The neural basis of visual material properties in the human brain

by

Hua-Chun Sun

Dissertation submitted for the degree of

Doctor of Philosophy

School of Psychology

University of Birmingham

United Kingdom

September 2015

Abstract

The central issue of the dissertation is to investigate the neural basis of visual glossiness processing in the human brain. It contains three independent studies with human functional magnetic resonance imaging (fMRI) measurements. The research background and related articles were introduced in Chapter 1. The first study is to localize brain areas preferentially responding to glossy objects defined by specular reflectance. We found activations related to gloss in the posterior fusiform (pFs) and in area V3B/KO and the two areas process gloss information differently (Chapter 2). The second study is to investigate how the visual-induced haptic sensation is achieved in our brain. We found that in secondary somatosensory area (S2) was distinguishable between glossy and rough surfaces, suggesting that visual information about object surfaces may be transformed into tactile information in S2 (Chapter 2). In the third study we investigate how the brain processes surface gloss information conveyed by disparity of specular reflections on stereo mirror objects and compared it with the processing of specular reflectance. We found that both dorsal and ventral areas were involving in this processing. We also found that in ventral part of left V3B/KO the pattern activations that were distinguishable between different specular reflectance can be transferred to distinguish different glossiness conveyed by binocular gloss information but not vice versa. This result implicates that in this region the processing of stereoscopic gloss information has a pattern of activation that is additional to the representation of specular reflectance. In Chapter 5 the main findings were summarized and compared across the three studies. Overall, the three studies contribute to our understanding about the neural basis of visual glossiness and material processing in the human brain.

Included papers and contribution

Journal articles

1. **Sun, H.-C.**, Ban, H., Di Luca, M., & Welchman, A. E. (2015). fMRI Evidence for the Processing of Surface Gloss in Human Visual Cortex. *Vision Research*, 109, 149-157.
2. **Sun, H.-C.**, Welchman, A. E., Chang, D. H. F & Di Luca, M. (submitted). Look but don't touch: visual cues to surface structure drive somatosensory cortex.
3. **Sun, H.-C.**, Di Luca, M., Ban, H., Muryy, A., Fleming, R. W., & Welchman, A. E. (submitted). Differential processing of binocular and monocular gloss cues in human visual cortex.

Conference abstracts

1. **Sun, H.-C.**, Di Luca, M., Fleming, R. W., Muryy, A., Ban, H. & Welchman, A. E. (2015). The Perception of Glossiness in the Human Brain. Paper to be presented at the *15th Annual Meeting of the Vision Science Society (VSS)*, Florida, USA
2. **Sun, H.-C.**, Welchman, A. E. & Di Luca, M. (2014). Visual Material Perception in the Human Brain: Glossiness versus Roughness. Paper presented at *Neuroscience 2014 meeting*, Washington DC, USA
3. **Sun, H.-C.**, Welchman, A. E. & Di Luca, M. (2014). Different Networks of Visual Glossiness and Roughness Processing in the Human Brain. Paper presented at *Vision Leads to Action Conference*, Birmingham, UK
4. **Sun, H.-C.**, Welchman, A. E. & Di Luca, M. (2014). Visual Roughness and Glossiness Perception in the Human Brain. Paper presented at *Research Poster Conference*, Birmingham, UK

5. **Sun, H.-C.**, Ban, H. & Welchman, A. E. (2014). The Perception of Glossiness in the Human Brain. Paper presented at the *14th Annual Meeting of the Vision Science Society (VSS)*, Florida, USA

Contribution

The candidate took the major part in experimental design, stimuli generation, experimental programming, pilot testing, data collection, data analysis and manuscript writing in all the studies of this dissertation.

Table of contents

Chapter 1: Literature review

1.1 Material perception

Identifying the material of objects around us determines how we respond, interact or make use of the objects in our environment. For example, fabric objects like shirts are soft and can make us warm. Shiny metals and glasses are usually slippery and we need to apply extra force and care when grabbing them. Our visual system seems to be expert in recognizing a variety of material properties such as metal, fabric, wood, plastic and so on. This ability makes us able to identify objects easily in the real world with few errors. This also implies that some neural circuitries in our visual system might be specialized in processing material information (Hiramatsu, Goda, & Komatsu, 2011).

Our perception of outside world is multisensory. In material perception, we can distinguish the materials in the environment via different modalities such as vision (e.g the appearance of objects), touch (e.g. roughness, hardness and coldness of objects), smell (the odour of objects) and audition (e.g. the sound when we hit the object). For humans and other primates, the major sensory input is vision which covers the largest part of our cerebral cortex (Nakayama, He, & Shimojo, 1995). Most importantly, we perceive materials around us through vision first (in most of the cases, since we can view objects in a certain distance) and then we decide whether we need to further identify the material with other modalities (Bergmann Tiest, W. M. & Kappers, 2007; Giesel & Zaidi, 2013; Whitaker, Simões-Franklin, & Newell, 2008). Therefore, understanding how our visual system processes material information is a crucial issue and we specifically focus on the role of vision throughout our studies here.

Neural basis of material processing

Due to the complexity of the visual system and its hierarchal property, it is usually divided into three main levels according to its function: lower-level, mid-level (or intermediate level) and higher-level. Material perception is assumed to involve mid-level visual processes which are related to the representation of surface properties, while lower-level vision is related to local information encoding such as orientation, luminance, colour and contrast and higher-level vision is related to object or face recognition and categorization (Anderson, 2011; Nakayama et al., 1995). So mid-level vision might be the possible stage that our visual system processes material information, the next question would be: where might it be processed in the brain? What would be the candidate areas?

In previous human neuroimaging studies of material perception, material information was found to be represented in higher ventral visual areas such as fusiform gyrus (FG), inferior occipital gyrus (IOG) and collateral sulcus (CoS) (Cant, Arnott, & Goodale, 2009; Cant & Goodale, 2007, 2011; Cavina-Pratesi, Kentridge, Heywood, & Milner, 2010a, 2010b; Hiramatsu et al., 2011; Peuskens, et al., 2004). Studies showed that when participants attend to the material/texture of objects, CoS was activated more than when attending to other properties such as object shape, orientation or motion (Cant & Goodale, 2007; Peuskens et al., 2004). Subsequent research from the same group using fMRA (functional magnetic resonance adaptation) paradigm also found that anterior CoS was most sensitive to material change rather than object form change or color change (Cant et al., 2009). Cavina-Pratesi, Kentridge, Heywood and Milner (2010a, 2010b) further found that posterior CoS extracts information about surface properties (texture) and it is different from anterior CoS which mainly responded to color and it is also different from lateral occipital cortex that mainly responded to shape. Cant and Goodale (2011) further found that in addition to posterior CoS,

the Parahippocampal Place Area (PPA) also processes material information. All these studies indicated the importance of higher ventral areas around CoS in visual material processing.

It appears that in addition to higher ventral areas, low-level visual areas are also involved in extracting material information. Giesel and Zaidi (2013) found that participants' precepts of material properties (thickness, roughness and undulations) varied with spatial frequency bands and could be altered by adaptation to different frequency bands of noise. This result suggests that material information could be extracted from the processing of spatial frequency in V1 (Giesel & Zaidi, 2013). In addition, Hiramatsu, Goda and Komatsu (2011) used human fMRI and let participants viewed the images of nine-categories of material images (e.g. wood, fabric, fur etc.). They found that the activation patterns in early visual areas (V1, V2) corresponded to monocular information of materials (e.g. skew, kourtosis, color etc.) while activation in higher areas (FG/CoS) corresponded to perceptual properties of materials (e.g. the ratings in smooth-rough/dry-wet scales) (Hiramatsu et al., 2011). Similarly, a monkey fMRI study also found that monkey V1 primarily represented low-level image properties while V4 and the posterior inferior temporal (PIT) cortex represented the visuotactile properties of material (e.g. human rating scores in glossy-matte, smooth-rough, dry-wet etc.), similar to the human ventral cortex FG and CoS (Goda, Tachibana, Okazawa, & Komatsu, 2014). Single-cell recording in monkeys showed that most neurons in the inferior temporal (IT) cortex were tuned to specific material texture (e.g. sandpaper, bread, leaf etc.) regardless of shape changes (Köteles, De Mazière, Van Hulle, Orban, & Vogels, 2008). Another single-cell recording study in monkey V4 found that the majority of V4 neurons were selective to specific textures and these preferences remained with changes of illumination angles (Arcizet, Jouffrais, & Girard, 2008). Okazawa, Tajima, and Komatsu (2014) further found that texture selectivity of monkey V4 neurons can be explained by their responses to particular combinations of image statistics such as position,

orientation and scale. To sum up, these studies indicate that the neural basis of material perception is along ventral visual pathway.

Fleming (2014) suggested that the underlying mechanism of material perception can be specified as two interdependent processes: categorization and estimation. Categorization identifies what the material is (e.g. wood, stone, metal etc.). In contrast, estimation judges specific characteristics of a material (e.g. degree of glossiness, roughness, softness etc.). The two processes closely interact with each other that categorization facilitates estimation by extracting related properties within that category and estimation also helps categorization by providing information in feature spaces (Fleming, 2014; Fleming, Wiebel, & Gegenfurtner, 2013). The aforementioned neural basis of material perception is based on categorization processing and results showed that higher ventral visual cortex (especially in CoS) are involved in this procedure. However, how estimation processing occurs in the brain is not clear yet. The three studies here were designed to further examine this issue. We primarily focus on the estimation of two material properties: glossiness and roughness, which are described in the next two sections respectively.

1.2 Glossiness perception

Surface gloss changes dramatically with material composition and surface smoothness (Nishio, Goda, & Komatsu, 2012). It conveys some crucial information of object properties such as whether the object is soft or hard, smooth or rough, and dry or wet. Objects in our environment have different levels of glossiness, from very matte to very glossy. Matte surfaces reflect light diffusely in all directions while specular surfaces reflect light in a small, limited range of directions (Norman, Todd, & Orban, 2004). Lambertian surface is an extreme idealized example of matte surface which scatters light equally into all possible directions and lacks specular reflection (Anderson, 2011). In contrast, perfect mirror is an extreme example in another way. It reflects light in an angle equal to its incidence angle (Wendt, Faul, Ekroll, & Mausfeld, 2010). Most objects in our daily life have reflections in between, with a combination of both diffuse and specular reflection.

One of the most important model which is generally used in glossiness studies is Ward model. There are three parameters that are assumed to be important in defining surface gloss in Ward model: diffuse reflectance (ρ_d), specular reflectance (ρ_s) and roughness (α) (Ward, 1992). Diffuse reflectance decides the proportion of incident light reflected by the diffuse component. It specifies the lightness of an object. Specular reflectance decides the proportion of incident light reflected by the specular component. It controls the intensity of highlights. Roughness determines the blurriness and spread of highlights (Fleming, Dror, & Adelson, 2003; Nishio et al., 2012; Nishio, Shimokawa, Goda, & Komatsu, 2014; Olkkonen & Brainard, 2010). (Note that the 'roughness' parameter in Ward's model is different from the roughness that we described in the next section since the former indicate roughness in objects' microscale level while the latter is in mesoscale level). In general the three parameters are approximately independent to each other. A study using matching task

(observers adjusted diffuse and specular reflectance of a test sphere to match a reference sphere with different illumination field) found that the changing on diffuse reflectance component is independent of specular reflectance. Also, the changing on specular reflectance component is independent of diffuse reflectance, and both of them are independent of roughness component (Olkkonen & Brainard, 2010).

There are many factors that affect perceived surface gloss, from low-level to higher-level factors and from the properties of object itself to outside conditions. Interestingly, small variances of these factors sometimes changes perceived gloss dramatically. For low-level factors, some researchers assume that people judge surface gloss based on a simple parameter: the skewness of image luminance histogram (Landy, 2007; Motoyoshi, Nishida, Sharan, & Adelson, 2007). They found that when the skewness of image luminance histogram increases (positively skewed), the image is rated glossier. In other words, skewness is correlated with perceived glossiness positively. Moreover, after adaptation to a negative skewness image, surface looks glossier than the surface before adaptation. This aftereffect of perceived gloss implied the participation of early visual areas in gloss perception (Motoyoshi et al., 2007).

However, in Kim and Anderson's study (2010), they found that some of the gloss aftereffects cannot be explained by skewness. For example, they found that zero-skew adaptor can also reduce perceived gloss as positive-skew adaptor, while negative-skew could not facilitate perceived gloss as Motoyoshi et al. (2007) found (Kim & Anderson, 2010). Therefore, researchers argue that gloss perception is not just mediated by simple statistics. It involves in more complex mechanism which cannot be explained just by skew *per se* (Anderson & Kim, 2009; Kim & Anderson, 2010).

For higher-level factors that influence surface gloss, it is well known that objects with highlights on it look glossier than without highlights because glossy surface reflects the light form its environment easily. However, the position of highlights on object is also important. Rotation or translation of highlights from its original position can diminish perceived gloss while the skewness remained the same (Anderson & Kim, 2009; Kim, Marlow, & Anderson, 2011; Marlow, Kim, & Anderson, 2011). In particular, rotation reduced perceived gloss dramatically even if the highlights are rotated for small degree (Anderson & Kim, 2009; Marlow et al., 2011). These results cannot be explained by skewness since the skewness of image luminance histogram remains the same after rotation and translation. In other words, both orientation congruence (whether specular highlights and the diffuse shading surrounding them share the same orientation) and brightness congruence (whether specular highlights appear near to the brightest areas of the diffuse shading) are assessed for recovering surface gloss (Kim et al., 2011; Marlow et al., 2011).

Not only bright specular highlights but also dark specular lowlights can affect perceived gloss. Specular lowlights are defined as dark luminance extreme on surfaces. It was found that blurring or rotating specular lowlights also reduced perceived gloss (Kim, Marlow, & Anderson, 2012). These results suggest that the appropriate position and alignment of specular lowlights with the diffuse shading surrounding it is necessary to induce surface gloss, just like the properties of specular highlights mentioned above. It implied a common process underlying specular highlights and lowlights processing (Kim et al., 2012).

Another crucial factor that influences gloss perception is the disparity of highlights. For convex surfaces, only the highlights lay inside the surface (which is physically correct depth) were rated glossy and it looked less glossy when the highlights lay above the surface (which is physically incorrect depth) (Blake & Bülthoff, 1990; Kerrigan & Adams, 2013). However, for concave surfaces, both highlight disparities were perceived glossy. Therefore

the authors suggest that our visual system may use some simple heuristics (i.e highlights occur inside objects) in surface gloss judgment (Kerrigan & Adams, 2013). Another study found that highlight disparity strongly increased the strength of perceived glossiness, gloss constancy, and the authenticity of perceived gloss comparing with the stimuli without highlight disparity (Wendt et al., 2010; Wendt, Faul, & Mausfeld, 2008). Other studies found that perceived gloss was stronger when viewing a surface in stereo than viewing in non-stereo circumstance (monocular) (Obein, Knoblauch, & Viéot, 2004; Sakano & Ando, 2010). To sum up, all the studies consistently show that stereo cues play an important role in gloss perception.

Perceived gloss is not the same across a surface. It was found that gloss ratings decreased with the distance of probe point from highlight centre (Berzhanskaya, Swaminathan, Beck, & Mingolla, 2005). In addition, surface continuity is also important for gloss perception. When the highlight on an object was partially occluded (by occluders or by gaps), perceived gloss also decreased accordingly. These results demonstrated that surface representation is closely connected to gloss perception (Berzhanskaya et al., 2005).

Surface relief and illumination direction also affect surface gloss. One study found that the perceived illusory gloss of a stretched Lambertian surface increases with increasing its relief (depth range) but decreases with increasing illumination direction from frontal to grazing. Therefore the high relief surface with frontal illumination elicited the highest gloss perception (Wijntjes & Pont, 2010). A further study found interaction effect between surface relief and illumination direction. That is, under frontal illumination shallow reliefs were judged glossier than deep reliefs, but under oblique illumination deep reliefs were judged glossier than shallow reliefs (Marlow, Kim, & Anderson, 2012). Marlow and Anderson (2013) further argued that the failure of gloss consistency might because that our visual system relies on some image cues for judging surface gloss such as contrast, sharpness and

coverage of highlights. However, these cues may be modulated by object geometry and illumination direction, which are unrelated to specular reflectance (Marlow & Anderson, 2013).

Different types of illumination can also affect gloss perception. For example, the same material of objects looks glossier under small light sources illumination (which produce sharp highlights) than under diffuse light sources (Dror, Willsky, & Adelson, 2004). Moreover, real-world illumination makes objects look glossier than using point light source illumination (Dror et al., 2004). Different illumination fields may also induce a small difference in perceived gloss, but in general observers can retain (but not perfectly) surface reflectance properties such as specular reflectance, diffuse reflectance and roughness across different real-world illumination fields (Fleming et al., 2003; Olkkonen & Brainard, 2010), and most participants had certain glossiness transitivity (the gloss consistency under different light fields) (Doerschner, Boyaci, & Maloney, 2010). These results indicated that the illumination in the real world can also affect perceived gloss even if surface material and specularity remain the same (Doerschner, Boyaci, et al., 2010). However, artificial illuminations such as point light sources or pink-noise/white-noise light sources decreased observers' accuracy in surface reflectance estimation. Thus they suggested that human may rely on some assumptions (such as a dominant direction of illumination, extended highlight edges, and the organization of the edges) that usually present in real world to interpret surface gloss (Fleming et al., 2003).

Color information is also important in gloss perception. It was founded that objects were judged glossier when it is presented in front of a black background than when presented in front of a white background (Doerschner, Maloney, & Boyaci, 2010). Another study found that colour image facilitated gloss constancy better than greyscale image (Wendt et al., 2010). Still another study found that objects looked glossy when diffuse and specular reflectance

shared the same color or when specular reflectance was white and diffuse reflectance was colored (e.g., white on red). However, if it is in the opposite way (colored specular reflectance and white diffuse reflectance), perceived gloss reduced. The reason is that it is an unusual experience to see a white object with colored highlights. Thus, the colored highlights were perceived as pieces of colored foil stuck on a white matte object. It suggests that our visual system can correctly take into account of color information when estimate surface gloss (Nishida, et al., 2008).

Although surface properties like glossiness seems like to be a property of static objects, motion cues can also have great impact on gloss perception. One study found that perceived gloss was enhanced when observer's retinal image changed due to head motion comparing with static viewing (Sakano & Ando, 2010). This is because that the reflected features on the surfaces changed accordingly during self-motion which strongly signalled glossiness. Moreover, not only self-motion but also objects motion can change reflected features on objects that improved gloss perception (Wendt et al., 2010). Another study showed that participants judged rotating objects shinier when it was rendered with normal specular reflections than when rendered rigidly attached (painted) reflections during rotation (Doerschner, et al., 2011). The authors thus suggested three motion cues that our visual system may use for distinguishing between glossy and matte surfaces: coverage, divergence, and 3D shape reliability. Coverage is defined as the proportion of trackable features on an object when it is moving. If an object is very shiny, coverage will be low since the reflected features on its surface is distorted so that the trackability of their features will be impaired when the object is moving rapidly. Divergence is defined as the expansions and contractions caused by the optic flow on an object during moving. Shiny objects have high divergence since the reflected features on its surface will expend toward convexities and contract toward concavities while matte objects don't have this feature. 3D shape reliability is defined as the

consistency of the optic flow vectors on an object that are constrained by its shape geometry during moving (i.e. optic flow moves in the same way as object motion). The inconsistency usually signals shinny surfaces because the reflected features are not stuck to its surface. Similarly, observers performed better in judging gloss property differently when they were allowed to interact with the object (e.g. rotating the images by mouse) comparing with watching them passively (Scheller Lichtenauer, Schuetz, & Zolliker, 2013). This is because that they had visual feedback from object movement which facilitated their judgments of surface properties. To sum up, optic flow on moving objects provides important information about surface gloss (Doerschner et al., 2011).

Neural basis of gloss processing

The visual system seems to use a variety of cues in estimating surface gloss, including skewness, the position and orientation of highlights and lowlights, disparity, surface relief, illumination direction, illumination fields, occlusion, colour, and optic flow during motion. Although there are many studies investigating the factors that may have influence on gloss perception, however, the neural basis of it is still unclear. A monkey neuron-recording study in gloss perception showed that superior temporal sulcus (STS) selectively responds to different surface gloss properties. For example, some cells responded specifically to glossy objects with sharp highlights and other cells responded specifically to shiny objects with blurred highlights. Still other cells responded to matte objects without highlights (Nishio et al., 2012). Importantly, these cells maintained their selectively in preferred gloss even if the shape or illumination of the objects changed. This result implied that surface gloss information is represented in STS, which is within the monkey inferior temporal (IT) cortex. A monkey fMRI study found that specular objects, comparing with matte objects and

scrambled conditions, elicited more activation along the ventral visual pathway, form V1, V2, V3, V4 to IT (Okazawa, Goda, & Komatsu, 2012). A recent monkey neuron-recording study also found that gloss-selective neurons in IT systematically changed with physical gloss parameters (ρ_d, ρ_s, and α as described before) (Nishio et al., 2014). To sum up, higher-order ventral visual area around IT, which is the likely homolog of human LOC/FG/CoS (Goda et al., 2014; Köteles et al., 2008), is assumed to be important in surface gloss representation in the brain.

Although there are some monkey electrophysiological and fMRI studies in gloss perception (Nishio et al., 2012; Nishio et al., 2014; Okazawa et al., 2012), how the human brain processes surface gloss information is still not clear. A recent human fMRI study found that surface gloss activated V2, V3, V4, VO-1, VO-2, CoS, LO-1 and V3A/B (Wada, Sakano, & Ando, 2014). Moreover, gloss perception is assumed to be special that is processed independently from other material properties like colour and texture. This is evident from a brain-damage patient who had deficits in colour and texture discrimination but still showed clear ability in gloss perception (Kentridge, Thomson, & Heywood, 2012). Here we further examined this issue with different ways in manipulating glossiness (using both monocular and binocular gloss cues) and control conditions (spatial scrambling in Study 1, highlight rotation in Study 2, and highlight disparity in Study 3). The aim is to identify the brain areas preferentially responding to glossy object in the human brain. Importantly, we investigated how the brain processes different gloss cues and whether there is a 'gloss centre' in the brain which integrates and represents a variety of gloss cues (Study 3). Note that gloss information is not only processed by neurons preferentially responding to glossy objects than matte objects but also processed by neurons preferentially responding to matte objects than glossy objects as found in previous studies (Nishio et al., 2012; Nishio et al., 2014). We therefore applied multi-voxel pattern analysis (MVPA), seeking to detect pattern activation caused by

both types of glossiness-related neurons. In addition, glossiness is not the only material property we want to investigate. Here we also examined how the brain processes another material property (i.e. roughness) and the difference with gloss processing.

