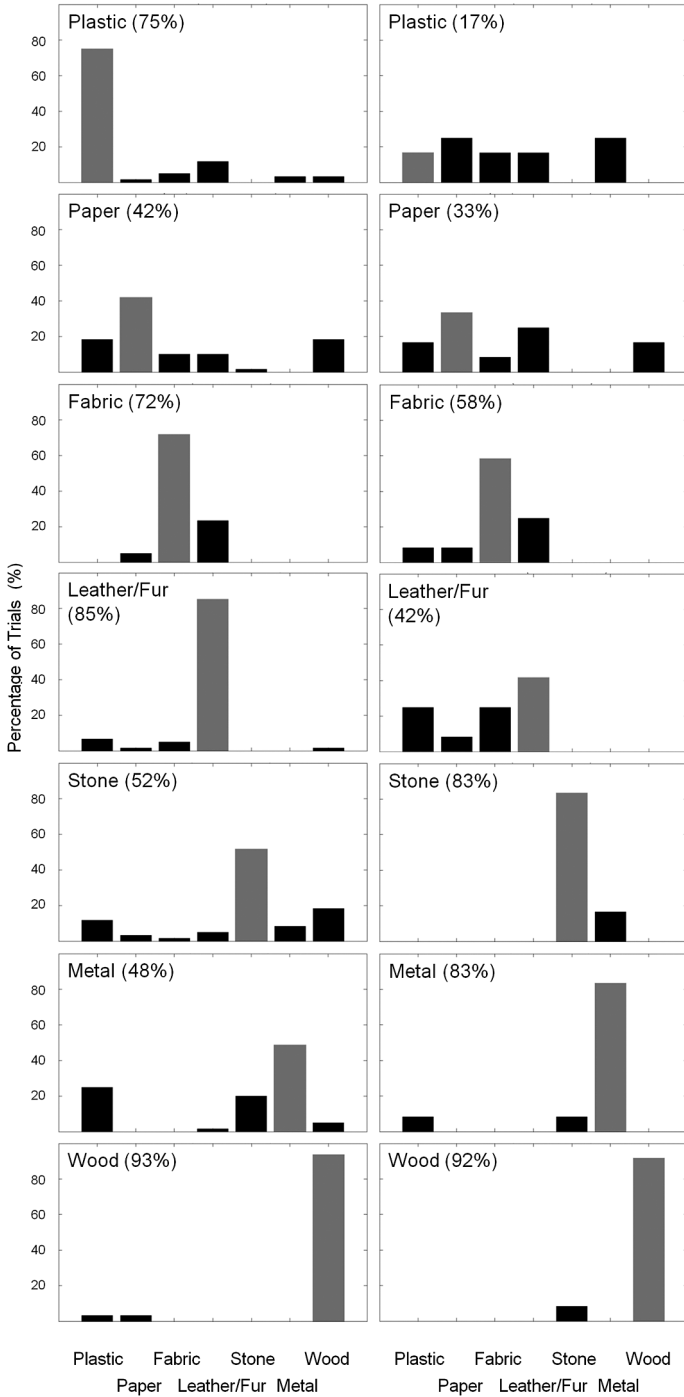
3.2. Categorization

Seeing that material categories grouped nicely in the PCA space, we wanted to assess participants' performance in assigning the stimuli to these categories. In addition to the rating task, we therefore asked participants to categorize or identify our 84 material stimuli based on either visual or haptic information alone.

3.2.1. Haptic Categorization

On average, 66% of the 84 stimuli were consistently assigned to their material class. We found substantial differences between material classes; see Fig. 13 (left column). Lowest consistencies were found for metal, stone and paper. In particular, metal stimuli were very often assigned to the plastic category. Highest consistency was achieved in the categorization of wood, and leather

Figure 13. Categorization and classification results for the haptic modality: the plots in the left column show the categorization data collapsed across five observers as well as the mean categorization accuracy for each material class separately. The plots in the right column show the classifications of the machine learning algorithm based on participants' property ratings of our stimuli.



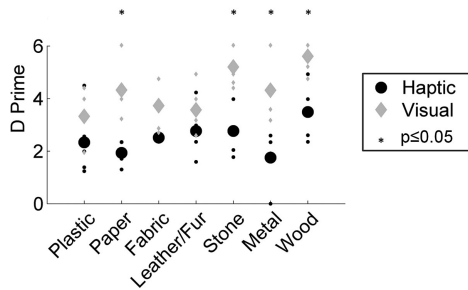


Figure 14. Sensitivity measures (d') for visual and haptic categorization. Small symbols represent a single participant's d' , large symbols represent the mean d' of five participants. Significant differences between the visual and the haptic condition are indicated by an asterisk ($p \leq 0.05$, Bonferroni corrected).

and fur. Materials which had similar surface properties were confused most often. Metal, paper and plastic often have smooth uniform surfaces, while fabric for instance showed much more distinctive surface properties and was mainly confused with leather and fur.

3.2.2. Comparison with Visual Categorization

On average, participants categorized 90% of the stimuli into their predefined categories. Thus, in contrast to the haptic categorization task, participants showed higher agreement when visually assigning each stimulus to its material category.

In order to compare the visual and the haptic performances, we calculated participants' sensitivity indices (d') in both conditions (Fig. 14). For paper, stone, metal and wood, sensitivity measures significantly differed between participants who completed the task haptically and those who completed it visually (two sample t -test, $p \leq 0.05$, Bonferroni corrected).

3.2.3. Classification of Material Categories

We compared the visual and haptic material categorization behavior of our five participants to the classification results a linear classifier would achieve based on the material property ratings in each modality. As a limitation of the following results, it should be mentioned that the performance of a classifier partly depends on the chosen classification algorithm. Therefore these results must be interpreted with some caution. We applied a naïve Bayesian classifier to our 84 material samples with a leave-one-out cross-validation. This was implemented in Matlab using the 'classify'-function (The MathWorks Inc., 2007, Natick, MA, USA). The function essentially performs a linear discriminant analysis with a diagonal covariance matrix. In leave-one-out cross-validation, the classifier is run several times, each time a different stimulus is withheld from the training set. After training, the classifier is tested on that stimulus.

Specifically, the classifier was run 84 times. Each single observation (i.e. each stimulus) was used once as the test sample, while all other observations were used as training samples. As features, the classifier was given the ten different material property ratings (averaged across participants) for each stimulus. This was done for each modality separately. In the visual modality, the classifier achieved an accuracy of 71% (58.3% plastic; 58.3% paper; 50% fabric; 58.3% leather and fur; 83.3% stone; 91.7% metal; 100% wood). In the haptic modality, the classifier achieved an accuracy of 58% (16.7% plastic; 33.3% paper; 58.3% fabric; 41.7% leather and fur; 83.3% stone; 83.3% metal; 91.7% wood) (see Fig. 13, right column). Given that we performed a 7-way classification (i.e. chance performance would be at 14.3%), classification performance is relatively high. The difference in categorization performance between the two modalities which we observed in the categorization results is captured to a certain extent in the outcome of the classification analysis. The visual ratings seem to contain somewhat more information about material class membership than the haptic ratings. However, the pattern of category-wise misclassifications does not entirely match the pattern of our participants' categorizations. This is particularly evident for stone and metal stimuli which usually were correctly labeled by our classifier, whereas our participants' performance for those stimuli was among the most inconsistent of the different material categories. Plastic stimuli, on the other hand, were categorized rather consistently by our observers whereas the classifier seemed to have more difficulties with this rather diverse category.