1.3 Roughness perception

Roughness is one of the most important material properties that signals important status of objects. For example, from the skin of fruits we can tell whether the fruit is fresh and juicy or dry and old. However, unlike glossiness which is mainly perceived through vision, roughness is a special surface property that not only vision but also haptic systems are able to estimate it (Bergmann Tiest, W. M. & Kappers, 2007). Rough surfaces produce vibrations when we stroked, and they create different levels of pressure on skin when touching (Bergmann Tiest, W. M., 2010). For the perception of roughness, tactile information might be even more dominant than visual input. A study has shown that visual roughness perception was modulated by tactile inputs when the two information modalities were incongruent but not vice versa. That is, participants made more errors when they had to judge on visual roughness with incongruent tactile distractors. However, when judging on tactile roughness, incongruent visual distractors did not affect much (Guest & Spence, 2003). This result suggests that roughness perception is more dominated by haptic than by visual information. Probably that is why studies about roughness perception primarily focus on haptic modality rather than vision (Bergmann Tiest, W. M., 2010; Connor & Johnson, 1992; Guest, Catmur, Lloyd, & Spence, 2002; Kahrimanovic, Bergmann Tiest, & Kappers, 2009; Kitada, et al., 2005; Kitada, Sadato, & Lederman, 2012; Klatzky & Lederman, 1999; Lawrence, Kitada, Klatzky, & Lederman, 2007; Lederman, S. J. & Klatzky, 2009; Libouton, Barbier, Berger, Plaghki, & Thonnard, 2012; Libouton, Barbier, Plaghki, & Thonnard, 2010; Roland, O'Sullivan, & Kawashima, 1998; Schütz-Bosbach, Tausche, & Weiss, 2009; Suzuki, Gyoba, & Sakamoto, 2008; Yoshioka, Craig, Beck, & Hsiao, 2011; Yoshioka, Gibb, Dorsch, Hsiao, & Johnson, 2001). However, the understanding of visual roughness perception is relatively less and we further focus on this next.

The most powerful factor that influences perceived roughness is the spacing between the elements (i.e. groove width) that constitute the surface (Bergmann Tiest, W. M., 2010; Connor & Johnson, 1992; Eck, Kaas, Mulders, & Goebel, 2013; Kahrimanovic et al., 2009; Kitada et al., 2012; Lederman, S. J. & Klatzky, 2009; Whitaker et al., 2008; Yoshioka et al., 2001). A study found that the peak of roughness was perceived when inter-element spacing of dot patterns was about 1.25 mm (viewing distance 75 cm) and 3-4 mm for visual and haptic modality respectively due to the better spatial resolution of visual modality than haptic modality (Eck, Kaas, Mulders, et al., 2013).

Similar to gloss perception, the outside conditions such as the direction of illumination can also affect visually perceived surface roughness. Participants perceived surfaces to be rougher with smaller illumination angle (more grazing angle) relative to surfaces (Ho, Landy, & Maloney, 2006). Another study further found that participants' viewpoint also had influence on perceived surface roughness (Ho, Maloney, & Landy, 2007). Their explanation for these failures in visual roughness consistency was that the greater proportion of visible shadows contained in a surface, the rougher the surface will be perceived (Ho et al., 2007). In other words, proportion of shadows which changes with illumination angle and viewpoint affects our interpretation of surface roughness.

Neural basis of roughness processing

Due to the lack of evidence about how the brain process visual roughness information, studies about visual material processing and haptic roughness processing may provide some insights about this. As aforementioned, neural basis of general material processing (categorization) and visual glossiness processing are mainly in higher ventral visual cortex. Therefore visual roughness processing may also involve in the nearby areas. In addition, due

to roughness is a special material property that both visual and haptic systems are able to perceive it (Bergmann Tiest, W. M. & Kappers, 2007; Guest & Spence, 2003), it would be possible that visual roughness may activate some common areas (bimodal areas) as by haptic roughness. Thus visual roughness processing may also relate to the areas responsible for haptic processing.

Human brain imaging in haptic roughness perception. The neural basis of haptic roughness perception involves mainly in somatosensory areas. A human repetitive transcranial magnetic stimulation (rTMS) found that low-frequency (1 Hz) rTMS applied to somatosensory cortex dampened overall tactile roughness discrimination (roughness rating) while tactile distance discrimination (rating for dot spacing: far to close) remained intact. This evidence suggests the engagement of somatosensory cortex in tactile roughness perception (Merabet, L., et al., 2004). More precisely, parietal operculum (PO) which is inside secondary somatosensory cortex (S2) was found to be important for tactile roughness perception in many studies. A brain imaging study using positron-emission tomography (PET) measurement found that lateral parietal opercular cortex (PO) was activated more when participants doing roughness discrimination of objects with their right hand than doing length or shape discrimination. In contrast, the anterior part of the intraparietal sulcus (IPA) was activated more when participants discriminated length or shape of objects with their right hand than when discriminating roughness. This suggests that PO and IPA in the somatosensory areas have separate functional contributions in different submodalities of tactile inputs (Roland et al., 1998).

Another human fMRI study by Lederman et al. had slightly different findings. They found that contralateral postcentral gyrus (PCG) (where the somatosensory areas are located in) were activated when participants classified objects along the three roughness levels (‘smooth’, ‘medium’, and ‘rough’) and classified objects by three different shapes. PO was

activated only when participants classified objects along the three hardness levels ('soft', 'firm', and 'hard'). In other words, this study did not find the involvment of PO in tactile roughness classificaiton, which is conflicted with other studies (Servos, Lederman, Wilson, & Gati, 2001).

Despite of this, the other studies all consistently found the importance of PO in tactile roughness processing. A human fMRI study showed that parietal operculum (PO), insula and right lateral prefrontal cortex were involved in tactile roughness perception. Brain activations in these areas showed negatively graded response (activation increased with decreased perceived surface roughness) during roughness estimation task (roughness rating for each test surfaces). More specifically, PO and insula were also activated during no-estimation task (simply to attend to the stimulus without rating roughness) comparing with rest condition. By contrast, right lateral prefrontal cortex was activated only during estimation task comparing with no-estimation task and rest condition. Thus they assumed that PO and insula might be related to sensory processing of surface roughness estimation while right lateral prefrontal cortex might be involved in cognitive processing of roughness estimation (Kitada et al., 2005).

Further studies also showed consistent results. Simões-Franklin, Whitaker and Newell (2011) found that PO activation was strongly affected by roughness of sandpapers that coarse surfaces elicited greater activation than fine surfaces and its activation was not affected by exploration type (active touch *vs.* passive touch). In contrast, primary somatosensory area (S1) was affected by the latter (stronger activation for active touch than passive touch) but not the former. These results indicated a hierarchy in tactile information processing, from exploration procedure (active or passive) to stimulus properties (e.g. surface roughness) (Simões-Franklin, Whitaker, & Newell, 2011).

Another human fMRI study by Kaas, van Mier, Visser, and Goebel (2013) also emphasised the importance of PO in encoding and maintaining tactile texture information. Participants were asked to judge whether the texture of a sandpaper matched with one of previous samples with memory load manipulation (1 to 4 samples to choose from). In location judgment they judged whether location of a sandpaper matched with one of previous samples. They found both task effect (activation for texture judgment > activation for location judgment) and load effect (stronger activation for larger load) in PO, suggesting its involvement for encoding and maintaining of tactile texture information (Kaas, van Mier, Visser, & Goebel, 2013).

Monkey brain activities in haptic roughness perception. A monkey single-cell recording study also found that S2 (where PO is located in) neurons code tactile groove-width information that firing rate of some neurons changed with groove-width positively and some changed negatively. Since groove width is closely related to roughness perception, this evidence suggests a direct involvement of S2 in roughness perception (Pruett, Sinclair, & Burton, 2000). Put together with previous human researches, studies consistently showed that the processing of haptic roughness information occurs in PO.

Visual-haptic texture perception. We develop multisensory experience everyday such as we touch an object while look at it. This experience may establish the connection of visual frequency information and tactile vibration (e.g. viewing high-spatial frequency components while feeling high-frequency vibrations on the hands). Indeed, evidence showed that participants were able to adjust tactile pulse rates to match visual spatial frequency showed by Gabor and the result is consistent across participants, indicating visual-haptic crossmodal association (Guzman-Martinez, Ortega, Grabowecky, Mossbridge, & Suzuki, 2012). The study of Eck, Kaas and Goebel (2013) also supports this viewpoint. They conducted a human fMRI study with presenting dot pattern visually, haptically, and both. They found that cortical

activations changed in early visual and somatosensory areas when presenting stimuli bimodally comparing with visual only or haptic only condition. Therefore they suggest crossmodal interactions of texture processing happened in early visual and somatosensory cortex. (Eck, Kaas, & Goebel, 2013).

Stilla and Sathian (2008) conducted a human fMRI study in which they investigated texture and shape perception with visual and haptic stimuli presentation respectively. They found that comparing with haptic shape discrimination, haptic texture discrimination elicited stronger activaion in ventral somatosensory area including PO and posterior insula, and in right medial occipital cortex (MOC, mainly in V2). In vision, visual texture discrimination also elicited stronger activaion in MOC than in visual shape discrimination. In other words, right MOC responded to texture information bimodally. The authors suggested the involvement of MOC was trigged by microgeometric, fine-grained discrimination in both visual and haptic texture task (Stilla & Sathian, 2008). Similarly, Sathian et al. (2011) found visual texture discrimination elicited stronger activation in the left MOG and right posterior fusiform gyrus (pFG) comparing with visual location discrimination. Haptic texture discrimination elicited stronger activation in PO, MOG, ventral premotor cortex (PMv), the orbitofrontal cortex (OFC), and inferior frontal gyrus (IFG). Bisesonsory texture-selective activations was found in MOG, left PMv, left lingual gyrus (LG), and left IFG extending into the inferior frontal sulcus (IFS). Further connectivity analysis showed that haptic texture information went through modality-specific area PO to bisensory texture-selective area MOG, and this pathway is different from location processing, which went through dorsal frontoparietal cortex (Sathian, et al., 2011). Although the two studies showed that MOC is a texture-selective area in both visual and haptic modality, however, a recent study showed that visual texture and haptic texture information activated adjacent but distinct areas within MOC (Podrebarac, Goodale, & Snow, 2014).

To sum up, no studies so far have found visual-haptic texture (or roughness) selective regions in somatosensory areas. Potential bisensory texture-selective areas were only found in visual area MOC (Eck, Kaas, & Goebel, 2013; Sathian et al., 2011; Stilla & Sathian, 2008). Although Eck, Kaas and Goebel (2013) reported that cortical activations in somatosensory areas can be modulated by visual inputs by comparing the activation in visual-haptic condition (presenting dot pattern in both modalities) with activation in haptic condition (with visual control stimuli), however, whether visual texture information *per se* can affect somatosensory activation is not clear. Here we examined this issue directly by presenting glossy and rough objects visually and measuring both visual and somatosensory activation (chapter 3). Note that it is still not clear whether roughness (which variances in meso-scale level) and glossiness (which variances in micro-scale level) are coded in a single continuum or different continuums and whether there is a difference between visual and somatosensory modalities. However, the different activation patterns caused by roughness and glossiness should all be detectable with multivoxel pattern analysis either way.

1.4 Research purposes

Studies investigating the neural basis of material perception are primarily based on categorization processing (identifying different materials). The results consistently showed that higher ventral visual cortex (especially in CoS) are involved in this procedure (Arcizet et al., 2008; Cant et al., 2009; Cant & Goodale, 2007, 2011; Cavina-Pratesi et al., 2010a, 2010b; Goda et al., 2014; Hiramatsu et al., 2011; Köteles et al., 2008; Okazawa, Tajima, & Komatsu, 2014; Peuskens et al., 2004). However, material properties are usually different across material categories. For instance, objects in 'metal' category usually have high glossiness and coldness and low roughness and softness. In contrast, objects in 'fabric' category usually have low glossiness and coldness and high roughness and softness. How the estimation processing of specific material properties occurs in the brain is not clear yet. One of the most important material properties that gains lots of research interest is glossiness. Several studies about glossiness processing have been established based on the macaque model (Nishio et al., 2012; Nishio et al., 2014; Okazawa et al., 2012). Studies about how the human brain processes gloss information is still rare, and the first one was published only recently (Wada et al., 2014). In addition, all the human and macaque studies described above focused on monocular gloss cue only. How and where the brain processes other gloss cues is still unknown. Therefore, the main purpose of the current studies is to enhance the understanding of how the human brain processes information about surface gloss. We used human fMRI measurement throughout all the studies here.

In the first study (chapter 2), we examined how the human brain processes the conventional monocular gloss cue – specular reflectance. Different from the study by Wada

et al. (2014) who contrasted glossy and matte objects under bright and dim illumination to exclude the confounding of luminance, we used spatial scrambling for controlling instead. This is a different approach of perturbing global image arrangement while preserving local image features to target mechanisms of the global gloss cue synthesised by specular reflectance while dissociating the role played by local image features. In addition, we performed special localizers to define important brain areas. Wada et al. (2014) only had retinotopic mapping localizer and hMT+ localizer. Here we used two more important localizers to define LO, pFs and V3B/KO. Since previous studies found ventral visual areas such as FG/CoS were important, we sub-divided ventral regions based on object localizer to define LO and pFs. Moreover, unlike Wada et al. (2014) who combined V3A/B together for analysis, here we defined V3B/KO with a separate localizer and analyzed V3A and V3B/KO separately so that we sought to understand the difference of the two areas more precisely. We also defined STS since this area was found crucial for glossiness processing in monkey studies. Finally, we had different data analysis approach than Wada et al. (2014). In particular, we used Granger causality analysis (GCM) in mapping information flows in each ROI based on this approach.

In the second study (chapter 3), we further investigated the neural basis for tactile predictions from visual stimulation. We focus on not only visual areas but also somatosensory areas because somatosensory cortex might be important in this prediction processing. We analyse brain activation elicited by objects in four conditions with different visual appearances: glossy, glossy control, rough, or rough control. Glossy objects are similar as those in the last study with specular highlights on them. We used rotated specular components for its control, which is different from spatial scrambling in Study 1. Rough objects were generated by applying wave textures on objects geometry while in its control the same wave textures were applied to objects surface only. We contrasted glossy objects with

rough objects because they have very different sense of touch, and we also contrasted between their controls for comparison because they have the same sense of touch (all matte) with similar low-level image features as their original counterparts.

In addition to the localizers for visual areas described in previous studies, we used specialized somatosensory localizer to localize primary and secondary somatosensory areas. Previous studies failed to find evidence of somatosensory areas in processing visual texture information with conventional BOLD signal change analysis (Sathian et al., 2011; Stilla & Sathian, 2008). Here we performed multi-voxel pattern analysis (MVPA), seeking to detect subtle pattern activation in somatosensory areas. From this study we could further clarify visual-haptic crossmodal association in material perception.

In the third study (chapter 4), we investigated how the brain process different gloss cues. We identified the brain areas preferentially responding to glossy object with different ways in manipulating their glossiness. Except for conventional monocular gloss cue (specular reflectance), here we firstly examined how the how the brain process binocular gloss cue by manipulating binocular disparities of specular reflections. We manipulated the disparities of specular reflections of the environmental on random shapes to make it look glossy (normal reflection like pure mirror objects) or matte (specular reflections coincident with the surface disparities like painted). We further compared brain processing between monocular and binocular cues about surface gloss. Importantly, we performed transfer analysis in MVPA to test whether a classifier trained to discriminate glossy vs. matte objects with monocular cues was able to discriminate those with binocular cues, and vice versa. The rationale is that the area which shows the transfer effect is more likely to be the locus that integrates different gloss cues (at least for the integration of monocular and binocular gloss cue). From this study we were able to check whether there is a general ‘gloss centre’ in the brain that integrates and represents different gloss cues together or different gloss cues are processed separately.

Chapter 2:

fMRI evidence for areas that process surface gloss in the human visual cortex

This chapter has been published in

Sun, H.-C., Ban, H., Di Luca, M. & Welchman, A. E. (2015). fMRI Evidence for the Processing of Surface Gloss in Human Visual Cortex. *Vision Research*, 109, 149-157.

Foreword

In this chapter we investigated how the visual system processes glossiness information in the human brain, since most studies about it have been established based on the macaque model (Nishio et al., 2012; Nishio et al., 2014; Okazawa et al., 2012) and the first human fMRI study was just published recently (Wada et al., 2014). Different from the study by Wada et al. (2014) who contrasted glossy and matte objects under bright and dim illumination to exclude the confounding of luminance, we use spatial scrambling for controlling instead. We also used special localizers to define LO, pFs and V3B/KO, since ventral areas and V3A/B were found important so we sub-divided these areas and investigated their contributions to gloss processing more precisely. Finally, we exerted different data analysis than Wada et al. (2014), in particular using Granger causality analysis (GCM) in mapping information flows during the processing about surface gloss.

Abstract

Surface gloss is an important cue to the material properties of objects. Recent progress in the study of macaque's brain has increased our understating of the areas involved in processing information about gloss, however the homologies with the human brain are not yet fully understood. Here we used human functional magnetic resonance imaging (fMRI) measurements to localize brain areas preferentially responding to glossy objects. We measured cortical activity for thirty-two rendered three-dimensional objects that had either Lambertian or specular surface properties. To control for differences in image structure, we overlaid a grid on the images and scrambled its cells. We found activations related to gloss in the posterior fusiform (pFs) and in area V3B/KO. Subsequent analysis with Granger causality mapping indicated that V3B/KO processes gloss information differently than pFs. Our results identify a small network of mid-level visual areas whose activity may be important in supporting the perception of surface gloss.

Keywords

Surface gloss; material perception; posterior fusiform; V3B/KO; fMRI

2.1 Introduction

Surface gloss provides an important cue to an object's physical material and its microstructure (Nishio et al., 2012). From a perceptual perspective, it has particularly intriguing properties because there are cases where glossiness is specified only by small image areas containing highlights (Beck, 1972). Unlike other aspects of material, a slight change in an object (e.g. minor change of material or smoothness) can cause huge differences in the perceptual impression of gloss (Fleming, 2012). While a number of image cues have been proposed to modulate gloss perception, it is an open challenge to understand how this information is processed to infer surface material.

Psychophysical studies suggest that the brain uses a variety of visual signals to estimate gloss. For instance, low-level factors such as the image luminance histogram skew can bias perceived gloss and cause perceptual aftereffects (Gegenfurtner, Baumgartner, & Wiebel, 2013; Landy, 2007; Motoyoshi et al., 2007). Mid-level factors such as specular reflections (Nishio et al., 2012; Norman et al., 2004; Okazawa et al., 2012; Okazawa, Koida, & Komatsu, 2011; Wendt et al., 2010) and surface relief (Anderson, Marlow, & Kim, 2012; Marlow & Anderson, 2013; Marlow et al., 2012; Wijntjes & Pont, 2010) also influence the impression of gloss. Highlights play a particularly important role in affecting judgments of material, and this can relate to their position and orientation (Anderson & Kim, 2009; Anderson, Kim, & Marlow, 2011; Kim & Anderson, 2010; Kim et al., 2011; Marlow et al., 2011), their colour (Nishida, Motoyoshi, & Maruya, 2011; Wendt et al., 2010), and their binocular disparity (Kerrigan & Adams, 2013; Muryy, Fleming, & Welchman, 2012; Wendt et al., 2010; Wendt et al., 2008). Here we chose to investigate how manipulating surface appearance through highlights gives rise to changes in brain activity. In particular, we use

fMRI to identify the cortical regions that respond preferentially to visual gloss depicted by highlights.

Recent studies have suggested candidate areas in macaque brain that may play an important role in processing gloss (Nishio et al., 2012; Okazawa et al., 2012). For instance, specular objects elicited more fMRI activation along the ventral visual pathway, from V1, V2, V3, V4 to inferior temporal (IT) cortex compared to matte objects and phase-scrambled images of the objects (Okazawa et al., 2012). Single-unit recordings from the superior temporal sulcus (STS) within IT cortex identified neurons that were selective for gloss uninfluenced by changes in the 3D structure of the viewed object or by changes to the illumination (Nishio et al., 2012). Further, these gloss-selective responses reflect combinations of reflectance parameters that align to the perceptual dimensions guide judgments of surface properties (Nishio et al., 2014). These results from the macaque indicate that specular reflectance properties are likely to be encoded in ventral visual areas.

Despite this recent progress in the macaque model, we still have rather little insight into how the human brain processes gloss. Human brain imaging work examining the (more general) representation of material properties (e.g. wood vs. metal) implied a role of ventral visual areas, especially in fusiform gyrus (FG), inferior occipital gyrus (IOG) and collateral sulcus (CoS) (Cant et al., 2009; Cant & Goodale, 2007; Hiramatsu et al., 2011). This work employed stimulus changes in multiple image dimensions (e.g. colour, texture and gloss), meaning that activity related to gloss *per se* could not be determined. It is likely to be an important distinction as tests of a neuropsychological patient who had deficits in colour and texture discrimination showed that they were unimpaired on gloss judgments (Kentridge et al., 2012). This suggests that the cortical processing of gloss is (at least partially) independent from the processing of other material properties. Recently, Wada and colleagues (Wada et al., 2014) reported that fMRI activity related to surface gloss is evident in V2, V3, V4, VO-1,

VO-2, CoS, LO-1 and V3A/B. In particular, they contrasted glossy and matte objects under bright and dim illumination to exclude the confounding of luminance. Here we use the different approach of perturbing global image arrangement while preserving local image features to target mechanisms of the global synthesis of image cues when judging gloss. It is also different from Okazawa et al. (2012) who contrasted glossy objects with phase-scrambled versions of these objects. We presented observers with stimuli from four experimental conditions: Glossy, Scrambled Glossy, Matte and Scrambled Matte. Thereby we sought to discriminate Gloss vs. Matte renderings of objects while dissociating the role played by local image features. We looked into earlier visual areas (V1-V4), ventral visual areas (LO and pFs) and STS in this processing as these areas were found important for processing in material-related information previous studies. In addition, we also included V3A and V3B/KO as they were found involving in glossiness processing recently (Wada et al., 2014). hMT+/V5 was also included as in their study to double check its function in gloss processing.

2.2 Methods

2.2.1 Participants

Fifteen participants who had normal or corrected-to-normal vision were recruited for the experiment. Two were authors (H.-C. S. and H. B.) and the remainder were naïve participants. All were screened for normal stereoacuity and MRI safety before being invited to participate. All participants had previously participated in other fMRI studies in which fMRI localiser data (see 'ROI definition') and a T1-weighted anatomical scans (see 'MRI data acquisition') were acquired. The age range was 19 to 35 years old, and 13 of the 15 were male. All participants gave written informed consent before taking part in the experiment. The study was approved by the STEM Ethical Review Committee of the University of Birmingham. After completing the experiment, non-lab member participants received monetary compensation.