3.2.4. Classification of Participants' Categorizations

To examine our participants' haptic categorization more closely, we also applied a classifier to participants' categorization data. Specifically, we again applied a classifier to the property ratings, this time trying to predict each individual participant's categorizations instead of the actual categories. (We excluded one participant (of five) from this analysis because he/she categorized no stimulus as metal.) The classifier's mean (averaged across participants) accuracy was 57.4% (SD = 7.7%). In order to get an idea of which properties are informative about category membership and may be used by participants to form a decision in the categorization task, we looked at the classifier weights. These indicate how much each property contributes to the boundary equation, i.e. the weight the classifier gave each material property when forming its decision. For both the classifier predicting the actual categories as well as the classifier predicting participants' categorizations, the highest weights were found for elasticity, hardness, and temperature (see Fig. 15). These three material properties are also those that project highest on the second principal component. As mentioned before, while the first principal component only shows within-category variance, the second component shows inter-category

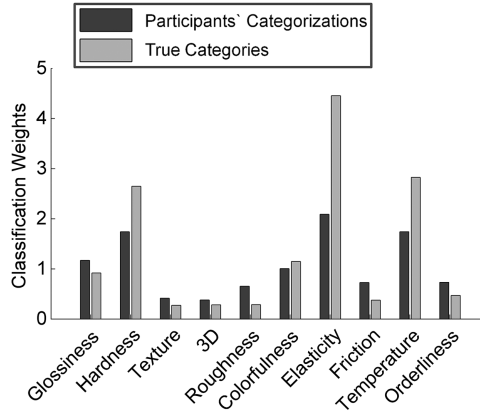


Figure 15. Feature weights for the classification of four participants' categorizations (averaged across participants) and 'true' categories.

variance. Thus, this component contains information about material categories and so it would make sense for participants to make use of it when categorizing materials.

4. Discussion

We intended to compare visual and haptic judgments of material qualities. Our results show that participants are able to reliably retrieve similar information about various material qualities from a wide variety of stimulus samples — both within the visual and the haptic modality. This is impressive since the information that is available to these two senses differs substantially. It indicates that vision and haptics resort to congruent perceptual representations. Participants were highly consistent in their judgments, which demonstrates that they relied to a large degree on very similar underlying representations of material qualities.

4.1. Order Effects

The fact that we found virtually no effects of the order of conditions (visual and haptic) indicates that the correspondence we observed for visual and haptic exploration of the material stimuli is not the result of mere learning or familiarity with the particular stimulus set. Participants who started with the visual ratings did not achieve the highly similar haptic ratings by learning all of our stimuli by heart in the first condition and then identifying (and picturing) the stimuli in the second (i.e. haptic) condition. This could be a good strategy, especially for 'unimodal' properties (e.g. colorfulness) and it is likely participants used this strategy in a few trials. However, the procrustes analyses we performed on the unimodal property spaces between participants starting with the visual ratings and participants starting with the haptic ratings show

that participants' representations are highly similar, irrespective of the order of conditions. To look more closely into this, we computed the mean inter-participant correlations for our property ratings separately for participants who completed those ratings in the first and the second half of the experiment (see Figs 11 and 12). If learning during the course of our experiment had played a role, we would expect the inter-participant correlations of participants doing the haptic ratings in the second half of the sessions to be higher for primarily visual properties because they have previously seen all stimuli. For visual ratings we would expect the same effect for primarily haptic properties. However, we did not find a clear relationship in the inter-participant correlations between properties and rating order in our data. We conclude that memory effects over the course of the rating sessions do not seem to play a dominant role in our data and their interpretation.

However, we assume that long-term learned associations played an important role in how participants solved the task. For some judgments, participants were even forced to rely on them. For example, for assessing the temperature of our stimuli visually, participants might have relied on either their representation of the stimulus in memory, or heuristics. Such a heuristic could be that they rated the quality based on the probability of the co-occurrence of two different qualities (e.g. "this feels smooth so it is more likely to be glossy") or the probability of this quality for the material category they perceived (e.g.: "This looks like stone so it is likely to be colder than most of the other stimuli."). These associations might stem from actual knowledge about physical properties of materials (e.g. metal has a better thermal conductivity than fabric or paper, therefore it is likely to be colder to the touch), but it is more likely to be based on learning over the whole lifespan. Most of the times we have touched metal or stone, we have experienced them to feel cool. Recent studies have suggested that humans often tend to make use of heuristics when perceiving our surroundings rather than knowledge about actual physical properties of objects (Fleming, 2012).

The strength of these associations is somewhat reflected in our inter-participant correlations. These were lower for haptic colorfulness ratings compared to visual colorfulness ratings. Although we have learned an association between materials and their probability of colorfulness, we do not have a perfect proxy for inferring colorfulness haptically. Some of these proxies can be quite precise, though. Visual temperature inter-participant correlations, for example, are only slightly lower than haptic ones. We obviously have rather accurate visual knowledge about how warm or cool a material will feel.

In line with this assumption, Fleming *et al.* (2013) were able to show that associations about different materials are rather consistent across participants, even when they have to rely solely on memory and stored knowledge. In ad-

dition, they showed that this semantic representation closely resembles the representation of visually presented material images.

4.2. PCA Dimensions

The PCAs we conducted on the ratings yielded very similar solutions for both modalities and revealed the same material clusters. One dimension distinguished between categories, and a second dimension showed substantial within-category variance. The dimensions of the two-dimensional PCA solutions might be interpreted as ‘hardness/softness’ and ‘roughness/smoothness’, respectively. Whereas hardness is a material quality in the proper sense and therefore mostly similar across a category, roughness is more of a surface feature that can strongly be influenced by the manufacturing process (e.g. stone can be natural [i.e. rough] or polished [i.e. smooth]).

The factor we have now called ‘roughness’ has the highest loadings from friction, roughness and three-dimensionality both in the visual and the haptic PCA. The properties that load highest on the second factor (‘hardness’) are the material qualities hardness and elasticity. The third highest load is temperature for the visual and colorfulness for the haptic PCA. This seems curious since these two properties are not directly accessible in the respective modality. We cannot sense temperature visually; neither can we sense colorfulness haptically. Thus, ratings along these dimensions must have been highly influenced by learned associations with hardness and elasticity. Hard materials like metal and stone are often cold to the touch and less colorful than soft materials like fabric. The fact that these properties received such high factor loadings might indicate that observers used very similar and simple heuristics to assess these properties, like coupling the evaluation of that quality to another quality. This is reflected in their contribution to the second largest principal component.

It should be noted that the substantial within-category variance of one of the components means that participants did not base their judgments solely on material categories. Participants could have used this as a heuristic to solve the task quickly and easily (e.g. this is paper so it is likely to be smooth). However, even in the visual condition, where categorization performance is very consistent, participants differentiated between individual samples of one category. Also note that the factor loadings are very consistent between the two modalities. That seems striking because the two dimensions (roughness and hardness) are generally considered very prominent in haptic perception, and previous haptic studies on the space underlying material perception have rather consistently found these two dimensions (Okamoto *et al.*, 2013). Does this mean that the haptic modality is in some sense dominant over the visual one? Of course, PCA solutions on data that are collected using the semantic differential method do rather heavily depend on both stimuli and rating scales. We did, however,

take care in using a large collection of stimuli and a variety of both visual and haptic material properties. The high correlations between visual and haptic assessments of properties that load on the factors roughness, friction, hardness and elasticity show that both hardness and roughness can be assessed very reliably with both senses (Bergmann Tiest and Kappers, 2007). Importantly, these two properties provide important information about how to interact with stimuli in our environment. Since we used real material samples (as opposed to photographs or Brodatz textures in previous visual studies about surface representation), environmental meaning and thereby haptic features of the stimuli were emphasized in our setting. As a matter of fact, in the visual condition, participants sometimes reported that they felt compelled to touch the materials. This affordance of material stimuli strengthens the importance of the haptic sense in material perception. Therefore, we do not think of the haptic sense as the dominant one in material perception but rather that it provides very important information in the everyday interaction with materials.