2.2.2 Apparatus and Stimuli

Stimuli. The stimuli comprised 32 2-D renderings of 3-D objects generated in Blender 2.67a (The Blender project: http://www.blender.org/). The objects were spheres and tori whose surfaces were perturbed by random radial distortions to produce slightly irregular shapes. The diameter of the stimuli was 12 deg on average and they were presented on a mid-gray background. We illuminated the objects using a square light source located front and above the objects. We chose this simple light source to be able to increase the influence of our scrambling manipulation. We created versions of the stimuli for each object that made up

the four conditions of the experiment: Glossy, Scrambled Glossy, Matte and Scrambled Matte (**Figure 1**). In the Glossy condition, objects were rendered using a BSDF (bidirectional scattering distribution function) mixed shader with 90% diffuse and 10% glossy components. We rendered objects in the Matte condition by setting the BSDF reflectance function to Lambertian (100% diffuse component). We controlled the luminance of the stimuli so that the mean luminance of the stimuli was 60.54 cd/m^2 and the absolute maximum was 103.92 cd/m^2 which corresponded to 57.55% and 98.78% of the display maximum luminance respectively. All the objects were rendered without background then we set background colour to gray before further manipulations as described below.

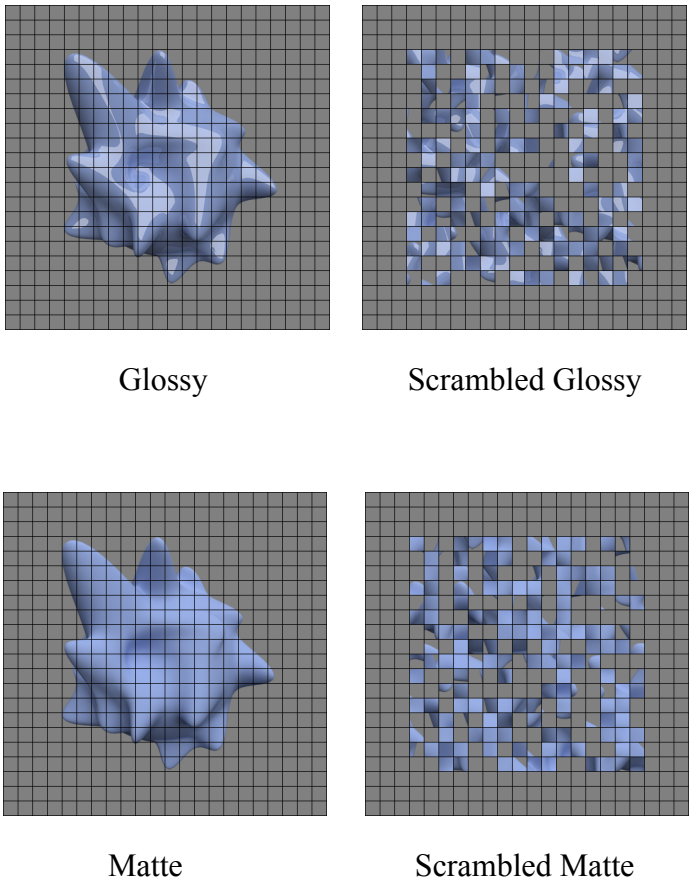


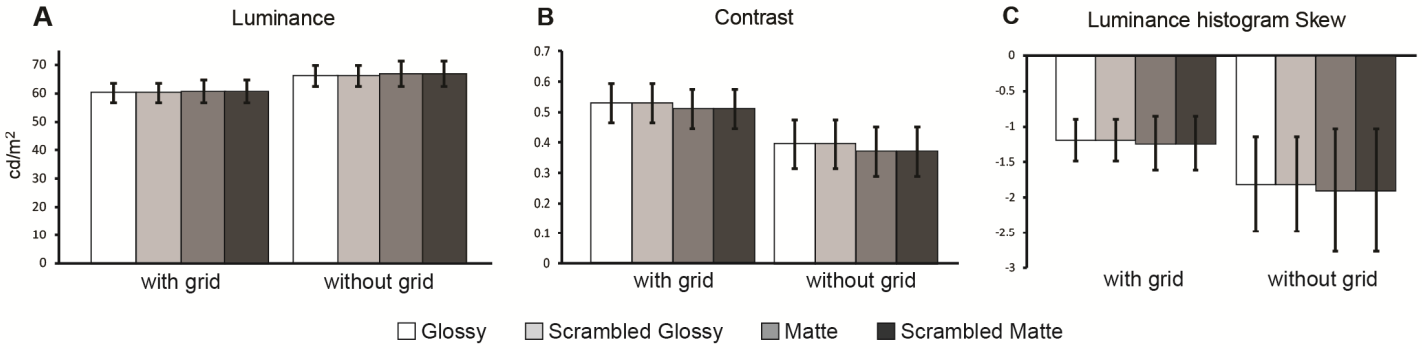
Figure 1. The four experiment conditions (Glossy, Scrambled Glossy, Matte and Scrambled Matte) rendered on an example object. Glossy components were shown in the Glossy

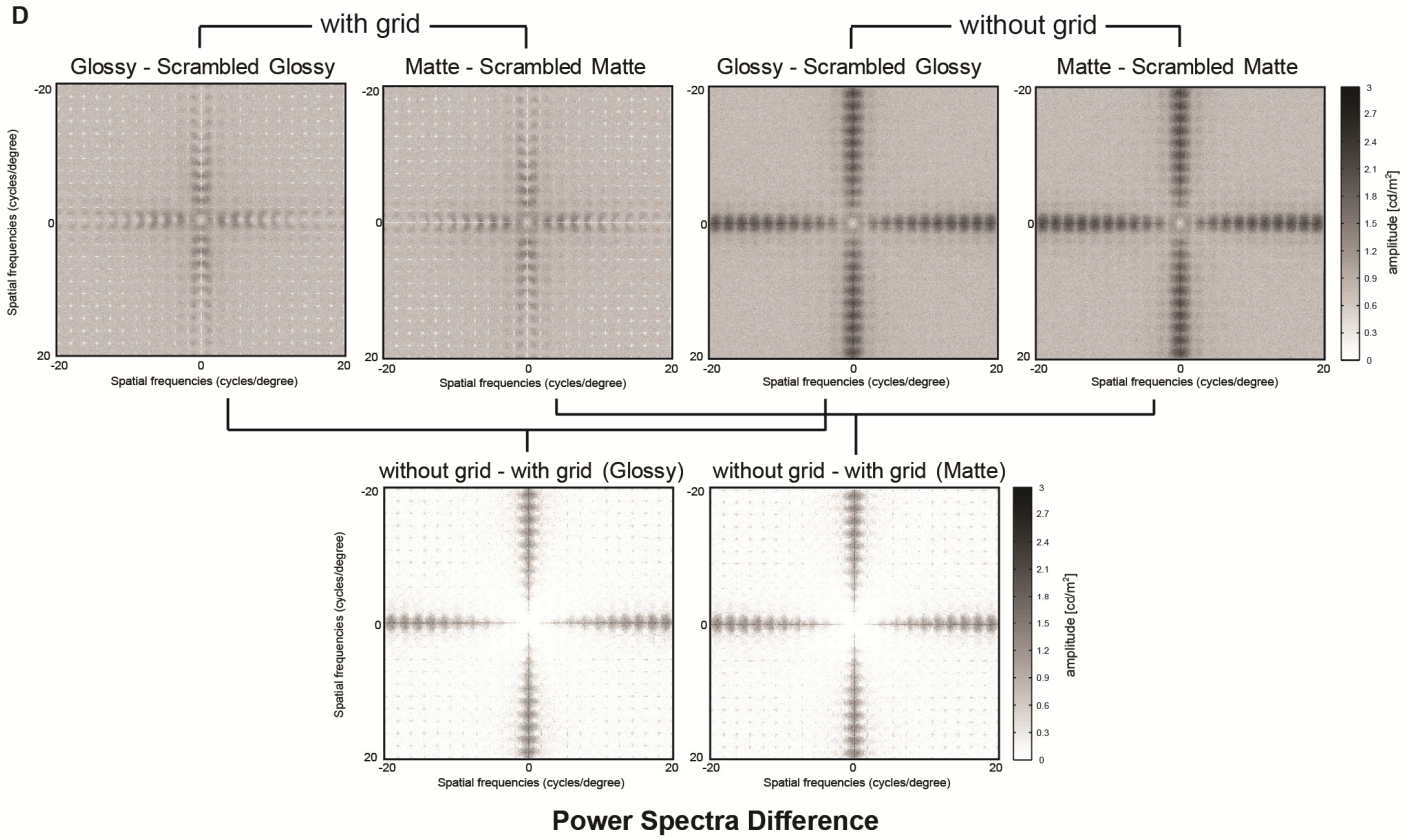
condition while only the diffuse components (Lambertian reflectance function) were shown in Matte condition. In the scrambled conditions, a 22 × 22 grid was superimposed over the images and then squares were randomly relocated within the grid.

To produce spatial scrambling, we superimposed a 22 × 22 1-pixel black grid over the images and then randomly relocated squares (0.55 deg of side) within the grid (Kourtzi & Kanwisher, 2000; Malach, et al., 1995). This approach differs from phase scrambling (Okazawa, et al., 2012) as blur, contrast, and luminance are only marginally affected. Moreover, the mosaic spatial scrambling approach we used interrupts object shape, shading, and specular highlights while all the local information (e.g., luminance, contrast, luminance histogram skew) is unchanged. Previous work indicates that highlight congruence with surface geometry and shading is crucial for perceived glossiness (Anderson & Kim, 2009; Kim et al., 2011; Marlow et al., 2011). Thus our stimuli strongly attenuate the impression of gloss by disrupting the relationship between highlights and global object structure.

Note that the superimposed grid was presented for both intact and scrambled versions of the stimuli. This greatly attenuates the amount of additional edge information that results from the spatial scrambling manipulation. Formally, we assessed differences in image structure by computing possible image cues that might drive the fMRI response. In particular, we found that the image statistics of mean luminance, luminance root-mean-square contrast, and luminance histogram skew were matched across the four conditions (**Figure 2**) indicating that there was more variation within the same class of stimuli than there was between classes. This is trivial for the scrambled versions of the stimuli (they must have the same values of skew, contrast and luminance, by definition), however, it is important that matte and glossy stimuli were well matched. In such a case, although the addition of a grid affects all these

measures, it did not create any consistent difference across the four conditions, thus the interpretation of the results should not be affected. Furthermore, the power spectra of the stimuli in the different conditions (**Figure 2D**) indicate that the grid is effective in equalizing the spatial frequency content of the images, particularly when contrasted with scrambled images without a superimposed grid. The grid adds high frequency components to intact images creating a pattern that is very similar to the one due to the scrambling procedure. In this way, frequency spectra are made more similar across conditions. Although there are still some spatial frequency differences across the four conditions with the grids, however, the difference between [Glossy vs. Scrambled Glossy] and [Matte vs. Scrambled Matte] is very similar. Therefore, after contrasting intact objects with their scrambled counterparts, spatial frequency difference between Glossy and Matte conditions should be very limited.





Power Spectra Difference

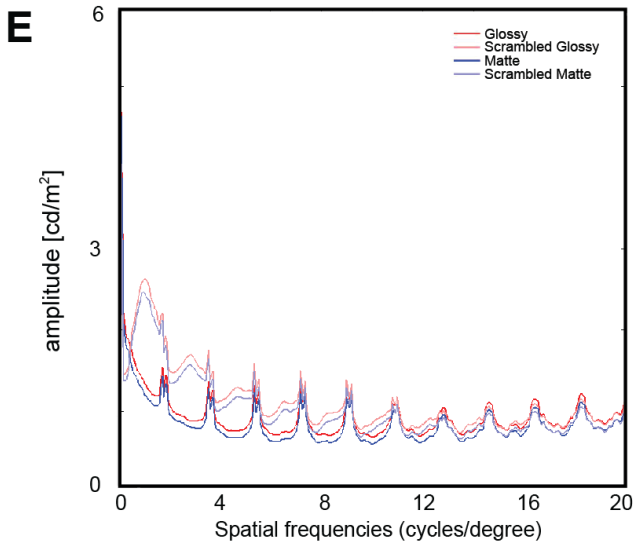


Figure 2. Image statistics of (A) pixelwise luminance, (B) contrast, (C) histogram skew, (D) difference in power spectra across the 32 images with and without the superimposed grid, and (E) spatial frequencies of the four conditions (with grids) averaged across images. Luminance was calculated by averaging the mean luminance of all pixels in each image then averaging across images. Contrast was calculated with pixelwise lumincance's standard deviation divided by its mean for each image (SD/Mean), averaged across images. Skew was calculated as the third standardized moment of the luminance histogram of each image, averaged across images (Motoyoshi et al., 2007). The absolute difference in power spectra was calculated for each image pair and then averaged across images.

Apparatus. Stimulus presentation was controlled using MATLAB (Mathworks Inc.) and Psychtoolbox (Brainard, 1997; Pelli, 1997). The stimuli were back projected from a JVC DILA SX21 projector onto a translucent screen inside the bore of the magnet. Participants viewed the stimuli binocularly via a mirror fixed on the head coil with a viewing distance of 64 cm. Luminance outputs were linearized and equated for the RGB channels separately with colorimeter measurements. A five-button optic fiber button box was provided to allow responses during the 1-back task.

MRI data acquisition. A 3-Tesla Philips scanner and a 32-channel phase-array head coil were used to obtain all MRI images at the Birmingham University Imaging Centre (BUIC). Functional whole brain scans with echo-planar imaging (EPI) sequence (32 slices, TR 2000 ms, TE 35 ms, voxel size 2.5 × 2.5 × 3 mm, flip angle 80 deg, matrix size 96 × 94) were obtained for each participant. The EPI images were acquired in an ascending interleaved order for all participants. T1-weighted high-resolution anatomical scans (sagittal 175 slices,

TR 8.4 ms, TE 3.8 ms, flip angle 8 deg, voxel size: 1 mm^3) were obtained from previous studies.

2.2.3 Design and Procedure

A block design was used. Each participant took part in 8 to 10 runs with 368 s length of each run in a 1.5 hour session. Each run started with four dummy scans to prevent startup magnetization transients and it consisted of 16 experimental blocks each lasting 16 s. There were 4 block types (i.e., one for each condition), repeated four times in a run. During each block, eight objects were presented twice in a pseudo-random order. Stimuli were presented for 500 ms with 500 ms interstimulus interval (ISI). Participants were instructed to maintain fixation and perform a 1-back matching task, whereby they pressed a button if the same image was presented twice in a row. They were able to perform this task very well (mean d'>3). Five 16 s fixation blocks were interposed after the third, fifth, eighth, eleventh and thirteenth stimulus blocks to measure fMRI signal baseline. In addition, 16 s fixation blocks were interposed at the beginning and at the end of the scan, making a total of seven fixation blocks during one experimental run. An illustration of the scan procedure is provided in **Figure 3**.

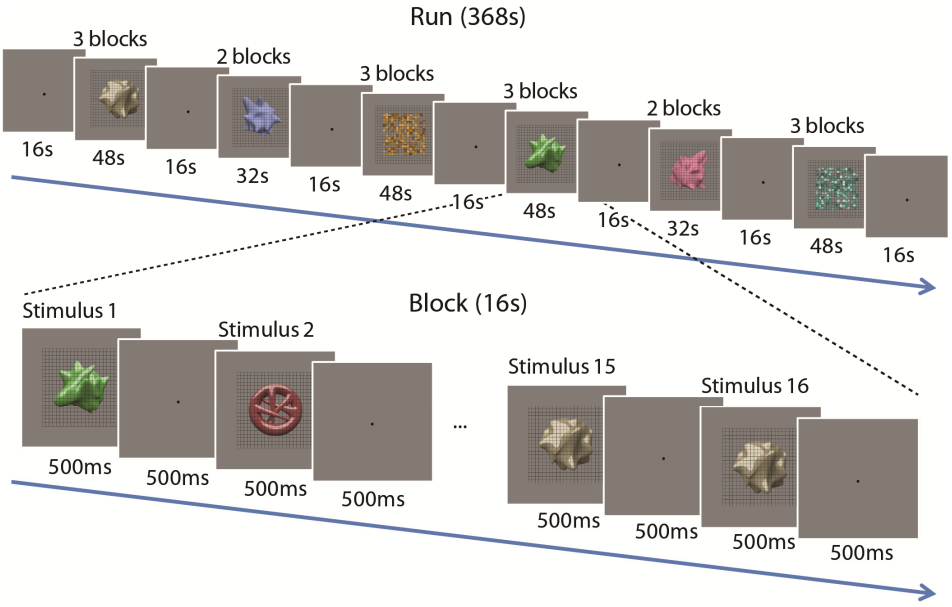


Figure 3. The fMRI procedure for one scan run. One each run there were 23 blocks (16 s each), including 7 fixation blocks and 16 experimental blocks. During each experimental block, stimuli were presented for 500 ms with 500 ms interstimulus interval (ISI). Participants were instructed to perform a 1-back matching task.

2.2.4 Data analysis

Functional MRI data processing. BrainVoyager QX version 2.6 (Brain Innovation, Maastricht, The Netherlands) was used for MRI data processing. Each participant's left/right cortical surfaces were reconstructed by segmenting gray and white matter, reconstructing the surfaces, inflating, cutting and then unfolding. All functional images were pre-processed with slice scan timing correction, 3D head motion correction, high-pass filtering (2 cycles per run cut-off) and linear trend removal. Functional images were co-registered with anatomical images and then transformed to Talairach coordinate space and aligned with each other. We computed the global signal variance of the blood oxygenation level dependent (BOLD) signal for each run using the whole-brain average of activity across volumes. If this exceeded 0.16% the scan run was excluded from further analysis to avoid the influence of scanner drifts, physiological noise or other artifacts (Junghöfer, Schupp, Stark, & Vaitl, 2005). On this basis, 17/146 runs across 15 participants were excluded from further analysis. A 3D Gaussian spatial smoothing kernel with 5 mm full-width-half-maximum (FWHM) was applied before analysing the data using a group-level random effects (RFX) general linear model (GLM).

ROI definition. A total of 11 regions of interest (ROIs) were defined. For each participant V1, V2, V3d, V3A, V3v, V4 were drawn by visual inspection of the data obtained from a standard retinotopic mapping scan preceding the experiment (Preston, Li, Kourtzi, & Welchman, 2008). V3B/KO (kinetic occipital region), hMT+/V5 (human motion complex)

and LO (lateral occipital region) were defined by additional functional localizers respectively in a separate session as in previous studies (Ban, Preston, Meeson, & Welchman, 2012; Dövencioğlu, Ban, Schofield, & Welchman, 2013; Murphy, Ban, & Welchman, 2013). For nine of the fifteen participants, V3B/KO and hMT+/V5 were defined according to Talairach coordinates ([x,y,z] = [42, -81, 6] for right V3B/KO; [-42, -81, 6] for left V3B/KO; [42, -62, 6] for right hMT+/V5; [-42, -66, 2] for left hMT+/V5) (Orban, et al., 2003; Sunaert, van Hecke, Marchal, & Orban, 1999). LO and pFs were defined by a localizer scan for all participants in which intact object images and their spatially-scrambled versions were contrasted. pFs was identified as the more anterior portion of the activation map obtained from this contrast. The average mass centre of LO and pFs across the 15 participants were [39, -63, -8] and [31, -43, -14] for right and [-40, -67, -4] and [-38 -48, -14] for left hemisphere. The superior temporal sulcus (STS) was defined according to Talairach coordinates ([57, -45, 10] for right and [-57, -45, 10] for left in superior temporal sulcus) (Sunaert et al., 1999). The detailed information about each localizer and ROI definition can be found in Appendix 1.

Additional fMRI analysis. We computed percent signal change (PSC) by subtracting the BOLD signal baseline (the average signal in fixation blocks) from each experimental condition and then dividing by the baseline. In addition, voxels used in the PSC analysis were masked with the *t*-value maps obtained by contrasting all stimulus conditions *vs*. fixation blocks for each individual participant. PSCs were examined within independently identified ROI under each experiment condition. We then computed the difference in PSC between intact and scrambled versions of Glossy and Matte objects, which we term ΔPSC.

Finally, we used random effects Granger causality mapping (RFX GCM) to probe the information flow between ROIs. Granger causality uses temporal precedence to identify the

direction of influence from a reference region to all other brain voxels (Tyler, Likova, Kontsevich, & Wade, 2006). The GCMs for each participant were calculated along the entire time course of each run (time lag = TR, 2s)then they were combined together with a simple t-test ($t > 0$) and cluster-size thresholding (25 mm^2). We computed GCMs with entire time course because we aimed to examine the difference of information flow between seed areas instead of between conditions. Choosing a specific condition (e.g. Glossy) may cause the problem that the difference of information flow between seed areas may be due to the processing of some low-level confounding (e.g. luminance, contrast) in that condition. Using all the four conditions may reduce this confounding (theoretically).

2.3 Results

To identify brain areas that preferentially responded to glossy objects, we used a conjunction analysis to find voxels that were activated more strongly in Glossy condition than in any of the other three conditions across the 15 participants. In particular, **Figure 4** shows the results of a random-effects GLM with statistical significant voxels ($p<.05$) and cluster-size thresholding (25 mm^2). The orange areas demark significantly higher activation in Glossy condition under the three contrasts, respectively: Glossy *vs*. Scrambled Glossy, Glossy *vs*. Matte, and Glossy *vs*. Scrambled Matte. The conjunction analysis of the three contrasts was the same as the first monkey fMRI study (Okazawa et al., 2012) which has similar design as ours so that we can compare with their results. In general, these areas were distributed along ventral visual pathway in both hemispheres including the ventral occipitotemporal cortex, consisting with the results of Okazawa et al (2012) who used phase scrambling instead of spatial scrambling as the controls. In addition, we found responses in the area around V3B/KO, which is traditionally thought to belong to the dorsal visual stream.