4.3. Vision and Haptics in Material Perception: Redundant and Complementary

As mentioned earlier, a very close correspondence between both modalities was established by studies examining object shape perception. However, the relationship between these two modalities remained unclear for the perception of objects' surfaces and materials. Sensory mechanisms put constraints on our perception of the world. Whereas vision's parallel nature allows it to process objects fast and globally, haptic processing is sequential, slow, relies on memory components and only acts in a very confined region of space. This implies consequences for the processing of materials. The answer to the question of the relationship between the visual and the haptic sense in material perception strongly depends on the examined material qualities. Some material qualities are best accessible to vision: color, for example, cannot possibly be felt. For others, however, the haptic modality has an advantage: softness, for example, is not directly accessible to the visual modality. In order to judge softness visually, one has to rely on learned associations and knowledge. Other qualities, like roughness, are very salient for the haptic sense but there are also image statistics that correlate with roughness and allow us to infer a surface's roughness by vision quite well (Bergmann Tiest and Kappers, 2007, but see Ho *et al.*, 2006). In a review, Whitaker *et al.* (2008) characterized the relationship between the visual and haptic modality in surface perception as "independent, but complementary". The visual sense seems to be more suited for the perception of macrogeometric object properties, while microgeometric properties can be better assessed by the haptic sense.

Therefore, for material perception, the haptic modality might play a more important role than for shape perception. While the visual system clearly dom-

inates in bimodal integration tasks for shape perception, this seems not to be the case for material perception. For example, studies with parametrically varied textured objects have shown that letting participants explore stimuli haptically increases the salience of surface features as opposed to shape features (Cooke *et al.*, 2007). Lederman and Abbott (1981) could show that introducing a visuo-haptic discrepancy about the roughness of a stimulus led to an equal weighting of information from both senses.

In our PCA data and correlation matrices, we found the visual and the haptic assessments to be largely congruent. However, we also observed some revealing differences. For example, judgments on colorfulness and texture (clearly visual properties) correlate to a much smaller extent than, for example, three-dimensionality. As mentioned before, while visual and haptic judgments of materials are to a large extent redundant, some qualities can be assessed best by only one sense. The notion that the visual and haptic senses on the one hand provide redundant information for many tasks but also complement each other sometimes seems evident in our daily behavior: we do not necessarily need to touch every material to gain information about it. When we see a wooden board we can identify it as such with ease. However, when we are shopping for furniture and want to distinguish between a shelf made from real wood and one from imitated wood, we probably will have to touch it in order to make this fine distinction. Thus, to make fine-grained distinctions and resolve ambiguities, it is helpful to use both senses.

To answer the question what roles the two senses play in their collaborative assessment of materials, we can also resort to the results of the categorization task. The haptic sense may be apt to make very fine judgments about some specific material properties but categorization of materials seems to be less consistent with haptic information alone than with visual information alone. This indicates that vision allows a better integration of material features. Fewer features might be available in haptic exploration, which can lead to confusions between similar categories (e.g. plastic and metal). However, our PCA results show that the basic representation of features is quite similar even when haptics and vision are tested alone.