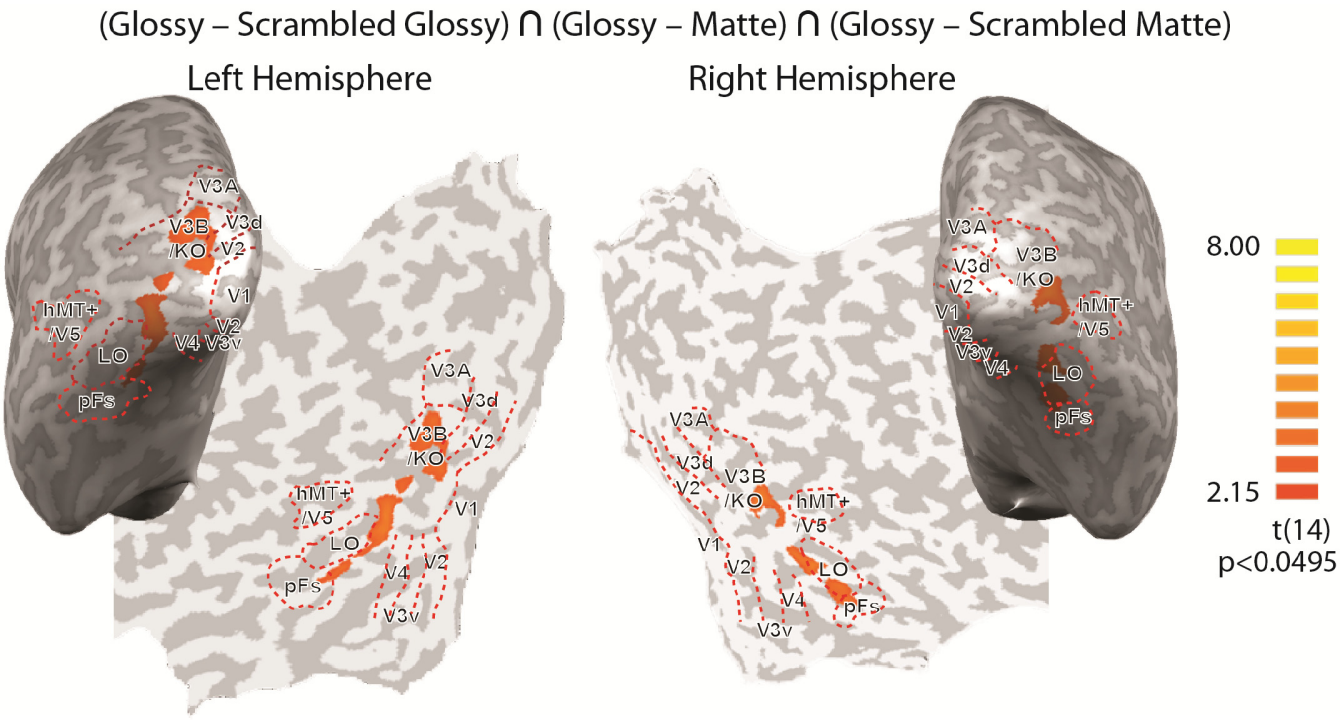


Figure 4. The result of conjunction analysis across the 15 participants. The Glossy condition was compared with the other three conditions. Significant conjunctions are presented on representative flat maps. Sulci are shown in dark gray and gyri are in light gray. The colour scale indicates *t*-values. The significance level was $p<.05$ with cluster-size thresholding 25mm^2. The orange areas represent activation in Glossy condition that is significantly higher than any of the other three conditions respectively.

To complement our whole brain contrast analysis, we also examined the percent signal change (PSC) within independently identified regions of interest. To identify responses to global objects with consistent surface properties, we contrasted the glossy and matte stimuli against their scrambled controls by subtracting PSC in scrambled conditions from their intact counterparts for Glossy (light bars) and Matte (dark bars) conditions leading to ΔPSC (**Figure 5**). We first tested whether activation differed for scrambled stimuli and their intact counterparts by testing if the ΔPSC deviated from zero. In early (V1, V2) and intermediate (V3d, V3v, V4) visual areas, we found stronger responses to the scrambled

stimuli than their intact counterparts (single sample *t*-test, two-tailed, Bonferroni corrected, on ΔPSC averaged across Glossy and Matte conditions. V1: t_{14}=8.5, p<.001; V2: t_{14}=7.2, p<.001; V3d: t_{14}=3.7, p<.022; V3v: t_{14}=5.0, p<.001; V4: t_{14}=3.8, p<.022), indicating that globally incoherent stimuli drive higher levels of activity. By contrast, in higher visual areas V3A, V3B/KO, hMT+/V5, LO and pFs we found stronger responses for intact versions of the stimuli (V3A: t_{14}=4.0, p<.011; V3B/KO: t_{14}=3.4, p<.044; hMT+/V5: t_{14}=11.9, p<.001; LO: t_{14}=12.7, p<.001; pFs: t_{14}=5.9, p<.001). Response magnitudes in the STS were low, and not significantly different from zero (t_{14}=0.5, p=.604).

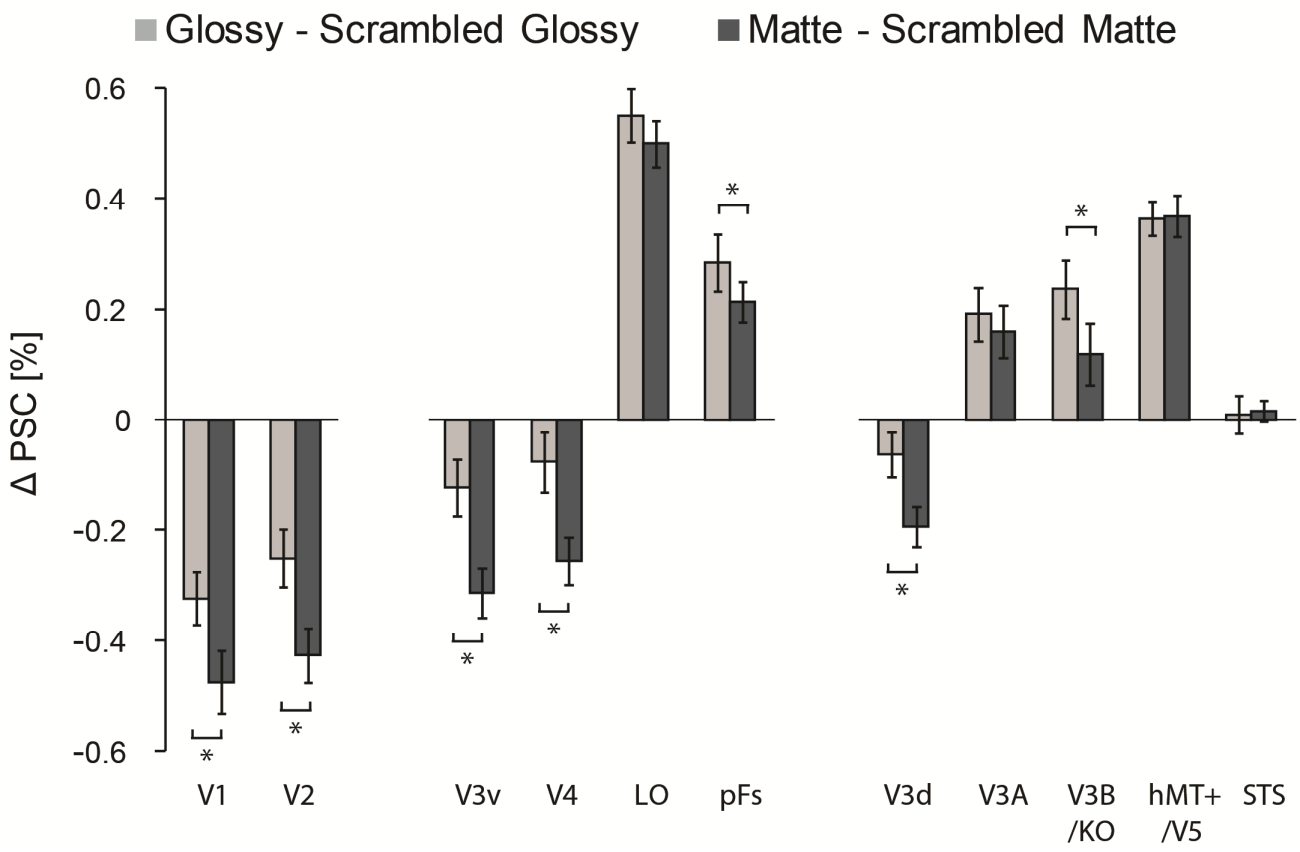


Figure 5. The ΔPSC for Glossy (light bars) and Matte (dark bars) conditions in all the ROIs. The ΔPSC in Glossy condition was calculated by [Glossy – Scrambled Glossy] PSCs and the ΔPSC in Matte condition was calculated by [Matte – Scrambled Matte] PSCs. The bars

reflect mean ΔPSC across 15 participants with ± 1 SEM. Asterisks represent significant difference between Glossy and Matte in ROIs (p<.01). The bars were arranges in three groups which represent the ROIs in early visual areas, ventral visual areas and dorsal visual areas respectively.

We then compared ΔPSC for Glossy (light bars) against Matte (dark bars) conditions (**Figure 5**) in all the ROIs. A two-way repeated measures ANOVA showed a significant difference between Glossy and Matte conditions ($F_{1,14}$=10.7, p=.006), an effect of ROI ($F_{10,140}$=102.5, p<.001), and a significant interaction ($F_{10,140}$=12.9, p<.001). Thereafter we tested for the differences between conditions in each ROI. Asterisks in Figure 5 represent significant differences in activation between the two conditions (Tukey's HSD post-hoc test at p < 0.01). We found that responses were significantly higher for objects with glossy than with matte surfaces in areas V3B/KO and pFs. Note that to compute ΔPSC we subtracted the activation in scrambled versions of the stimuli, so the glossy selectivity observed in V3B/KO and pFs is less likely to be explained by low-level differences in the images of the objects. Moreover, we found no significant difference in the percent signal change (PSC) between Scrambled Glossy and Scrambled Matte conditions (see **Figure 6**), suggesting that the significant differences in ΔPSC between glossy and matte stimuli were mainly due to the PSC difference between Glossy and Matte conditions rather than between their scrambled counterparts. ΔPSC in early visual areas (V1, V2, V3v, V3d, V4) were also significant, however response modulation in these areas was higher for scrambled stimuli than for intact ones. Since the PSC in Scrambled Glossy and Scrambled Matte conditions were similar (see **Figure 6**), we can conclude that the difference is mainly due to intact conditions. It is possible that some low-level difference still cannot be ruled out completely with spatial

scrambling procedure (e.g. spatial frequency difference in **Figure 2**; line segments and edges of highlights) and this residual effect caused the differences in ΔPSC between glossy and matte stimuli in earlier visual areas. It is also possible that some neurons in these areas selectively respond to glossy object (Okazawa, et al., 2012 and Wada, et al., 2014 also found the importance of V1-V4 in gloss processing), however, unlike V3B/KO and pFs, these areas respond prevalently to scrambled images rather than intact ones. This suggests that these areas primarily deal with low-level image features but also play a secondary role in gloss processing. As reviewed above, responses in STS were very low and not significantly different across conditions.

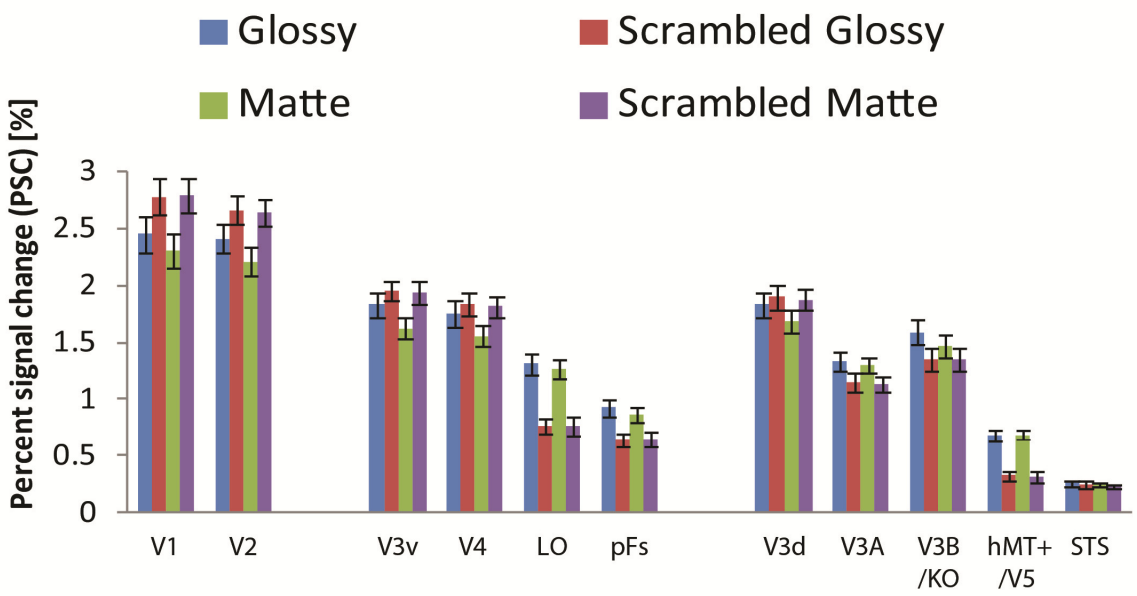


Figure 6. The PSC of Glossy condition (blue bar), Scrambled Glossy condition (red bar), Matte condition (green bar) and Scrambled Matte condition (purple bar). No significant difference was found between the two scrambled conditions in each ROI.

The preceding analysis indicates two brain areas (pFs and V3B/KO) that appear to be important in processing information about gloss. To quantify how these areas communicate

with other parts of the visual cortex, we used a random effects Granger causality mapping analysis (RFX GCM) to assess how these areas influence and depend on activity elsewhere. **Figure 7** shows the results using either pFs (**Figure 7A**) or V3B/KO (**Figure 7B**) as the reference region respectively. Blue areas indicate brain areas that are significantly influenced by the reference region, while the green colour map identifies locations that have a significant influence on the reference region. We found that activity in pFs had a strong influence on both dorsal and ventral areas. This may reveal that gloss-related activity is used for the processes of object processing (in ventral cortex) in addition to affecting depth estimates (estimated in dorsal areas). By contrast, the estimated connectivity in V3B/KO was quite different. V3B/KO mainly received information from ventral areas rather than having influence on them, perhaps indicating that gloss information in V3B/KO is inherited from a primary locus in ventral areas. In addition, we observed that V3B/KO also received some information from an area near the STS. Although our other analyses did not suggest the involvement of the STS, this analysis appears consistent with the role of the STS in gloss indicated by electrophysiological recordings (Nishio et al., 2012). We should note that we could not determine whether the information flow captured by the Granger Causality Mapping is specific to gloss signals. Nevertheless, as the preceding conjunction analysis and PSC results showed the importance of pFs and V3B/KO in processing gloss, it is quite possible that the GCMs show different information flows between pFs and V3B/KO for gloss processing.

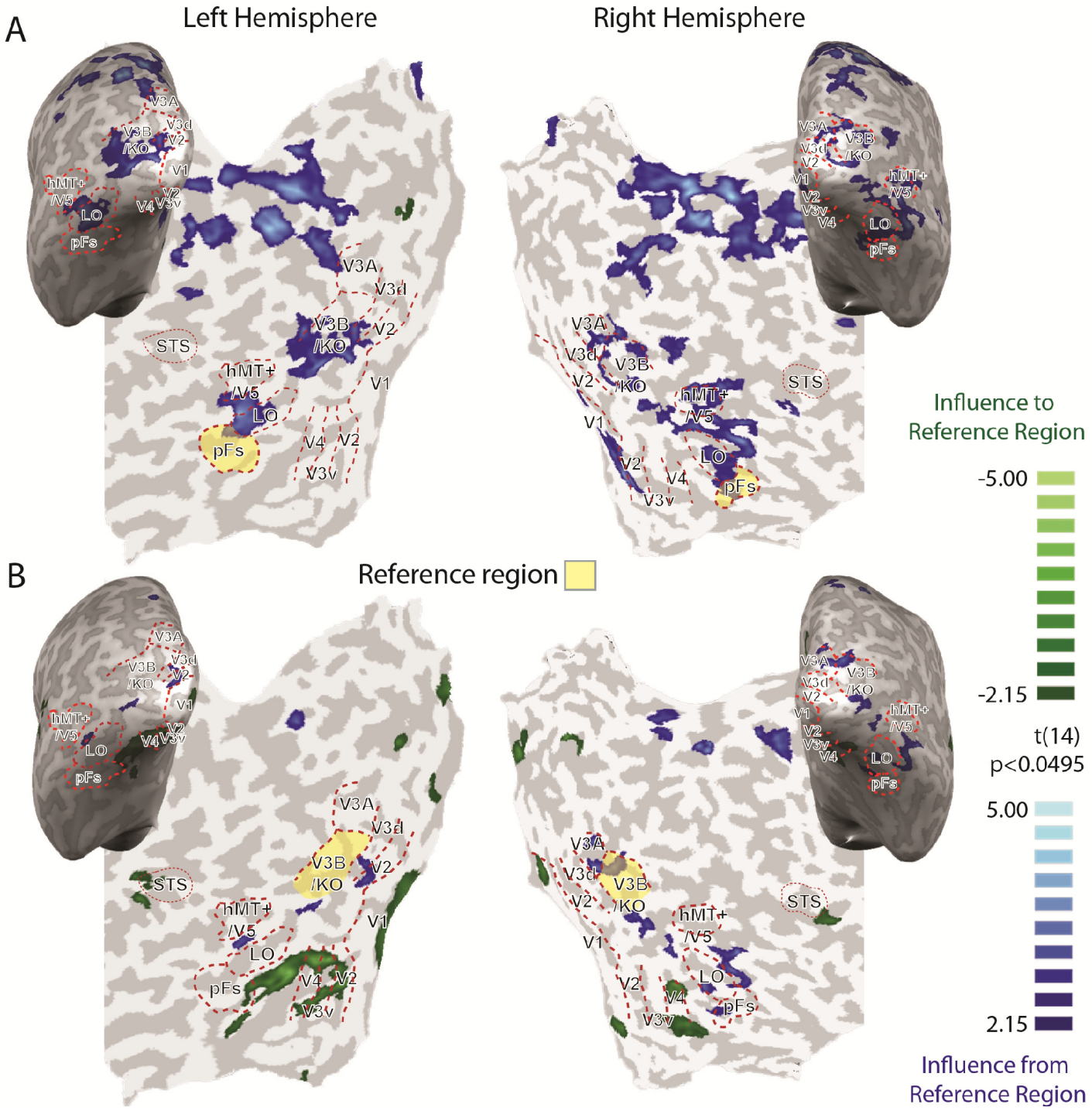

A
Left Hemisphere
Right Hemisphere
V3A
V3d
V3B
/KO
V2
V1
hMT+
/V5
LO
V4
V3v
pFs
STS
B
Reference region
Influence to
Reference Region
-5.00
-2.15
t(14)
p<0.0495
5.00
2.15
Influence from
Reference Region

Figure 7. RFX GCMs with (A) pFs and (B) V3B/KO as reference regions (yellow areas). Blue areas received significant influence from the reference region and green areas sent significant influence to the reference region ($p<.05$ for t-test on GCMs). Note that since the group GCMs were averaged across participants and then presented on representative flat maps, individual ROI boundaries may not perfectly fit the group level.

2.4 General discussion

The aim of this study was to localize the brain areas preferentially responding to glossy objects in the human brain. We did this by rendering glossy and matte versions of three-dimensional objects, and using scrambled images to control for low-level image cues. Our results point to a role for the posterior fusiform (pFs) and area V3B/KO in the processing of surface gloss: we found stronger responses to glossy objects than their matte counterparts, and this could not be explained by low-level stimulus differences. By assessing connectivity between brain areas while viewing glossy and matte stimuli, we observed that pFs exerted influence on ventral and dorsal brain areas, while V3B/KO was influenced by activity in midlevel ventral areas, which may imply a difference between areas in their use of information from gloss as a cue to material (pFs) vs. object shape (V3B/KO).

Recent imaging studies in macaques suggest that glossy objects elicit more activation along the ventral visual pathway form V1 to IT cortex (Okazawa, et al., 2012). We also found higher activation in the ventral stream, in particular in the pFs. Our results are reassuringly consistent with a very recent fMRI study that used a different image control approach (Wada et al., 2014). In particular, that study indicated the role of ventral areas and the combined areas V3A/B (which is very near to the V3B/KO that we identify). Since the ROI in our study were mapped using independent localisers before the experiment whereas Wada et al. considered only one area (V3A/B), our results pinpoint gloss-related activity more precisely, suggesting that the more lateral V3B/KO region is more important in gloss processing than V3A. The involvement of early visual areas (V1 to V4) is not clear. Although **Δ**PSC in earlier areas is significant due to higher activation for Glossy than for Matte objects (see Wada, et al., 2014; Okazawa, et al., 2012), however, unlike V3B/KO and pFs, response modulation in

these early areas is higher for scrambled stimuli. This suggests that these areas primarily deal with low-level features such as the area which occupies visual field, discontinued borders and high spatial frequency information which is more in scrambled than in intact conditions. Note that some low-level features might be affected by our scrambling technique. For example, there are more highlight boundaries (line segments and edges) on Glossy objects and scrambling decreases the number of these segments and edges. Thus, the PSC difference in V1 to V4 might be caused by such low-level image properties. It is also possible that these earlier areas also play a role in gloss processing as found in previous studies (Okazawa et al., 2012; Wada et al., 2014). Previous human fMRI studies found the modulation of fMRI responses by different object materials perception in the fusiform gyrus (FG) and collateral sulcus (CoS) (Cant et al., 2009; Cant & Goodale, 2007; Cavina-Pratesi et al., 2010a, 2010b; Hiramatsu et al., 2011). This work employed a wide variety of object materials (e.g., metal, wood, stone, glass) thus creating differences in surface gloss as well as differences in texture and colour. Here we focused on gloss, manipulating surface reflectance of untextured and homogeneously coloured objects. Despite this important difference between the studies, the surface-property-specific region (they denoted as CoS) they found is located very close to the area we denote as pFs based on a comparison of Talairach coordinates. Consistent with this, other work showed that a patient with colour and texture discrimination deficit could judge glossiness correctly, indicating that glossiness information does not exclusively depend on colour or texture processing (Kentridge et al., 2012). Taken together, this evidence suggests a dissociation between areas underlying material/texture from gloss. Nevertheless, the proximity of these areas may suggest a close interrelation and connection between material and gloss processing centres.

The role of V3B/KO in gloss processing

An important finding here is that the brain area V3B/KO seems to be involved in gloss processing. V3B/KO, located in dorsal visual stream, is well known to selectively respond to kinetic boundaries (Van Oostende, Sunaert, Van Hecke, Marchal, & Orban, 1997). It was also found to be involved in integrating different depth cues (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013; Tyler et al., 2006). Our study, together with the recent results by Wada et al. (2014), indicate that the activity in V3B/KO is modulated by surface gloss, although previous work has not highlighted the involvement of this area in processing material information. One possibility is that V3B/KO does not actually processes gloss information *per se*. The causality mapping suggests quite a different pattern of causal relationships in V3B/KO than in pFs, with V3B/KO primarily being influenced by signals from elsewhere, while pFs influences responses in other areas. It is possible that the effect we found in V3B/KO was due to the effect of adding internal boundaries to the shapes corresponding to the locations with highlights. Alternatively, because specular highlights are known to influence the perception of 3D shape (Blake & Bülthoff, 1990; Fleming, Torralba, & Adelson, 2004; Muryy, Welchman, Blake, & Fleming, 2013), it is possible that differences in activity in V3B/KO for glossy vs. matte objects relate to differences in the estimated 3D shape. This appears consistent with the recent work that indicates that V3B/KO integrates different cues to 3D structure (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013).

Human STS in gloss processing

The superior temporal sulcus (STS) of the macaque was found to show specific responses to glossy objects based on both fMRI (Okazawa et al., 2012) and single-unit recordings (Nishio et al., 2012). However, in our study we did not find strong evidence for

the involvement of human STS in glossiness processing: changes in signals in this area were low, although the causality mapping did indicate some modulation of activity near the STS. It is possible that there are functional differences between human brain and monkey brain. For example, studies found functional differences between the two species in V3A and the intraparietal cortex for three-dimensional structure-from-motion (3D-SFM) processing (Orban, 2011; Vanduffel, et al., 2002). It is also possible that the reasonably large voxel sizes used in our study limited our ability to detect responses to glossy stimuli in the human STS, and/or that the underlying population is spatially limited such that it did not survive the cluster threshold we applied.

The advantages of using mosaic spatial scrambling

In our study we chose to generate control stimuli using a scrambling technique applied to a visible grid. The presence of a grid reduces changes in low-level image properties due to scrambling (e.g., luminance, contrast, luminance histogram skew) while disrupting global properties of the shapes that are known to modulate the impression of gloss (Anderson & Kim, 2009; Kim et al., 2011; Marlow et al., 2011). The use of a superimposed grid over the stimuli was conceived to ensure that the amount of edge information in the stimuli was broadly similar between intact and scrambled conditions (**Figure 2**). This expedient overcomes the large difference in spatial frequency content that would be produced by scrambling alone (**Figure 2D**). Although there are slight differences in spatial frequency between intact objects and their scrambled counterparts (see **Figure 2D**), scrambling had similar effects for Glossy and Matte conditions. Therefore differences in the spatial frequency spectra could not be the only cause for the pattern of results found. Furthermore, image statistics (luminance, contrast, skew and spatial frequency) did not differ substantially

between Glossy and Matte conditions, ensuring that the results are not due to these properties as well. One could also argue that images with an overlaid grid could be amodally completed behind the occlusions. Such completion would be present for intact objects in both Glossy and Matte conditions. Therefore the completion-related activity would not bias the results. Similarly, even though scrambling clearly makes the stimuli occupy a larger portion of the visual field (**Figure 1**), our analysis procedures makes it unlikely that such differences contributed to the findings we report in the study. This is because our conjunction analyses were not based only on [Glossy vs. Glossy Scrambled] and on [Matte vs. Matte Scrambled] comparisons, but also on the contrast [Glossy vs. Matte]. Overall, the results we presented cannot be explained by local edges, contrast, or configuration changes as these factors were the same for Glossy and Matte conditions.

We should also note that during our experiments our participants were not making active perceptual judgments of gloss. It is possible that activations would have been stronger had we asked for concurrent perceptual judgments. However, this would likely have introduced attention-based differences between the intact and scrambled conditions, which we deliberately sought to avoid using a task at the fixation point.

Finally, it is interesting to consider whether the areas we identify here would be involved in other aspects of gloss processing. As discussed in the Introduction, gloss perception can be modulated by several factors including low-level image cues (i.e. low-level image statistics), image configurations (such as the position and orientation of highlights), scene variables including light source direction (Marlow & Anderson, 2013; Marlow et al., 2012; Wijntjes & Pont, 2010), light source style (Dror et al., 2004; Fleming et al., 2003; Marlow & Anderson, 2013; Olkkonen & Brainard, 2010; Te Pas, Pont, & van der Kooij, 2010) and background colour (Doerschner, Maloney, et al., 2010). Moreover, factors related to 3D structure from self motion and object motion (Doerschner et al., 2011; Sakano & Ando,

2010; Uehara, et al., 2013; Wendt et al., 2010) and stereo viewing (Sakano & Ando, 2010; Sawabe, Yamamoto, Nakaguchi, Yamauchi, & Tsumura, 2010) can change perceived gloss. Finally, even non-visual sources such as haptic cues (Kerrigan, Adams, & Graf, 2010) and interactions with objects (Scheller Lichtenauer et al., 2013) can lead to changes in surface appearance. It is an open challenge to understand whether these variables involve processing in pFs and V3B/KO, or whether additional areas are recruited.

In summary, This study reveals that V3B/KO and pFs are selectively active when processing images of glossy objects. This finding is consistent with other recent human fMRI studies and it suggests close networks for gloss and material processing in the ventral stream. Our results imply a different role of V3B/KO and pFs, suggesting that V3B/KO may be tuned to processing highlight boundaries or 3D shape properties rather than to glossiness processing. Overall, our study highlights a small network in the fusiform that may be important in supporting our perception of surface gloss.

Research limitations

The role of V3B/KO in gloss processing is still not clear in the current study which may confound with the processing of highlight boundaries and 3D shape and need to clarify with different manipulations in further studies. In this study we used spatial scrambling to control for the confounding of low-level image features, however, the effect of low-level factors still cannot be ruled out completely (though largely reduced). Also, the superimposed grids may also affect (impair) glossiness and object processing since they are not supposed to present in normal viewing conditions. The role of human STS is not clear from this study due to low spatial resolution of functional scans (2.5 × 2.5 × 3 mm voxel size) which cannot

detect subtle activations. Time resolution (TR = 2) of functional scans is also not perfect for GCM analysis which may not be able to detect important information flow.

Afterword

In this study we found two crucial areas for the processing of surface gloss: posterior fusiform (pFs) and V3B/KO. Our results are consistent with a very recent fMRI study that used a different image control approach (Wada et al., 2014). Particularly, our results pointed out the locus more precisely than before that in ventral stream pFs is more selective to glossy objects than other areas and in dorsal stream V3B/KO is more selective than V3A.

The main purpose of this study was to localize brain areas preferentially responding to glossy objects. Therefore it is reasonable to assume that these areas may have stronger activation for glossy than matte stimuli with GLM (and also PSC) analysis. Recent human and monkey fMRI studies also used the same approach to identify selective neural response to glossiness (Okazawa, Goda, & Komatsu, 2012; Wada, Sakano, & Ando, 2014) and we can compare with their results directly. However, some groups of neurons which involves in glossiness processing may respond in the other way (preferentially responding to matte objects than glossy objects) as founded in monkey electrophysiological studies (Nishio, Goda, & Komatsu, 2012; Nishio, Shimokawa, Goda, & Komatsu, 2014). We therefore used MVPA for the following studies which are able to decode brain activations that respond to glossy and matte stimuli in different ways.In this study we used a different approach than previous material studies in defining ventral ROIs. We used a functional localizer to define LO and pFs instead of defining CoS by anatomical structures because there might be some individual difference in brain function-structure correspondence. Sometimes a brain structure may involve a variety of functional processing and sometimes a specific functional processing may involve an area that covers different brain structures. Therefore, functional ROIs might be a better approach in examining the issues about material processing.

To compare with previous studies, CoS was defined based on anatomical structure (Weiner, et al., 2014) and was compared with functionally defined pFs (Appendix 2). As shown in Appendix 2, pFs we defined here is separated from CoS for most of the participants. pFs mainly located around fusiform gyrus (FG), mid-fusiform sulcus (MFS), occipito-temporal sulcus (OTS) according to Weiner, et al. (2014). It seldom overlaps with CoS. The PSC result showed that CoS did not have stronger responses to glossy objects than their matte counterparts and scrambled counterparts as pFs did (Appendix 3). In addition, MVPA also showed that in CoS the decoding performance for Glossy condition vs. Matte condition was nearly chance (52.38% prediction accuracy). Therefore we did not include CoS for the following studies.

Chapter 3:

Look but don't touch: visual information about surfaces decoded in somatosensory cortex

This chapter has been submitted as a journal article:

Sun, H.-C., Welchman, A. E., Chang, D. H. F & Di Luca, M. (submitted). Look but don't touch: visual information about surfaces decoded in somatosensory cortex.

Foreword

In the previous chapter we focus on the visual areas that were important for glossiness processing (pFs and V3B/KO). In this chapter we focus on not only visual areas but also somatosensory areas, to investigate the neural basis for tactile predictions from visual stimulation. We analyse brain activation elicited by objects in four conditions with different visual appearances: glossy, glossy control, rough, or rough control. Glossy objects are similar as those in the last study with specular highlights on them. We contrasted glossy objects with rough objects because they have very different sense of touch, and we also contrasted between their controls for comparison because they have the same sense of touch (all matte) with similar low-level image features as their original counterparts. We performed multi-voxel pattern analysis (MVPA), seeking to detect subtle pattern activation for different visual material information in somatosensory areas.

Abstract

To plan our interactions with nearby objects, our brains use visual information to estimate shape, material composition, and surface structure *before* we come into contact with them. For example, an object with spiky surface (i.e., a cactus) might cause us to predict a rough (and painful) sensory experience, leading to great care when planning physical contact with it. The neural basis for tactile predictions from visual stimulation is unclear. Here we analyse brain activation elicited by different visual appearances, measuring fMRI responses to objects that are: glossy, matte, rough, or textured. In addition to activation in visual areas, we found fMRI responses in the secondary somatosensory area (S2) that are evoked by looking at glossy and rough surfaces. This area activates differently for tactile-related visual properties (gloss, rough, and matte) while other visual properties (i.e., coloured texture) do not substantially change activity. Moreover, control measurements in which participants were explicitly asked to imagine different sensations were not sufficient to evoke activity in this area. These findings suggest that visual cues to an object's surface properties evoke activity in neural circuits associated with tactile stimulation in the somatosensory cortex in a relatively automatic way. This indicates a crossmodal network that may be responsible for predicting the likely consequences of different actions. This pre-activation can be seen as the a-priori probability of the physics of the interaction (i.e., the expectation of what friction to expect) that can be used for planning finger placement and grasping force.

3.1 Introduction

When we look at objects, we are able to predict how they will feel once we come into contact with them (Fleming, 2014; Xiao, Jia, & Adelson, 2013). For instance, shiny objects with *glossy* surfaces, like silverware and plastic, are expected to feel smooth and hard when pressed, and sliding our fingers over their surface may generate stick-slip interactions. Textured objects, like a tree bark and sandpaper, are expected to feel rough when pressed upon and can lead to abrasion if stroked. Matte objects, like wood and stone, are expected to feel irregular and can generate skin vibration if caressed. These expectations refine movement planning: e.g., slippery objects necessitate a more precise and powerful grip.

Studies of visual appearance perception have considered responses mainly in visual areas. All of these studies have consistently shown an importance of ventral visual areas in processing appearance information such as surface texture and categories of different materials (Cant et al., 2009; Cant & Goodale, 2007; Cavina-Pratesi et al., 2010a, 2010b; Goda et al., 2014; Hiramatsu et al., 2011). A recent human fMRI study showed that visual information about object category is also represented in primary and secondary somatosensory cortex (S1 and S2), in addition to the typical visual areas. They found that activation patterns in S1 and S2 could distinguish between three familiar visual object categories that most people had haptic experience with: wine glasses, mobile phones and apples (Smith & Goodale, 2015). Interestingly, S1 and S2 could not discriminate unfamiliar objects. The authors therefore proposed cross-modal connections may exist from visual areas to early somatosensory areas that transmit information about familiar object categories that contain content-specific information. This suggests that when looking at familiar objects, we may generate expectations of tactile sensation from memories which might promote a more

successful action planning with them. However, we are also able to generate expectations of tactile sensation and perform successful movement even when looking at novel objects that we see at the first time or objects that we have difficulties to identify or categorize (e.g. partially occluded objects or under weak illumination). It is still not clear whether the brain has a system to process this from objects' visual appearance directly without reorganization and categorization.

Recent studies have suggested a candidate area that may provide information about tactile sensation of objects' surface. The Parietal operculum (PO), a subregion of S2 has been shown previously to be important for tactile texture and roughness perception (Kaas et al., 2013; Kitada et al., 2005; Pruett et al., 2000; Roland et al., 1998; Sathian et al., 2011; Simões-Franklin et al., 2011; Stilla & Sathian, 2008). A human fMRI study found larger BOLD signal change in somatosensory cortex when presenting stimuli bimodally (vision and haptic) as compared to haptic only or visual only presentations (Eck, Kaas, & Goebel, 2013). However, other human fMRI studies did not find evidence for somatosensory activation in visual texture discrimination (Sathian et al., 2011; Stilla & Sathian, 2008). It is still not clear whether visual information about surface texture can be represented in somatosensory cortex.

Studies about haptic texture-selective areas found in visual cortex may provide the possibility that visual texture-selective areas may be found in somatosensory cortex. A human fMRI study showed that object-sensitive regions in occipitotemporal cortex (OTC, which includes the lateral occipital region (LO) and posterior fusiform (pFs)) represents information about object weight in advance of lifting that object. Moreover, when different object texture specified different object weight, this particular texture-weight association could also be represented in OTC through learning (Gallivan, Cant, Goodale, & Flanagan, 2014). Another human fMRI study found haptic object-selective activity in OTC (Amedi,

Malach, Hendler, Peled, & Zohary, 2001). Moreover, studies found haptic texture-selective area in the middle occipital cortex (MOC) and haptic shape- and location-selective areas in intraparietal sulcus (IPS) (Sathian et al., 2011; Stilla & Sathian, 2008). These results suggest that visual area OTC, MOC and IPS are actually visual-haptic bimodal areas which represent a variety of haptic information as well. These pieces of evidence increase the possibility that visual texture-selective areas may exists in somatosensory cortex as visual cortex and somatosensory cortex have similar hierarchy to some extent (Felleman & Van Essen, 1991).

Here we ask whether there are conditions in which there may be somatosensory involvement when viewing novel visual objects. In particular, we ask whether visual surface information *per se* can affect somatosensory activation. We measured human fMRI responses to visual images of computer-generated objects that had perceptually different surface characteristics. The stimuli were designed to evoke a visual impression of surface gloss or roughness, while the control conditions were designed to depict stimuli with similar image statistics that nevertheless gave rise to a different impression of surface properties. All the stimuli were novel objects to avoid issues of remembered sensations. We used multivoxel pattern analysis (MVPA) to test for visual and somatosensory areas that contained neuronal responses that supported reliable discrimination of different visual surface characteristics. The rationale is that if the brain has a system to generate expectations of tactile sensation when looking at objects with different surface properties, this difference should elicit different activation responses in somatosensory cortex. In addition to somatosensory cortex, we also looked into earlier visual areas (V1-V4), ventral visual areas (LO and pFs) and dorsal area V3A and V3B/KO as they were found important for processing of material-related information in the last chapter and also in other studies. We removed hMT+/V5 and STS from data analysis since they are not very relevant to the processing of material-related

information from the last chapter. We also included a control experiment to assess whether imagination of these objects can generate similar somatosensory activations.

3.2 Methods

3.2.1 Participants

Sixteen participants who had normal or adjusted-to-normal vision were recruited for this experiment. One was the author H.-C. S. and the remaining participants were naïve to the tasks and purpose of the study. All were screened for visual acuity and MRI safety before being invited to participate. The age range was 18 to 39 years old, and 5 of the 16 participants were male. All participants gave written informed consent before taking part in the experiment. The study was approved by the STEM Ethical Review Committee of the University of Birmingham. After completing the experiment, all participants (except the author) received monetary compensation or credits.

3.2.2 Apparatus and Stimuli

Stimuli. The study comprised three 3-D shaped objects generated by Blender 2.67a selected from a previous study (Sun, Ban, Di Luca, & Welchman, 2015) (**Figure 1A**). Stimuli were 12 deg in diameter on average, and they were presented on a mid-gray background. We created versions of the stimuli for each object that made up the four conditions of the experiment: Glossy, Glossy Control, Rough and Rough Control (**Figure 1B**). In the Glossy condition, objects were rendered using a mixed shader with 90% diffuse and 10% glossy components. In Glossy Control condition, the specular components rendered on Glossy objects were selected and were rotated by 45 degree in the image plane, which made the objects look matte since the important contextual information for gloss perception has been destroyed (Anderson & Kim, 2009; Kim et al., 2011; Marlow et al., 2011). In the Rough

condition, wave textures were applied to objects' 3-D geometry, resulting in bumps on the surface. The distortion parameter of wave textures was set to 30. In the Rough Control condition, the same wave textures were applied to the objects' surface, resulting in a painted texture on the surface. For each of the three objects there were five different levels. In Glossy and Glossy Control condition, there were five levels of the emission strength from the light source: 1, 1.2, 1.4, 1.6 and 1.8 (**Figure 1C**). In Rough and Rough Control condition, there were five levels of wave texture scale: 12, 17, 22, 27 and 32 (**Figure 1D**). The five levels of each object were presented in a random order to reduce adaptation of the fMRI response. A black fixation dot (dia = 0.5 deg) was shown during fixation blocks.

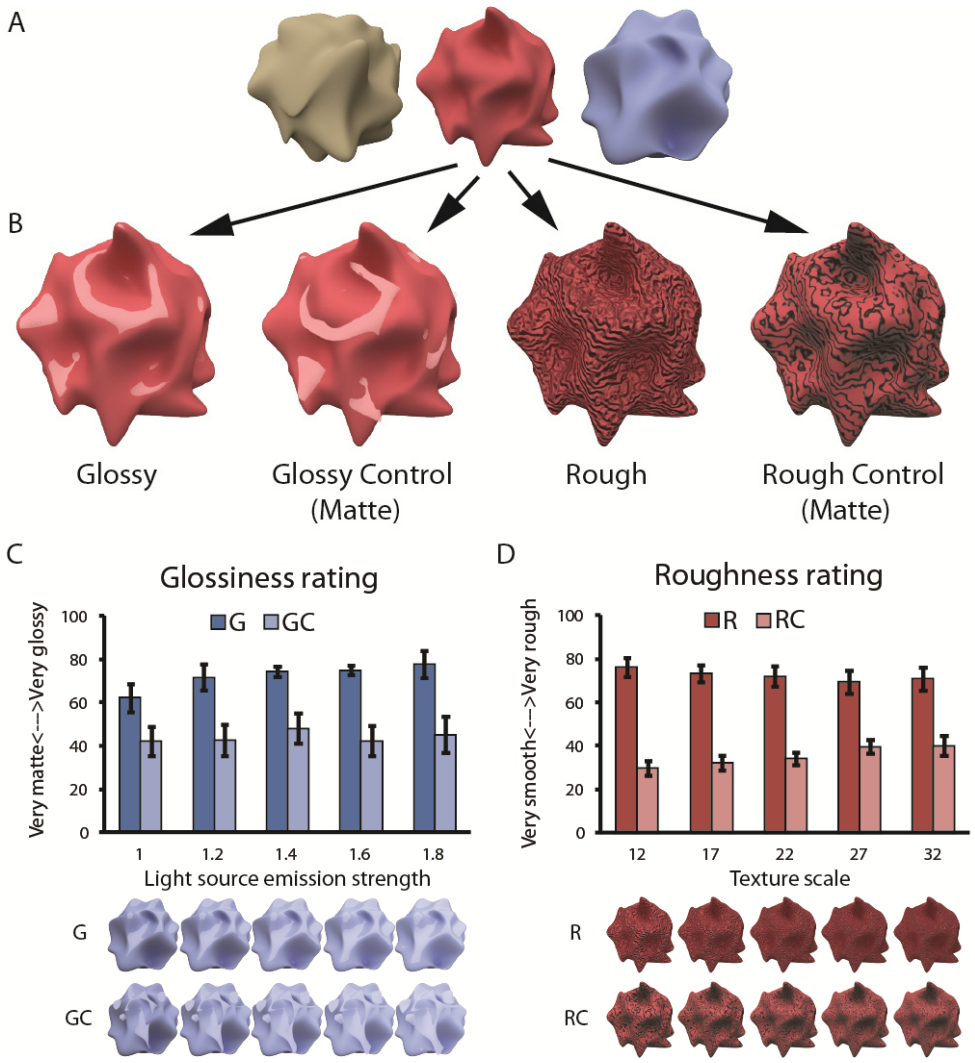


Figure 1. (A) The three objects used to create the stimuli shown to participants. (B) The four experimental conditions: Glossy, Glossy Control, Rough and Rough Control rendered on an example object. Specular components were shown in Glossy condition (90% diffuse and 10% glossy components) while the highlight areas were rotated and moved in the control condition to break the impression of surface gloss. In the Rough condition, wave textures were applied on objects' 3-D geometry; in the Rough Control condition, the same wave textures were applied to the reflectance of the surface to create the impression of a smooth surface with a painted texture. Glossiness rating and rough rating results are presented respectively in (C) and (D) under five levels of the emission strength from the light source and five levels of wave texture scale. The bars reflect mean rating scores across 7 participants with ± 1 SEM.

In the control experiment, 12 new objects were presented to participants in familiarization session before entering the scanner. The 12 objects were assigned to 4 groups for the four conditions in which the objects were rendered with the same color and surface properties within group. Then participants were presented with objects contours from the same objects filled with homogeneous color and were asked to imagine the surface properties of the four conditions. The color-coding of Glossy/Glossy control and Rough/Rough control was counterbalanced across participants. Participants were trained to associate the color cues with the four conditions and were able to make color-condition associations with 100% accuracy prior to entering the scanner (and upon re-test after the scan). During the scan, there were five levels of luminance scale for each object contour presented in a random order to avoid adaption effect in fMRI as in the main experiment.

Apparatus. The same apparatus were used as described in our previous paper (Sun et al., 2015). Psychtoolbox (Kleiner et al., 2007) was used for stimulus presentation. A JVC DILA SX21 projector was used for projecting stimuli on a translucent screen inside the bore of the magnet. Participants viewed stimuli via a mirror fixed on the head coil with a viewing distance of 64 cm. Luminance outputs were linearized and equated for the RGB channels separately with colorimeter measurements. A five-button optic fiber button box was provided for participants to make response when performing 1-back task in the scanner.

MRI data acquisition. A 3-Tesla Philips scanner and an 8-channel phase-array head coil were used to obtain all MRI images at the Birmingham University Imaging Center (BUIC). T1-weighted high-resolution anatomical scans (175 slices, TR 8.4 ms, TE 3.8 ms, flip angle 8 deg, voxel size: 1 mm^3) were obtained for each participant. Functional whole brain scans with echo-planar imaging (EPI) sequence (32 slices, TR 2000 ms, TE 35 ms, voxel size 2.5 × 2.5 × 3 mm, flip angle 80 deg, matrix size 96 × 94) were also obtained for

each participant. The EPI images were acquired in an ascending interleaved order for all participants.