One could consider this as a contradiction within our data: while participants' haptic assessment of material qualities is quite similar to their visual assessment, haptic categorization performance is less consistent. However, at closer inspection, this is not unexpected. After all, in real life, we rarely categorize materials based on haptic information alone. Even though participants are well able to retrieve all these individual material properties when forced to, they either do not use this information for categorization purposes or this information is alone not quite rich enough to support highly consistent categorization performance. This is, to a certain degree, reflected in the performance of the classifier we applied to the rating data. The seven-way classifier ap-

plied to the visual rating data predicted the material categories of our samples with a higher accuracy than the classifier applied to the haptic property ratings. The visual ratings, therefore, seem to contain more information about the material categories our samples belong to than the haptic ratings. However, as already mentioned in the results section, the pattern of misclassifications differed between the classifier and human participants — they demonstrated different biases. Whereas human participants did not consistently recognize paper, stone and metal in the haptic categorization task, the haptic classifier performed worst for plastic, paper, fabric, and leather and fur. We assume that a crucial factor for this is the within-category variance. Humans can deal with such variance because they know about various appearances of material classes (e.g. they know that fabric can range from silk to carpet, or that plastic can range from bubble wrap to vulcanized rubber), the classifier, however, can only deal with rather homogeneous categories, like stone, metal, or wood, where the test sample is likely to be similar to the training samples.

Another important aspect is that haptic categorization performance was most likely somewhat degraded compared to natural conditions because we masked all auditory information that would have been helpful in identifying the materials. When assessing materials, it is rather common to scratch or knock onto their surfaces. Even when we do nothing but stroke the surface laterally, we very often hear sounds that might be informative about the material. It is very likely that this information is inextricably linked to haptic exploration and very helpful for the perception and identification of materials in everyday live.

4.4. Conclusion

We have found a very close correspondence between the representations of material qualities in the visual and the haptic sense. Participants are surprisingly consistent in judging a variety of material qualities unimodally, even for qualities that are not typically associated with the respective modality. The analysis of the inferred visual and haptic representations suggests that these are closely linked.

Acknowledgements

This work was supported by grant DFG 879/9 to K.R.G. We thank Fiona Gericke for help with data collection.

References

- Amedi, A., Jacobson, G., Hendler, T., Malach, R. and Zohary, E. (2002). Convergence of visual and tactile shape processing in the human lateral occipital complex, *Cereb. Cortex* **12**, 1202–1212.

- Amedi, A., von Kriegstein, K., van Atteveldt, N. M., Beauchamp, M. S. and Naumer, M. J. (2005). Functional imaging of human crossmodal identification and object recognition, *Exp. Brain Res.* **166**, 559–571.
- Bergmann Tiest, W. M. and Kappers, A. M. L. (2007). Haptic and visual perception of roughness, *Acta Psychol.* **124**, 177–189.
- Bhushan, N., Rao, A. R. and Lohse, G. L. (1997). The texture lexicon: Understanding the categorization of visual texture terms and their relationship to texture images, *Cognit. Sci.* **21**, 219–246.
- Brodatz, P. (1966). *Textures*. Dover, New York, NY, USA.
- Buckingham, G., Cant, J. S. and Goodale, M. A. (2009). Living in a material world: How visual cues to material properties affect the way that we lift objects and perceive their weight, *J. Neurophysiol.* **102**, 3111–3118.
- Cant, J. S. and Goodale, M. A. (2011). Scratching beneath the surface: New insights into the functional properties of the lateral occipital area and parahippocampal place area, *J. Neurosci.* **31**, 8248–8258.
- Cooke, T., Jäkel, F., Wallraven, C. and Bulthoff, H. H. (2007). Multimodal similarity and categorization of novel, three-dimensional objects, *Neuropsychologia* **45**, 484–495.
- Fleming, R. W. (2012). Human perception: Visual heuristics in the perception of glossiness, *Curr. Biol.* **22**, 865–866.
- Fleming, R. W., Dror, R. O. and Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties, *J. Vision* **3**(5), 3.
- Fleming, R. W., Wiebel, C. B. and Gegenfurtner, K. R. (2013). Perceptual qualities and material classes, *J. Vision* **13**(8), 9.
- Gaissert, N., Bühlhoff, H. H. and Wallraven, C. (2011). Similarity and categorization: From vision to touch, *Acta Psychol.* **138**, 219–230.
- Gaissert, N. and Wallraven, C. (2012). Categorizing natural objects: A comparison of the visual and the haptic modalities, *Exp. Brain Res.* **216**, 123–134.
- Gaissert, N., Wallraven, C. and Bühlhoff, H. H. (2010). Visual and haptic perceptual spaces show high similarity in humans, *J. Vision* **10**(11), 2.
- Ged, G., Obein, G., Silvestri, Z., Le Rohellec, J. and Viénot, F. (2010). Recognizing real materials from their glossy appearance, *J. Vision* **10**(9), 18.
- Giesel, M. and Gegenfurtner, K. R. (2010). Color appearance of real objects varying in material, hue and shape, *J. Vision* **10**(9), 10.
- Hiramatsu, C., Goda, N. and Komatsu, H. (2011). Transformation from image-based to perceptual representation of materials along the human ventral visual pathway, *Neuroimage* **57**, 482–494.
- Ho, Y.-X., Landy, M. S. and Maloney, L. T. (2006). How direction of illumination affects visually perceived surface roughness, *J. Vision* **6**(5), 9.
- Hollins, M., Bensmaïa, S., Karlof, K. and Young, F. (2000). Individual differences in perceptual space for tactile textures: evidence from multidimensional scaling, *Percept. Psychophys.* **62**, 1534–1544.
- Hollins, M., Faldowski, R., Rao, S. and Young, F. (1993). Perceptual dimensions of tactile surface texture: a multidimensional scaling analysis, *Percept. Psychophys.* **54**, 697–705.
- Kim, J. and Anderson, B. L. (2010). Image statistics and the perception of surface gloss and lightness, *J. Vision* **10**(9), 3.