3.2.3 Design and Procedure

Subjective rating task. A group of 7 naïve participants were recruited for the rating experiment before the fMRI session. Participants were instructed to perform a rating task which lasting for about 30 minutes. Participants performed glossiness ratings on all Glossy and Glossy Control stimuli in one block and roughness rating on all the Rough and Rough Control stimuli in another block. The order of the two blocks was balanced across participants. Participants viewed stimuli presented on a CRT monitor with a viewing distance of 83 cm. Luminance outputs were linearized and equated for the RGB channels separately with colorimeter measurements. The diameter of the stimuli was 12 deg. Each image was presented for 500 ms after which participants were given unlimited time to rate the image along a scale of 'very glossy' to 'very matte' for glossiness rating block, or along a scale of 'very rough' to 'very smooth' in the roughness rating block. Participants were permitted to place their rating bar in any position between the two ends to indicate their rating and the rating value was calculated by computing the distance between the bar and one end divided by the whole scale length.

fMRI Session. A block design was used. Each participant took part in 9 to 10 runs with 368 s length of each run in a 1.5 hour session. Each run started with four dummy scans to prevent startup magnetization transients and consisted of 16 experimental blocks each lasting 16 s. There were 4 block types (i.e., one for each condition), repeated four times in a run. During each block, fifteen objects were presented once in a pseudo-random order and one of them was showed twice (the 'event' that participants had to respond). Stimuli were

presented for 500 ms with 500 ms interstimulus interval (ISI). Participants were instructed to maintain fixation and perform a 1-back matching task, whereby they pressed a button if the same image was presented twice in a row. They were able to perform this task well (mean d'=2.07; SEM=0.10). Five 16 s fixation blocks were interposed after the third, fifth, eighth, eleventh and thirteenth stimulus blocks to measure fMRI signal baseline. In addition, 16 s fixation blocks were interposed at the beginning and at the end of the scan, making a total of seven fixation blocks during one experimental run. An illustration of the scan procedure is provided in **Figure 2**.

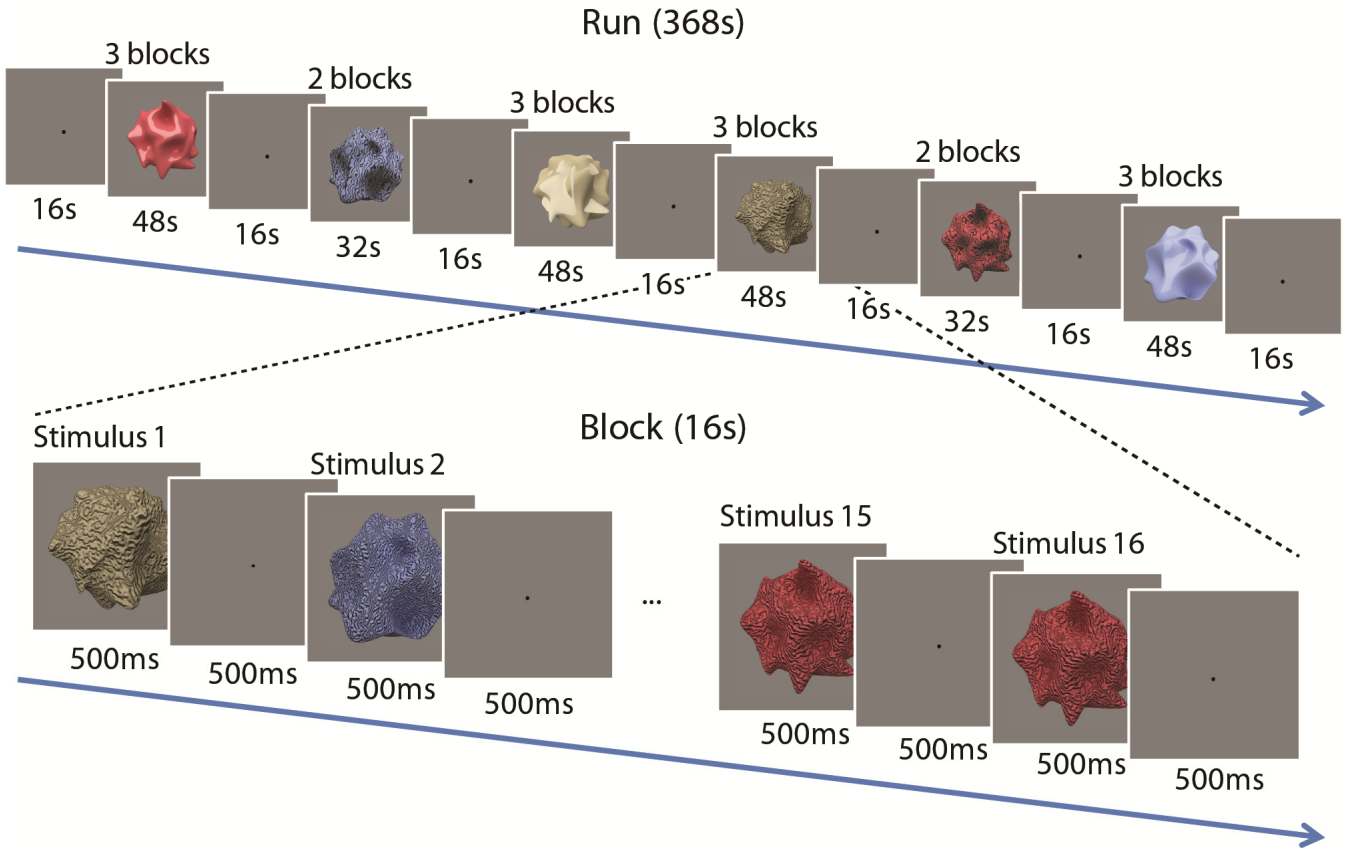


Figure 2. The procedure of one scan run. In each run there were 23 blocks (16 s each), including 7 fixation blocks and 16 experimental blocks. During each experimental block, stimuli were presented for 500 ms with interstimulus interval (ISI) 500 ms. Participants were instructed to perform a 1-back matching task during the scan.

In the control experiment the same block design was applied except that before the start of each block a color dot (the same colour as the object contours present next) was presented for 4 s for cuing the surface property that should be imagined. Participants were also instructed to perform a 1-back matching task while imaging surface property. They were able to perform this task as well (mean d'=2.48; SEM=0.16) as in the main experiment.

3.2.4 Data analysis

Functional MRI data processing. The basic data processing procedures for both the main and control experiment were the same as in our previous study (Sun et al., 2015). We also computed global signal variances of blood oxygenation level dependent (BOLD) signal as before and removed the scan runs which exceeded 0.23% of global signal variances. This criterion is a bit different from that in Study 1 (0.16%) since most of the participants were inexperienced and the global signal variances were higher than that in Study 1 in general. We therefore used a less strict criterion to retain basic number of runs for MVPA analysis. Seventeen runs out of 153 runs across 16 participants in the main experiment and 2 runs of 60 runs across 6 participants in the control experiment were excluded from further analysis based on this criterion.

ROI definition. Regions of interest (ROIs) of V1, V2, V3d, V3A, V3v, V4 V3B/KO (kinetic occipital region), LO (lateral occipital region) and pFs (posterior fusiform) were defined by localizers a separate session as in previous studies (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013; Sun et al., 2015). Somatosensory areas were defined by a somatosensory localizer. This localizer run contains 20 blocks, 10 air-on blocks and 10 air-

off blocks which showed alternately, each lasting 16s. In the air-on blocks, air puffs were sent to participants' ten finger tips by an air compressor (JUN-AIR) through plastic tubes (6 mm inner diameter) in 1-s on 1-s off cycles. No air was sent in air-off blocks. Somatosensory areas were defined by contrasting activations in air-on blocks with air-off blocks. Primary somatosensory area (S1) was defined with more dorsal/upper portion of the activations and secondary somatosensory area (S2) was defined with more ventral/lower portion of the activations based on anatomical grounds. The detailed information about each localizer and ROI definition can be found in Appendix 1. All the ROIs were defined by these independent localisers shown in **Figure 3.**

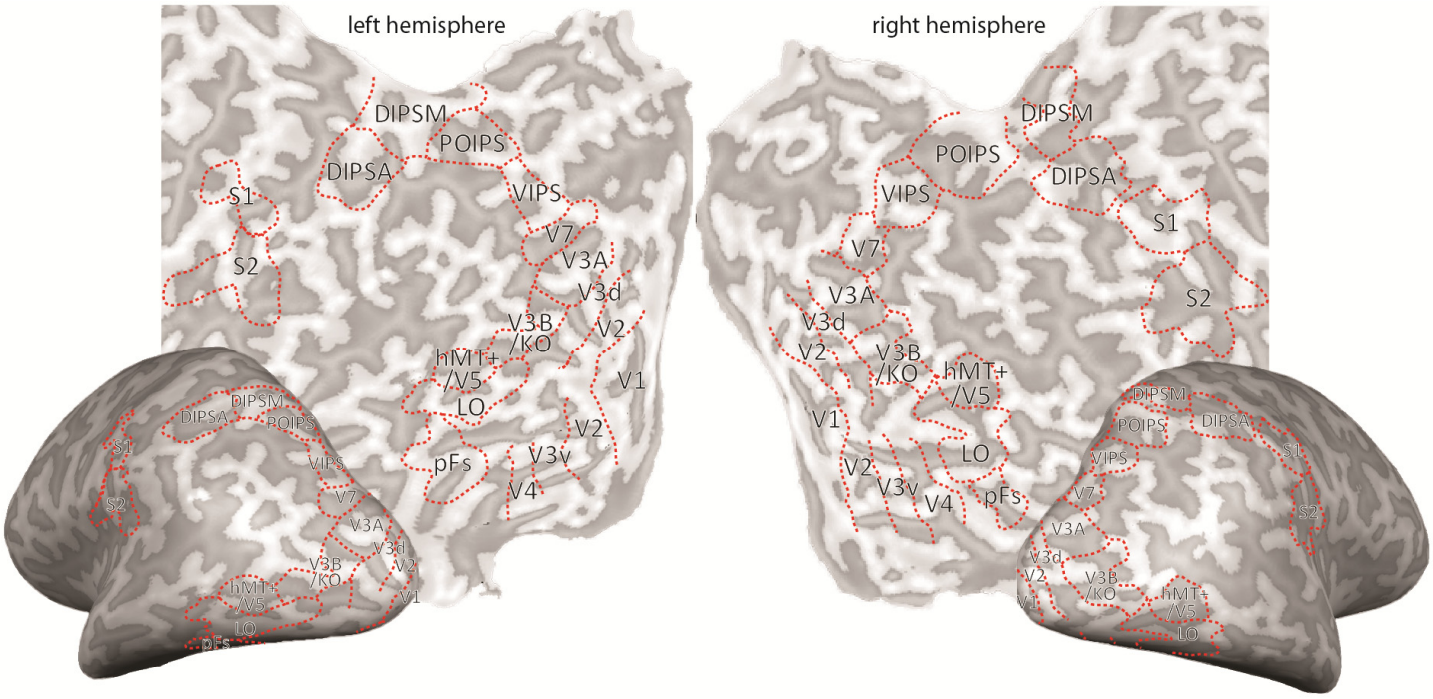


Figure 3. Red dotted lines are the ROI boundaries we defined with independent localizers for a representative participant, which include both visual and somatosensory cortex.

fMRI analysis. We used multivoxel pattern analysis (MVPA) to compute classification accuracies for different experimental conditions. For voxel selection, all voxels in each ROI were arranged with *t* value larger than 0 for the contrast of "all experiment conditions *vs.* fixation block" voxels in GLM t-value maps. The top 250 voxels were selected for classifications across ROIs. If a participant had fewer than 250 voxels in one ROI, we used the maximum number of voxels that had *t* values greater than 0. Less than 6% of ROIs across all participants had fewer than 100 voxels used in MVPA. On average each ROI has more than 221 available voxels for MVPA except S1 and S2, which had an average of 178 and 186 available voxels. After selecting the voxels, their time series were extracted and converted to z-scores. Then, the voxel-by-voxel signal magnitudes for a stimulus condition were obtained by averaging the signals over 8 time points (TRs) (= 1 block) separately for each scanning run. Before averaging, the time series were shifted 4 s to account for the hemodynamics response delay. The global baseline differences of these response patterns across the stimulus conditions and scanning runs were excluded by subtracting the mean of the patterns. These block-averaged signals were used as a response pattern in a ROI for the classification analysis. We used a linear Support Vector Machine (SVM) to discriminate between activities evoked by the different conditions in each ROI. In the training phase, 32 response patterns for each of the stimulus conditions were used as a training dataset for those participants that completed 9 runs and 36 response patterns were used for those who completed 10 runs. Then, 4 response patterns for each condition were classified by the trained classifier in the test phase. These training/test sessions were repeated and validated by a leave-one-run-out cross-validation procedure. The prediction accuracies were defined as a mean of these cross-validation classifications. The mean accuracies across participants were then tested against shuffled baseline with Bonferroni corrected, one-tailed single-sample *t*-

test, to check whether the accuracies are significantly above chance level (0.5 for all classifications in this paper as they are all binary classifications). Shuffled baselines were calculated with Permutation tests (1,000 repetitions for each ROI of each participant with randomly shuffling stimulus condition labels per test. The one-tailed, upper 95th percentile boundaries of accuracy distributions were averaged across all ROIs).

3.3 Results

To ensure that the stimuli we designed evoked a visual impression of surface gloss or roughness as we expected, we asked participants to rate the apparently glossiness (**Figure 1C**) or roughness (**Figure 1D**) of the objects. We found that the control versions of the stimuli were effective in reducing the appearance of gloss/roughness. In particular, objects in the Glossy condition were perceived to be significantly glossier than those in the Glossy Control condition (Wilcoxon signed-rank test, two-tailed, Z=–2.2, p<.05) and the objects in Rough condition were rated as significantly rougher than those in Rougher Control condition (Wilcoxon signed-rank test, two-tailed, Z=–2.4, p<.05). Moreover, to ensure that glossy and rough objects have similar image statistics of pixelwise luminance, contrast, histogram skew and power spectra with their controls despite different impression of surface properties, we quantified image statistics across conditions, finding that the control versions of the stimuli were well matched to their counterparts (**Figure 4**).

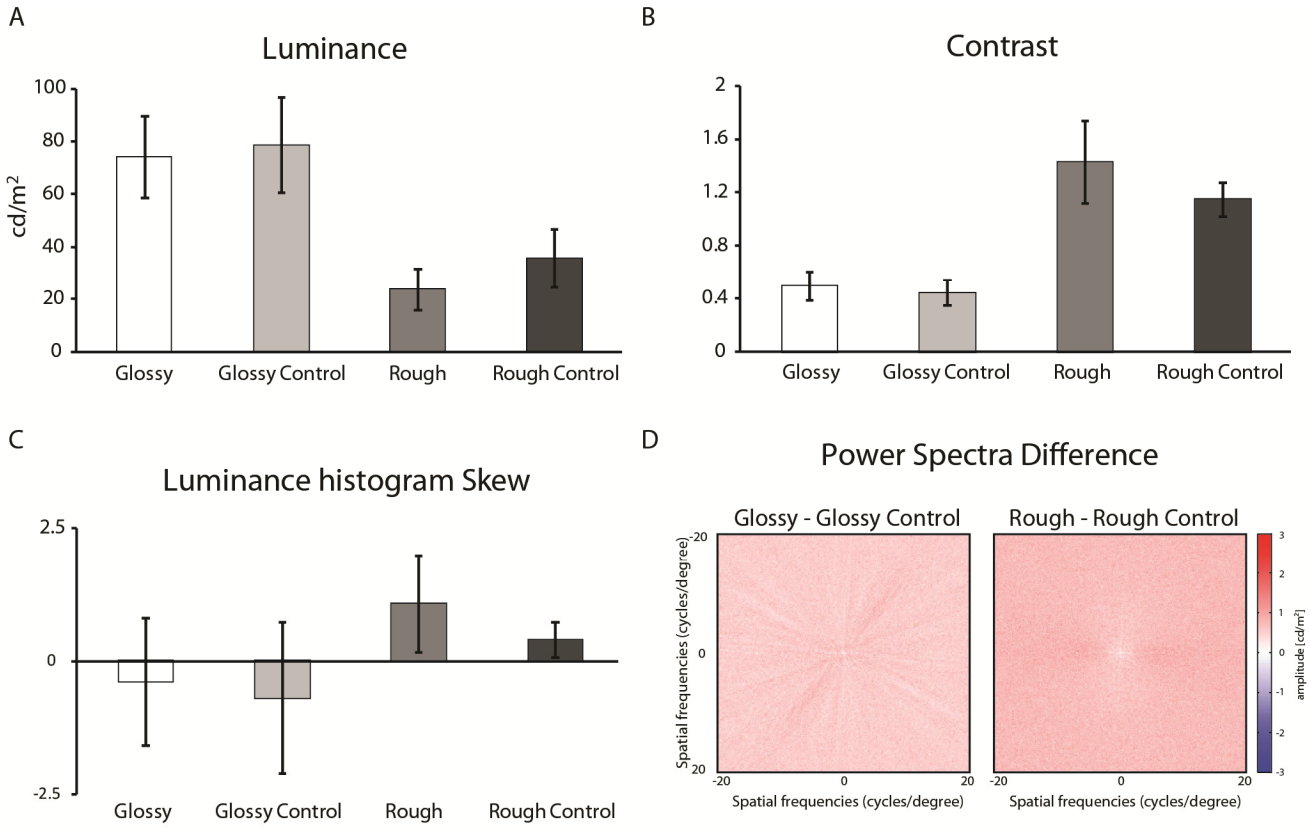


Figure 4. Image statistics of (A) pixelwise luminance, (B) contrast, (C) histogram skew, and (D) difference in power spectra across the 15 images (3 objects × 5 levels) under the four conditions. Luminance was calculated by averaging the mean luminance of all pixels in each image then averaging across images. Contrast was calculated with pixelwise lumincance's standard deviation divided by its mean for each image (SD/Mean), averaged across images. Skew was calculated as the third standardized moment of the luminance histogram of each image, averaged across images (Motoyoshi et al., 2007). The absolute difference in power spectra was calculated for each image pair and then averaged across images.

To analyse the fMRI data, we used multi-voxel pattern analysis (MVPA) to discriminate responses evoked by the different conditions. We found we could reliably decode differences between Glossy *vs.* Rough stimuli in all the regions of interest we had localised (**Figure 5A**, white bars all above the permuted chance baseline in each ROI). Contrasting the Glossy Control and Rough Control stimulus conditions allowed reliable predictions to be made across the visual cortex, although performance in somatosensory cortex was not reliably above chance (**Figure 5A**, black bars). We further ran a 2 (G *vs.* R and GC *vs.* RC) × 11 (ROIs) ANOVA to compare the difference between the two contrasts. We found significant ROI effect ($F_{10,150}$=62.7, p<.001), and a significant interaction between the two factors ($F_{10,150}$=2.3, p<.05). Tukey's HSD post-hoc tests (p<.05) showed that only in area S2 the decoding performance was significantly higher for the Gloss *vs.* Rough than the Glossy Control *vs.* Rough Control comparison. This suggests that in S2 the difference between glossy and rough surfaces that is evoked by the visual displays is greater than between control conditions that are similarly smooth.

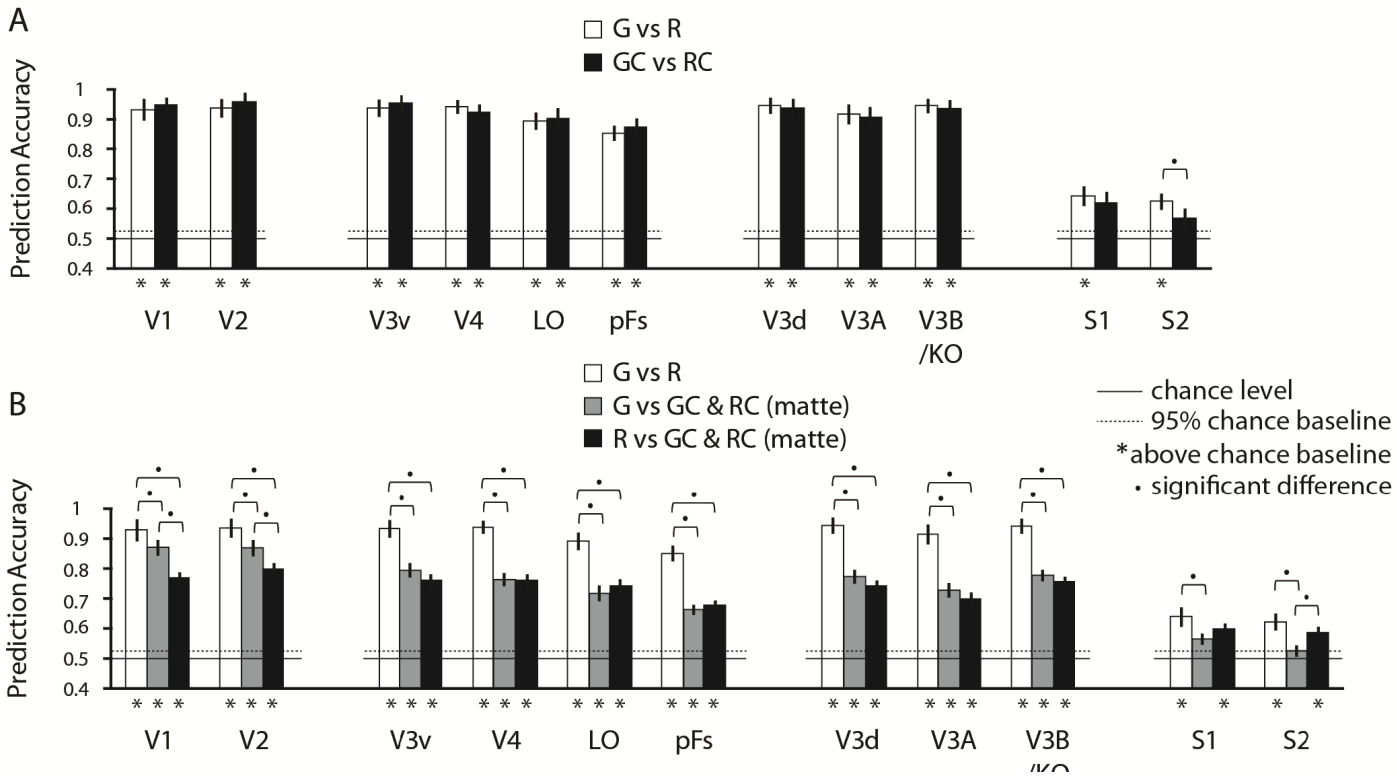


Figure 5. Classification performance of MVPA across 16 participants for (A) Glossy *vs.* Rough (white bars) and Glossy Control *vs.* Rough Control (black bars), and (B) Glossy *vs.* Rough (white bars) , Glossy *vs.* Matte (gray bars) and Rough *vs.* Matte (black bars). The bars reflect mean classification accuracy with ± 1 SEM. Chance level was 0.5 for all classifications as shown in solid horizontal lines. Dotted horizontal lines represent the upper 95th percentile with Permutation tests (see '*Data analysis*'). The one-tailed, 95% boundaries of accuracy distributions were averaged across all ROIs, which were 52.21% for G *vs.* R, 52.25% for GC *vs.* RC, 51.88% for G *vs.* M, 52.24% for R *vs.* M). Asterisks below the bars represent significant above-chance accuracies (single-sample *t*-test, one-tailed, Bonferroni corrected, $p<.05$). Dots above the bars show significant difference between the contrasts with Tukey's HSD post-hoc tests ($p<.05$).

In addition, we tested for differences in the ability to decode surface gloss and roughness. In particular, we used MVPA to contrast Glossy *vs.* Rough conditions as before, as well as the Glossy condition and the Rough condition against Matte conditions (corresponding to the combined GC and RC conditions). We found that all the areas we had localised could reliably decode differences between Rough *vs.* Matte and Rough *vs.* Glossy stimuli (**Figure 5B**, black bars and white bars all above the permuted chance baseline in each ROI). Glossy *vs.* Matte also showed the similar pattern in visual areas, however, performance in somatosensory cortex was not reliably above chance (**Figure 5B**, white bars). In all areas we found highest decoding performance when comparing Glossy *vs.* Rough visual stimuli. This may relate to larger differences in image statistics (**Figure 4**) between these conditions, or more heterogeneous activity evoked by combining data from different conditions into the same class. To compare the difference between the three contrasts, we ran a 3 (G *vs.* R, G *vs.* M and R *vs.* M) × 11 (ROIs) ANOVA. We found significant difference between the three contrasts ($F_{2,30}$=49.3, p<.001), significant ROI effect ($F_{10,150}$=65.2, p<.001), and a significant interaction between the two factors ($F_{20,300}$=7.2, p<.001). Tukey's HSD post-hoc tests (p<.05) revealed that performance in discriminating Glossy *vs.* Rough conditions was higher than in Glossy *vs.* Matte and Rough *vs.* Matte in all the visual areas, while in somatosensory areas Glossy *vs.* Rough was only higher than Glossy *vs.* Matte but not Rough *vs.* Matte. To sum up, the pattern of prediction accuracies differed in somatosensory areas from that in visual cortex. First, performance in discriminating Glossy and Matte conditions was at chance. Second, Rough *vs.* Matte conditions supported reliable decoding performance that was not significant in the Glossy *vs.* Matte comparison. This result was strongest in area S2 indicating that this somastosensory area processes visual surface information at the mesoscale level rather than microscale level (Ho, Landy, & Maloney, 2008), as Rough and Matte conditions are different in the former while Glossy and Matte are different in the latter. This is consistent with

previous studies that different processing of glossy and matte object was primarily found in visual cortex rather than somastosensory cortex (Kentridge et al., 2012; Okazawa et al., 2012; Sun et al., 2015; Wada et al., 2014).

These data suggest that activity in S2 can be driven through visually presented information. However, we also considered the possibility that the route to activity in S2 might be somewhat indirect. In particular, it was possible that viewing the stimuli simply caused the participants to imagine the surface of the objects, with this tactile imagery responsible for the fMRI responses we recorded. We therefore conducted an additional experiment in which we instructed participants to imagine objects with different surface construction, to assess fMRI responses in our regions of interest. Participants viewed the contours of the objects that were presented in the main experiment. These shapes were filled with a homogenous colour that indicated different surface properties that they should imagine (they learnt the association between colour and surface properties prior to being scanned). While we found we were able to decode differences in the visual appearance of the object across visual regions of interest (note that both colour and shape of visually-presented objects are also different across conditions), performance in somatosensory areas dropped to chance (**Figure 6**). In addition, to compare the difference between the two contrasts in **Figure 6A** and the three contrasts in **Figure 6B**, we ran a 2 (G *vs.* R and GC *vs.* RC) × 11 (ROIs) ANOVA and a 3 (G *vs.* R, G *vs.* M and R *vs.* M) × 11 (ROIs) ANOVA respectively. We only found significant ROI effect ($F_{10,50}$=37.0, p<.001; $F_{10,50}$=41.7, p<.001) in the two ANOVA. No significant difference across contrasts and interaction were found in both. To sum up, the results suggest that imagery *per se* is rather unlikely to underlie the responses we measured. Rather, it seems that viewing objects with different tactile-related visual properties causes activity in somatosensory cortex with little conscious effort on behalf of the participants.

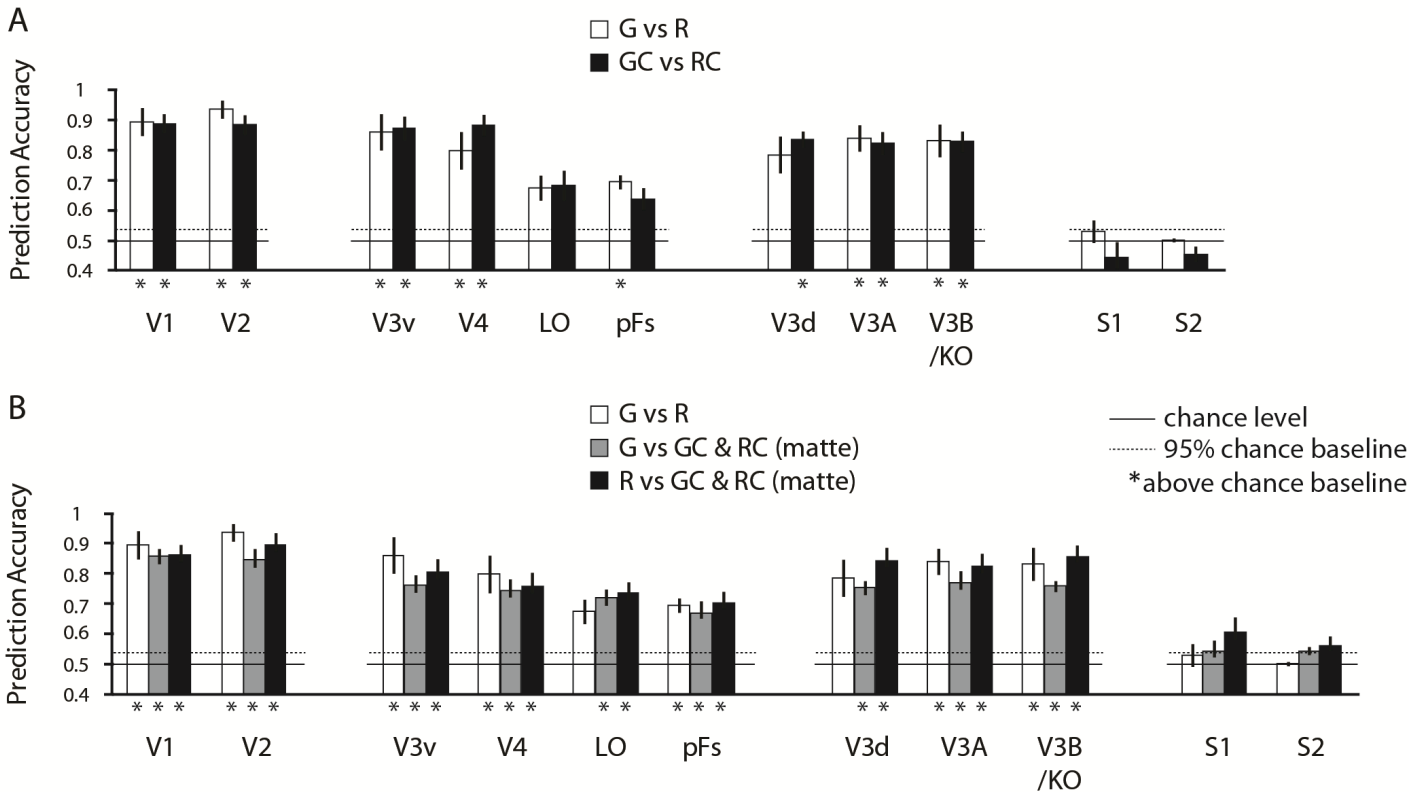


Figure 6. Classification performance of MVPA in the control experiment for the same contrasts as in Figure 2. Dotted horizontal lines represent the upper 95th percentile with Permutation tests (The one-tailed, 95% boundaries of accuracy distributions were 53.80% for G *vs.* R, 53.78% for GC *vs.* RC, 53.10% for G *vs.* M, 53.12% for R *vs.* M). Asterisks in the bottom of the bars represent significant above-chance accuracies (single-sample *t*-test, one-tailed, Bonferroni corrected, $p<.05$).

3.4 General discussion

Here we tested how surface properties of viewed objects evoke activities in different parts of the cerebral cortex. We found that visually responsive areas of the brain contained fMRI responses that discriminated between different classes of objects, as might be expected. Surprisingly, we found that these same images lead to differences also in somatosensory cortex that rough and smooth surfaces lead to differential activations of areas responsible for processing touch stimuli in the fingers. These findings show the importance of how expected surface properties pre-activate the brain through a visual-somatosensory crossmodal network. We speculate that this network may be responsible for predicting the outcome of our interactions (i.e. predict the friction and weight of the object) so promote a more successful action planning (e.g. determining the required force and precision when picking up nearby objects).

The differential activations of somatosensory cortex for rough and smooth surfaces we found were unlikely due to the confounding such as imagery, memory and other non-tactile visual characteristics. Such differences did not appear to be due to imagery processing as these effects disappeared in the imagery control experiment. Visual inputs were therefore likely to be necessary to elicit somatosensory activations and this processing is assumed to be an automatic mechanism that did not affect by top-down modulation. It also did not associate with remembered sensations as the stimuli we used were all novel and virtual. Moreover, it did not relate to non-tactile characteristics (e.g. object colour and patterns on it) as in the control experiment we used different object colour for the four conditions which did not lead to difference in somatosensory cortex. Even in the two control conditions (all matte) the

patterns on the objects were fairly different, no difference in somatosensory activation was found. In other words, somatosensory cortex might be specifically responsible for visual properties that could elicit different haptic sensation (e.g. roughness).

The imagery control experiment is used to test whether imagination of the object surfaces can generate similar somatosensory activations comparing with the main experiment. Therefore we made the design similar to the main experiment, including the 1-back task. One may argue that with the 1-back task imagery may not be successful. However, the 1-back task is assumed to be very easy (requires little cognitive resource/working memory) in which participants can easily detect repetition in shape and luminance. This 1-back task also helped us to evaluate whether participants really paid attention to the stimuli, which is necessary for imagery. The imagery task is also assumed to be easy that participants can easily recall the surface properties (glossy/matte/rough/texture) which remain unchanged through a block. Our data showed that participants did the 1-back task well and they clearly remember what surface properties to imagine for each colour counters (they all reached 100% accuracy in the tests before and after their scans). In addition, imagery tasks in other studies are much more difficult and people are successful doing it (e.g. mental rotation of objects or imaging two objects and judging on their material/geometry) (Kosslyn, Ganis, & Thompson, 2001; Newman, Klatzky, Lederman, & Just, 2005). Therefore the imagery experiment should be a valid manipulation that participants should have no difficulty doing it.

The activation patterns between the two somatosensory areas (S1 and S2) also differed in some respects. Somatosensory responses about visual roughness were particularly pronounced in area S2, where we found significant differences between fMRI responses to Glossy *vs.* Rough and Glossy Control *vs.* Rough Control conditions in S2 but not in S1. We also found clear differences between fMRI responses to rough *vs.* matte and glossy *vs.* matte

conditions in S2 but not in S1.The differences we observed between S1 and S2 can be related to the selective activation of S2 with tactile stimulation (Sathian et al., 2011; Stilla & Sathian, 2008) where rougher surfaces elicit stronger activation (Pruett et al., 2000; Simões-Franklin et al., 2011).

Our data demonstrate that the representation of surface properties in S2 is not based only on information from one sensory modality, and that its activation does not require tactile input. This finding challenges the view that visual and tactile surface information is processed largely independently as previously inferred by observing the qualitatively different encoding, processing, and representing of texture information in the two modalities (Bergmann Tiest, W. M. & Kappers, 2007; Eck, Kaas, Mulders, et al., 2013; Guest & Spence, 2003; Sathian et al., 2011; Stilla & Sathian, 2008; Whitaker et al., 2008). Instead, our results support perceptual equivalence between vision and touch (at least partially) on surface texture and roughness. This implies that surface texture information is represented in a similar way in vision and touch and it can be transferred between the two modalities to some extent, in line with evidence from human psychophysics (Jones & O'Neil, 1985; Lederman, Susan J & Abbott, 1981; Picard, 2006). Moreover, it extended the recent finding that only familiar visual objects which contains content-specific information (e.g. category properties) were represented in somatosensory cortex (Smith & Goodale, 2015). We demonstrate that somatosensory cortex represents visual surface information even if it is novel and cannot be recognized and categorized. This suggests that somatosensory cortex may receive information from the visual processing stage that is earlier than object reorganization and categorization.

A number of previous studies have shown an influence of visual input on somatosensory activation (Blakemore, Bristow, Bird, Frith, & Ward, 2005; Eck, Kaas, &

Goebel, 2013; Järveläinen, Schürmann, & Hari, 2004; Keysers, et al., 2004; Kuehn, Trampel, Mueller, Turner, & Schütz-Bosbach, 2013; Nakano, Ueta, Osumi, & Morioka, 2012). However, these studies did not identify bi-sensory texture-selective regions in somatosensory areas (Amedi et al., 2001; Eck, Kaas, & Goebel, 2013; Sathian et al., 2011; Stilla & Sathian, 2008). Instead, bi-sensory texture-selective areas were found in the middle occipital cortex (MOC), left lingual gyrus (LG), left ventral premotor cortex (PMv) and left inferior frontal gyrus (IFG) (Sathian et al., 2011; Stilla & Sathian, 2008). Here we found that S2, an area commonly thought to process tactile information about surface texture, can also be activated by visual information alone. It is possible that the subtle differences in activation patterns that are captured by the MVPA approach we used were not captured by previous work that used conventional general liner models or percent signal change analyses (Amedi et al., 2001; Sathian et al., 2011; Stilla & Sathian, 2008).

The visually-induced somatosensory activation found here is compatible with an anticipatory system that extracts surface properties from visual information, perhaps in preparation for a possible interaction with it. Such an anticipatory system might be crucial for providing information about surface and material properties that determine friction and dynamic properties (i.e., deformability), which in turn should be considered in planning an action (Di Luca & Ernst, 2014). For example, people expect metal objects to be heavier, stiffer, and smoother than wood objects and therefore contact them and grip them with more force (Bergmann Tiest, W. & Kappers, 2014; Buckingham, Cant, & Goodale, 2009). Moreover, previous studies showed that the anticipation of a sensory input activated similar networks as during real sensory stimulation, including S1 (Porro, et al., 2002; Porro, Cettolo, Francescato, & Baraldi, 2003) and S2 (Carlsson, Petrovic, Skare, Petersson, & Ingvar, 2000; Porro, Lui, Facchin, Maieron, & Baraldi, 2004). This suggests that the pre-activation of S2 by

visual material stimuli might activate the same network that is responsible for tactile perception of surface mesostructure and material.

This mechanism underlying the S2 processing by visual material inputs might be similar to observed touching, being touched, or observing other people using tools – all interactions that involve the mirror system (Blakemore et al., 2005; Järveläinen et al., 2004; Keysers et al., 2004; Kuehn et al., 2013; Nakano et al., 2012). In our experiment it is unlikely that participants retrieved tactile information from memory since the objects were all unfamiliar and abstract in nature. Rather, we speculate that during their lifetime, humans are exposed to cross-modal associations, i.e. a smooth tactile sensation with shiny objects and high-frequency spatiotemporal stimulation with rough ones. This repeated perceptual stimulation is stored as a memory of the tactile sensation associated with the view of objects' surface - i.e. a coupling prior (Ernst, 2006). This previous work, and the findings we present here, are consistent with the view that not only higher-order association cortex but also early sensory areas which were presumed to be unisensory can be modified by multisensory signals (Ghazanfar & Schroeder, 2006; Merabet, L. B., et al., 2007; Schroeder & Foxe, 2005). Such associations can be reciprocal: for instance, tactile stimulation has been shown to modify activity within the visual cortex of blindfolded participants (Merabet, L. B. et al., 2007).

Except for somatosensory cortex, we also found that responses in visual areas were discriminatable between different classes of objects including Glossy *vs.* Rough, Glossy *vs.* Matte and Rough *vs.* Matte (see **Figure 2**). This result is consistent with previous human and monkey neurophysiologic evidence about the involvement of early visual areas, ventral visual areas and dorsal visual areas (V3A and V3B/KO) in visual material processing (Cant et al., 2009; Cant & Goodale, 2007; Cavina-Pratesi et al., 2010a, 2010b; Goda et al., 2014; Hiramatsu et al., 2011; Nishio et al., 2012; Nishio et al., 2014; Okazawa et al., 2012;

Okazawa et al., 2014; Sun et al., 2015; Wada et al., 2014). Note that the differential responses in the three contrasts might also be due to the differences in low-level image features. Interestingly, the classification performance in discriminating Glossy *vs.* Rough conditions was higher than in Glossy *vs.* Matte and Rough *vs.* Matte in all the visual areas, probably because visual noise (from both low-level features and global components) increased after combining the two control conditions (Matte) so that classification performance decreased accordingly. Moreover, V1 and V2 showed better classification performance for discriminating Glossy *vs.* Matte than Rough *vs.* Matte, consisting with the evidence that early visual areas are responsible for processing surface gloss (Motoyoshi et al., 2007; Okazawa et al., 2012; Wada et al., 2014). We did not observe these effects in the imagery session (**Figure 6**) because it primarily involved in color decoding.

In summary, we found that somatosensory area S2 is responsive to the surface characteristics of roughness and glossiness that are conveyed by visual information. While visual areas respond to both surface properties and low-level image features, we found that area S2 primarily responds to visual surface properties that imply different tactile sensations. This area may constitute part of a circuit that predicts the outcome of our interactions with nearby objects to facilitate action planning.

Research limitations

It is still not clear whether roughness and glossiness are coded in a single continuum or different continuums and whether there is a difference between visual and somatosensory modalities from the current study. Further studies (e.g. single-cell recording) are needed to reveal this issue. Although we did not find visual surface relevant activation in

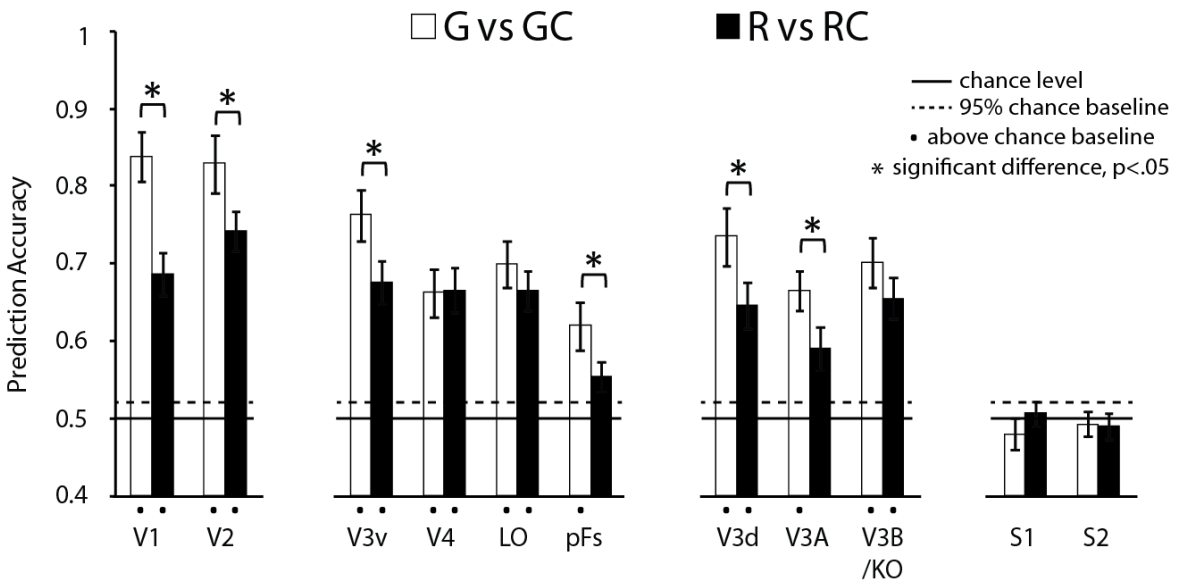
somatosensory cortex in imagery control experiment, however, it might be because participants did not perform vivid imagery of surface. In addition, we still cannot completely rule out memory effect in driving somatosensory activation because the virtual objects we used might have similar reproductions in real world that participants might have tactile experience with.

Afterword

This study is the first study which shows texture-selective regions in somatosensory cortex. This finding suggests that visual information about object surfaces is transformed into tactile information in somatosensory cortex. This visual-somatosensory cross-model network may provide neural evidence about the expectation of sense of touch through object's visual appearance.

Glossiness varies in microstructure (which is invisible) of surface properties while roughness varies in mesostructure of surface properties (i.e. bumpiness) (Ho et al., 2008). Glossiness and roughness are usually taken as two different material properties (Fleming et al., 2013; Granzier, Vergne, & Gegenfurtner, 2014). Comparing Glossy vs. Glossy Control and Rough vs. Rough control would be sensible under this scenario. This analysis can also answer the question that whether the network involved in processing glossiness is different from the network involved in processing roughness. We found that we could reliably decode differences between Glossy *vs.* Glossy Control stimuli in all the visual areas (white bars all above the permuted chance baseline in each ROI). We could also decode differences between Rough *vs.* Rough Control stimuli in most visual areas except for pFs and V3A (black bars). To compare the difference between the two contrasts in, we further ran a 2 (G *vs.* GC and R *vs.* RC) × 11 (ROIs) ANOVA. We found significant difference between the two contrasts ($F_{1,15}$=14.0, p<.01) significant ROI effect ($F_{10,150}$=29.6, p<.001), and a significant interaction between the two factors ($F_{10,150}$=4.8, p<.001). Tukey's HSD post-hoc tests showed that V1, V2, V3v, V3d, V3A and pFs had significantly higher classification accuracy for G *vs.* GC than R *vs*. RC (p<.05). These results suggested that most visual areas (in both dorsal and ventral stream) process information about glossiness and roughness quite differently and they

were more selective to glossiness than roughness in general, especially in V1, V2, V3v, V3d, V3A and pFs. This may implied that the two material properties could inform about complementary aspects of objects and then they were analysed separately. In addition, we replicated our finding in Study 1 that pFs and V3B/KO are crucial for glossiness processing with different control procedure (i.e. highlight rotation).



Additional figure. Classification performance of MVPA across 16 participants for (A) Glossy *vs.* Glossy Control (white bars) and Rough *vs.* Rough Control (black bars). The bars reflect mean classification accuracy with ± 1 SEM. Chance level was 0.5 for all classifications as shown in solid horizontal lines. Dotted horizontal lines represent the upper 95th percentile with Permutation tests (see '*Data analysis*'). The one-tailed, 95% boundaries of accuracy distributions were averaged across all ROIs, which were 52.35% for G *vs.* GC, 52.30% for R *vs.* RC). Dots below the bars represent significant above-chance accuracies (single-sample *t*-test, one-tailed, Bonferroni corrected, $p<.05$). Asterisks above the bars show significant difference between the two contrasts with Tukey's HSD post-hoc tests ($p<.05$).

Chapter 4:

Differential processing of binocular and monocular gloss cues in human visual cortex

This chapter has been submitted as a journal article:

Sun, H.-C., Di Luca, M., Ban, H., Muryy, A., Fleming, R. W., & Welchman, A. E. (submitted). Differential processing of binocular and monocular gloss cues in human visual cortex.