- Lederman, S. J. and Abbott, S. G. (1981). Texture perception: studies of intersensory organization using a discrepancy paradigm and visual versus tactual psychophysics, *J. Exp. Psychol.* **7**, 902–915.
- Liu, C., Sharan, L., Adelson, E. H. and Rosenholtz, R. (2010). Exploring features in a Bayesian framework for material recognition, *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference*, San Francisco, CA, pp. 239–246.
- Motoyoshi, I. (2010). Highlight-shading relationship as a cue for the perception of translucent and transparent materials, *J. Vision* **10**, 6.
- Motoyoshi, I., Nishida, S., Sharan, L. and Adelson, E. H. (2007). Image statistics and the perception of surface qualities, *Nature* **447**, 206–209.
- Okamoto, S., Nagano, H. and Yamada, Y. (2013). Psychophysical dimensions of tactile perception of textures, *J. IEEE Trans. Haptics* **6**, 81–93.
- Olkkonen, M. and Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination, *J. Vision* **10**, 5.
- Olkkonen, M., Witzel, C., Hansen, T. and Gegenfurtner, K. R. (2010). Categorical color constancy for real surfaces, *J. Vision* **10**(9), 16.
- Picard, D., Dacremont, C., Valentin, D. and Giboreau, A. (2003). Perceptual dimensions of tactile textures, *Acta Psychol.* **114**, 165–184.
- Rao, A. R. and Lohse, G. L. (1996). Towards a texture naming system: Identifying relevant dimensions of texture, *Vision Res.* **36**, 1649–1669.
- Sharan, L., Rosenholtz, R. and Adelson, E. H. (2009). What can you see in a brief glance? *J. Vision* **9**, 784.
- Stilla, R. and Sathian, K. (2008). Selective visuo-haptic processing of shape and texture, *Hum. Brain Mapp.* **29**, 1123–1138.
- Whitaker, T. A., Simões-Franklin, C. and Newell, F. N. (2008). Vision and touch: Independent or integrated systems for the perception of texture? *Brain Res.* **1242**, 59–72.
- Wiebel, C. B., Valsecchi, M. and Gegenfurtner, K. R. (2013). The speed and accuracy of material recognition in natural images, *Atten. Percept. Psychophys.* **75**, 954–966.