Foreword

In the former chapters we investigated how the brain processes the conventional monocular gloss cue – specular reflectance. How the brain process different gloss cues is not clear yet. In this chapter we investigated how the brain processes surface gloss conveyed by binocular disparities of specular reflections (binocular gloss cue) and compared it with the processing of conventional monocular gloss cue. Importantly, we performed transfer analysis in MVPA to test whether the SVM classifier trained to discriminate glossy vs. matte objects with monocular cue is able to discriminate those with binocular gloss cue and vice versa. Therefore we sought to find the brain area that integrates different gloss cues and represents information about surface gloss in general.

Abstract

Monocular cues to the surface reflectance properties of objects have been shown to involve processing in ventral areas of the human visual system. Here we tested for neural circuits involved in extracting binocular cues to surface appearance. We used human functional magnetic resonance imaging (fMRI) to investigate which brain areas are selectively involved in the processing of binocular cues to gloss. Using computer graphics rendering techniques, we manipulated binocular information to create four conditions that differed in their disparity structure and the impression of surface gloss that they evoked. We performed multi-voxel pattern analysis to test for fMRI responses that distinguish different classes of stimuli based on their depth structure vs. material appearance. We show that higher dorsal areas play a role in processing binocular gloss information in addition to known ventral areas involved in material processing. We then tested for similarities between the representation of gloss from binocular cues and those evoked by (monocular) image-based cues. In particular, we test for transfer in the decoding performance of an algorithm trained on glossy vs. matte objects defined by (a) binocular or (b) monocular cues. We found transfer effects from monocular to binocular cues in left V3B/KO, suggesting the integration of the two cues in this area while other areas process them differently. However, we did not find any transfer from binocular to monocular cues. Our results suggest a more complex processing for binocular gloss information which required greater cooperation between brain regions (i.e. IPS) and additional activation patterns.

Keywords

Surface gloss; material perception; specularity, MVPA, fMRI

4.1 Introduction

Surface gloss provides important information about the characteristics of a view object: for instance, shiny metal is usually newer and has better conductance than rusty metal, and fresh apples have glossier skin than rotten ones. Glossiness is also an crucial material property which helps identifying what the material is as it changes with material categories systematically (Anderson, 2011; Fleming et al., 2013). In this study we aim to understand how gloss information is processed and represented in the brain.

Our visual system is able to extract information about surface gloss efficiently and effortlessly. However, to achieve this is never easy given that the light entering our eyes contains a complex mixture of reflections from the environment. In most cases in the real world, light reflected by an object contains both diffuse reflections and specular reflection. Diffuse reflections carry information about the object itself (e.g. colour, texture) while specular reflections carry information about the environment (e.g. the contents of illumination field). In this case the main goal to achieve gloss perception in the visual system is to decompose specular reflection from diffuse reflections and calculate the relative proportion of the two (Anderson, 2011). Several neurophysiological and neural imaging studies have been mostly focused on investigating the neural basis of this (here we call it 'monocular gloss information') (Kentridge et al., 2012; Nishio et al., 2012; Nishio et al., 2014; Okazawa et al., 2012; Sun et al., 2015; Wada et al., 2014). Studies in monkey functional magnetic resonance imaging (fMRI) and single-cell recordings have shown that monocular gloss information is processed along ventral visual pathway from V1, V2, V3, V4 to superior temporal sulcus (STS) and inferior temporal (IT) cortex (Nishio et al., 2012; Okazawa et al., 2012). Similar findings were also revealed in the human brain recently where

gloss information was found to be processed primarily in ventral visual areas such as V2, V3, V4, ventral–occipital (VO-1/VO-2) area, lateral occipital (LO) area, collateral sulcus (CoS) and posterior fusiform (pFs) (Sun et al., 2015; Wada et al., 2014). In addition, those human studies have shown the involvement of V3B/KO in gloss processing, an area that belongs to dorsal processing stream.

In a more extreme case such as a pure mirror or chrome-plated object (which is rare but still exists in the real world), the extraction of surface gloss is more complicated. In this case no diffuse reflection is shown and all the reflections are from the illumination field with distortion by the surface geometry (Anderson, 2011; Fleming et al., 2003). It is a challenging for the visual system to recover information about surface gloss and shape with the reflections from the environment in this case. Moreover, in the real world we view objects with the two eyes, and this binocular disparity leads to a mismatch between perceived surface from mirror reflection and the true surface of an object. With a more complex shapes (e.g. potatoes), there are even mismatches between the reflections to the two eyes due to large vertical or horizontal disparities, which makes some areas of object unfusible (Muryy et al., 2013). However, little is known about how the visual system works in this complex condition and extract gloss information successfully, for example, to identify whether an object is really a glossy, mirror object which reflects light from the environment or a matte object with the similar painted patterns on the surface (Doerschner et al., 2011; Fleming et al., 2003). The visual system needs to identify specular reflections on the object binocularly and check whether the reflections of environment follow the law of reflections or not to determine perceived gloss. As a result binocular information becomes a crucial source to decode surface gloss in this case.

This paper is concerned with binocular cues, as disparity has been shown to also play an important role in glossiness perception (Blake & Bülthoff, 1990; Kerrigan & Adams,

2013; Muryy et al., 2012; Obein et al., 2004; Sakano & Ando, 2010; Wendt et al., 2010; Wendt et al., 2008). For example, it has been found that the disparity of a highlight affects perceived glossiness, leading to the suggestion the human visual system must have an internalised model of the physic of specular reflections and such model is employed to estimate object gloss. It has been also suggested that perceived gloss is the result of the weighting of a variety of cues, such as highlight intensity, sharpness, coverage and depth of specular reflections (Marlow & Anderson, 2013; Marlow et al., 2012). This theory can explain why high-level factors not directly related to specular reflectance (e.g. object geometry, direction and type of illumination) can modulate perceived gloss, since these factors also affect highlight intensity, sharpness, coverage and depth (Marlow & Anderson, 2013; Marlow et al., 2012).

While previous work has revealed areas involved in processing signals about surface cues from monocularly visible features, here we ask if any additional areas are involved when viewing gloss information indicated by binocular signals. Using computer graphics rendering, we were able to manipulate stimuli to produce four different conditions. First, we used physically correct rendering of mirrored objects. Second, we modified the rendering process so that the disparity signals were wrong for physically correct specular reflections, but key statistics of the binocular signals were matched, and the stimuli evoked an impression of viewing a glossy surface. Third, a 'painted' condition in which the pattern of reflections was 'stuck' onto the surface of the object with the effect that monocular features were almost identical to a gloss object, but when stimuli were viewed stereoscopically the object appeared matte. Finally, we used stereoscopically flat images of reflective objects. We thereby sought to test for neural responses relating to changes in binocular signals vs. the perceptual interpretation of surface material properties. In this study we focused on not only the areas found important in Study 1 and Study 2 (which are earlier areas, ventral visual areas and

dorsal area V3A and V3B/KO) but also we investigated higher dorsal areas in processing binocular gloss information because elaborative stereopsis is assumed to be processed there (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013; Neri, Bridge, & Heeger, 2004; Vanduffel et al., 2002).

4.2 Methods

4.2.1 Participants

Twelve participants who participated in Study 2 took part in this experiment. One was author H.-C. S. and the remainder were naïve participants. Three were male, and age ranged between 19 to 39 years. Participants were screened for normal stereoacuity, MRI safety and provided written informed consent. The study was approved by the STEM Ethical Review Committee of the University of Birmingham. Non-lab member participants received course credits or monetary compensation after the experiment.

4.2.2 Apparatus and Stimuli

Stimuli. A central fixation square (0.5 deg side length) was displayed in the background to provide a constant reference to promote correct eye vergence. We performed the experiment in two sessions: a binocular gloss session, and an monocular gloss session. We first describe the binocular gloss session. We created three different 3D objects (‘potatoes’) rendered using three different illumination maps (**Figure 1A**). (For a detailed description of the rendering process see (Muryy, Fleming, & Welchman, 2014; Muryy et al., 2013)). The objects were approximately 7 deg in diameter, and presented at the centre of the screen with ± 0.4 degree jitter from the centre to reduce the build up of adaptation across repeated presentations at the centre of the screen.

To produce stimuli for the four experimental conditions (‘mirror’, ‘painted’, ‘anti-mirror’, ‘flat’) in the binocular gloss session, we made subtle modifications to the stimulus

rendering process. Under standard mirror reflection (**Figure 1B**), stimuli are rendered by finding the pixel value of point $\boldsymbol{P}$ in the image of left eye ($\boldsymbol{E_L}$) and right eye ($\boldsymbol{E_R}$) by reflecting the viewing vectors $\boldsymbol{V_L}$ and $\boldsymbol{V_R}$ around the surface normal, $\boldsymbol{n}$, to calculate the reflected ray vectors $\boldsymbol{\omega_L}$ and $\boldsymbol{\omega_R}$. These point to particular image intensities in the spherical illumination map, determining the pixel intensities that should be presented to $\boldsymbol{E_L}$ and $\boldsymbol{E_R}$. Using computer rendering, we can manipulate this process to change the locations from which the objects are imaged for the purpose of defining the pixel intensities of the object, while keeping the stereoview frustum constant (**Figure 1C, D**) (this method is described in detail by Muryy et al. (2014)). In the 'mirror' condition, stimuli are generated following the normal specular reflection (**Figure 1B**), creating the impression of a mirrored object. In the 'painted' condition (**Figure 1C**), the specular reflections were effectively stuck onto the surface of the object, so that the specular reflections have the same stereoscopic depth as the object's surface; in this case, the 3D structure of the object is clearly visible, but the impression of surface gloss is very considerably reduced. In the 'anti-mirror' condition (**Figure 1D**), we reversed the locations from which image intensities in the environment are determined between the two eyes. This changes the disparity structure of the images, but stimuli are nevertheless appreciably 'glossy' ((Muryy et al., 2012). We also created a 'flat' condition (**Figure 1E**) in which the same image of the object was presented to both eyes. Participants' judgments also consisted with our manipulation that mirror and anti-mirror objects looked glossier than painted and flat objects (**Figure S1**).

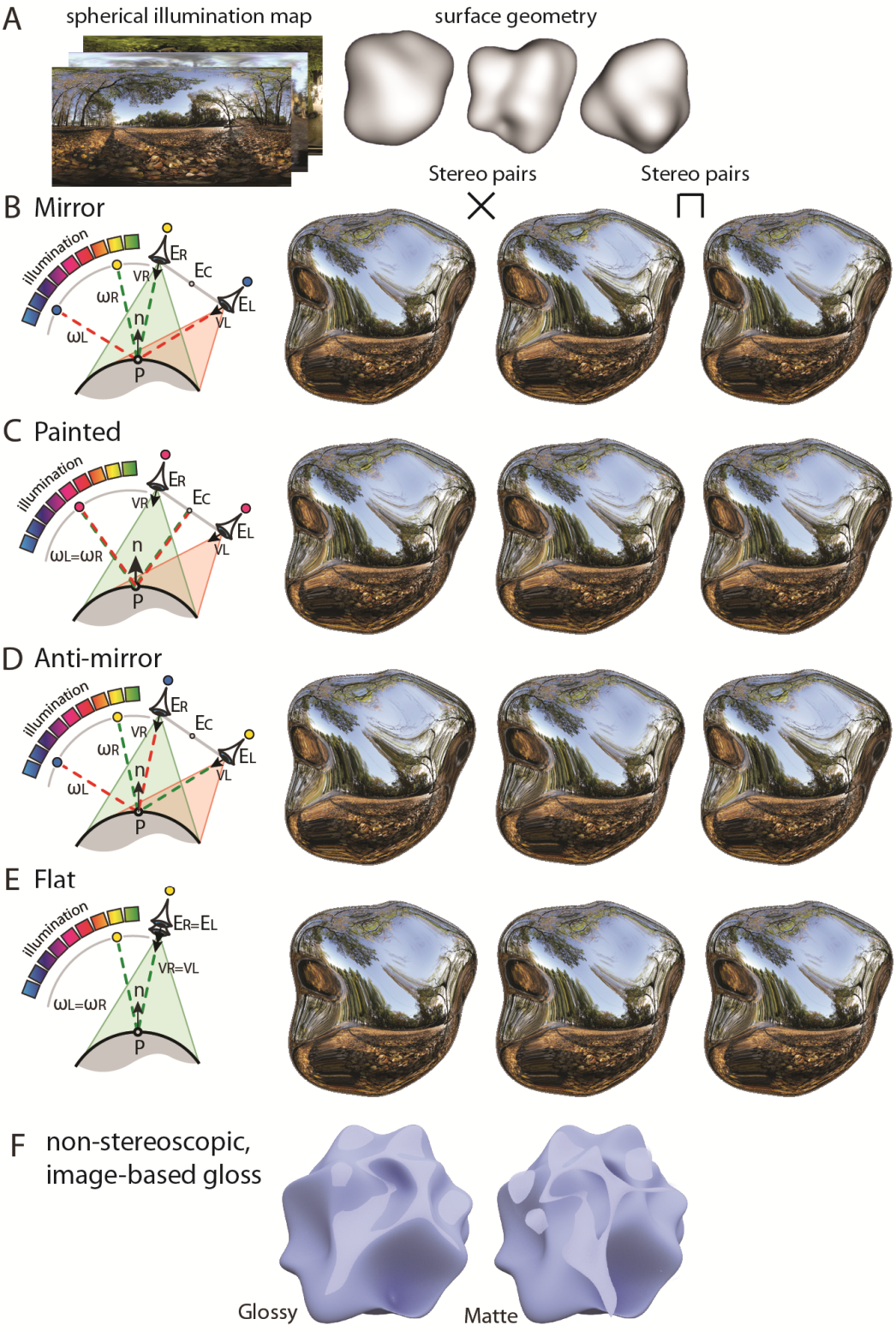


Figure 1. Experimental stimuli used in binocular (A-E) and monocular (F) gloss session. (A) Three illumination maps were rendered on the three objects (‘’potatoes’’) under four experimental conditions. (B-E) Schematic description of rendering procedure of each

condition shows in the left and an example stereogram shows in the right with cross-fusion and parallel-fusion separately. In the mirror condition (B), the reflections entering each eye follows the law of specular reflection so that normal specular reflections were presented creating a physically correct image of a polished object reflecting its surrounding environment (schematically illustrated using the color spectrum). In the painted condition (C), pixel intensities for each location on the surface of the object are determined based on the reflection of a ray cast from midway between the participant's eyes (the cyclopean point, $\boldsymbol{E}_{\mathbf{C}}$). The object is imaged from the true positions of the two eyes, meaning that the environment effectively acts as a texture painted onto the surface of the object. In anti-mirror condition (D), the reflected ray vectors are reversed for the two eyes, so the left eye images a portion of the environment appropriate for the right eye. This alters the disparities produced by reflection, but the object appears glossy (Muryy et al., 2012). In the flat condition (E), we randomly select the image of one eye (the right eye in the example) and present to both eyes. Objects look flat and made specular reflections have the same apparent depth as the image plane. In this condition objects look painted as in painted condition (see **Figure 2**). (F) An example stimulus in monocular gloss session (which is the same experiment as in Study 2). Specular components are presented in Glossy condition while in Glossy Control condition those components are rotated by 45 degree, making it a matte surface.

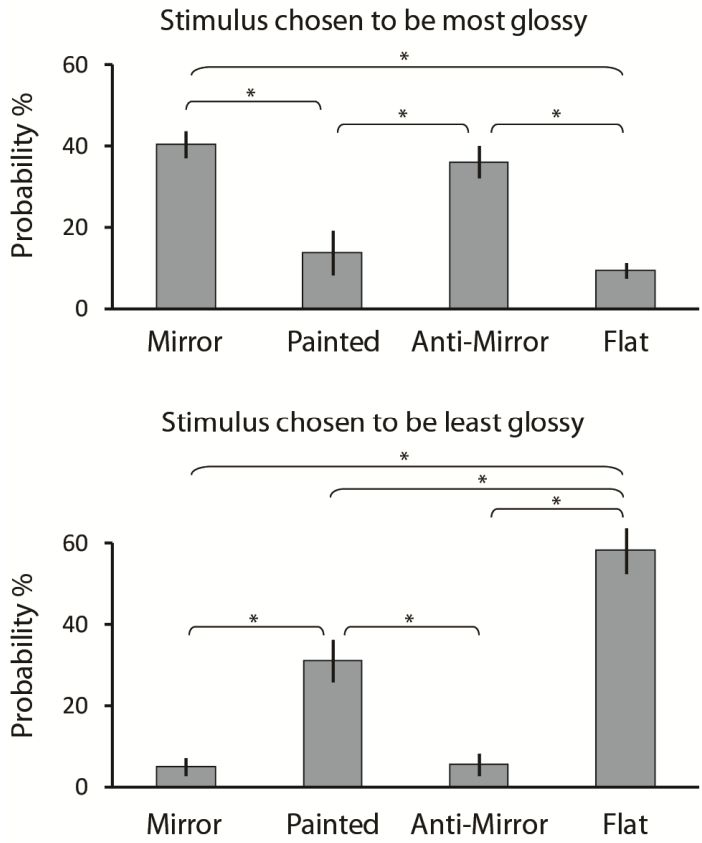


Figure 2. The result of behaviour task outside the scanner from 6 new naïve participants. Partipicants were presented 4 objects at the same time in each trial (corresponding to the 4 experimental conditions, in which the surface geometry and illumination maps were randomly arranged for each object). They were asked to choose the most gloss object in one block and the least glossy object in the other block. The order of the 2 blocks was balanced across participants. There were 180 trials in each block. The probability for choosing in each condition was averaged across participants and plotted in two graphs for the most and the least glossy block respectively with ± 1 SEM. One-way ANOVA (mirror, painted, anti-mirror, flat) was significant for both blocks ($F_{3,15}$=12.0, p<.001 for most glossy block; $F_{3,15}$=27.3, p<.001 for least glossy block). Tukey's HSD post hoc test showed that (p<.01) mirror and anti-mirror conditions were judged glossier than painted and flat conditions, and also the latter two were judged less glossy than the former two. In addition, flat objects were judged less glossy than painted objects.

The stimuli used in the monocular gloss session were the same experiment as in Study 2. Only data from the Glossy and Glossy Control conditions are presented here as roughness is not directly relevant to the current study. An example of object rendered in Glossy and Glossy Control conditions is shown in **Figure 1F**.

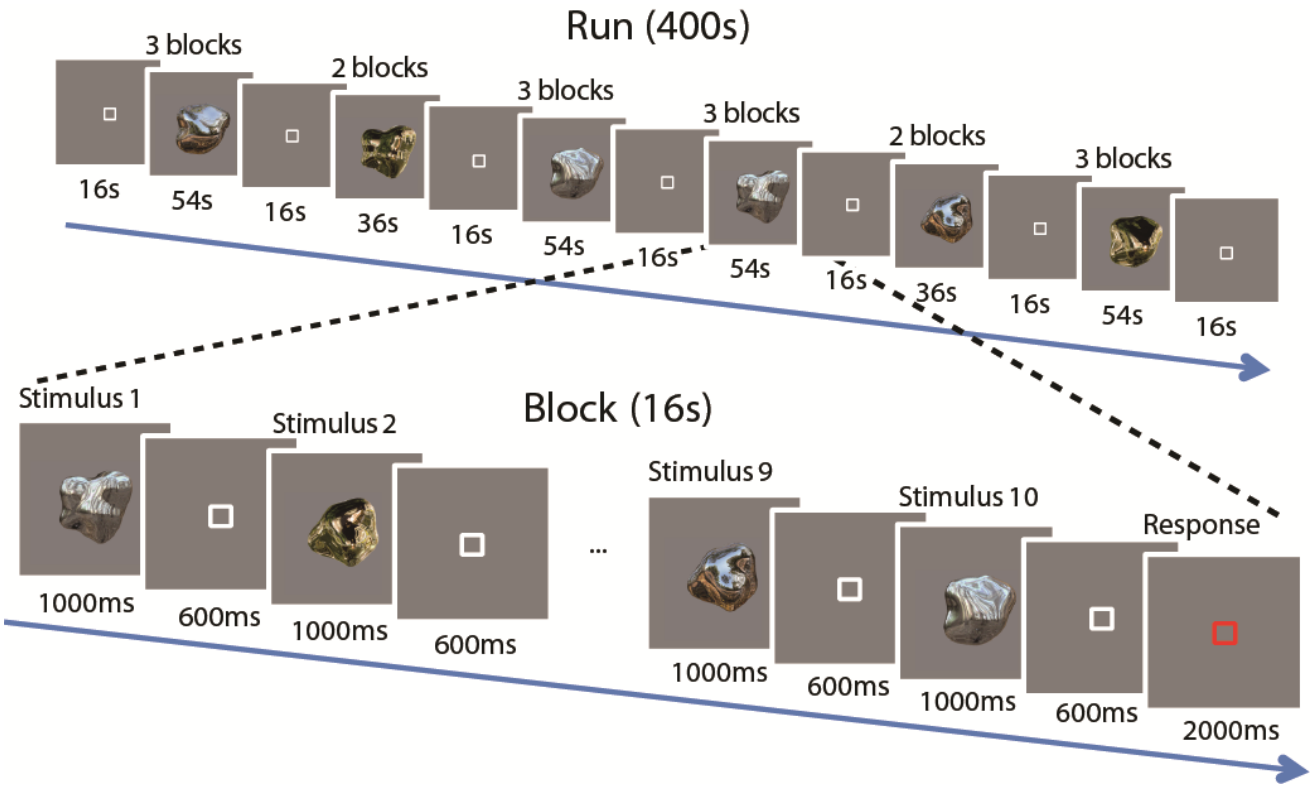
Apparatus. Stimulus presentation was controlled using MATLAB (Mathworks Inc.) and Psychtoolbox (Brainard, 1997; Pelli, 1997). The stimuli were back projected by a pair of projectors (JVC DILA SX21) onto a translucent screen inside the bore of the magnet. To present stereoscopic stimuli, the projectors were fitted with spectral comb filters (INFITEC, GmbH) – see Preston et al, (2009, J Neurosci, 29, 1688-98). Participants viewed the stimuli binocularly via a front-surface mirror fixed on the head coil with a viewing distance of 68 cm. In the monocular gloss session, participants viewed stimuli (binocularly) without wearing the Infitec glasses. Luminance outputs from the projectors were linearized and equated for the RGB channels separately with colorimeter measurements. Participant responses during the scan were collected using a five-button optic fiber button box.

MRI data acquisition. A 3-Tesla Philips Achieva scanner with an 8-channel phase-array head coil were used to obtain all MRI images at the Birmingham University Imaging Centre (BUIC). Functional whole brain scans with echo-planar imaging (EPI) sequence (axial 32 slices, TR 2000 ms, TE 35 ms, voxel size 2.5 × 2.5 (inplane) × 3 (thickness) mm, flip angle 80 deg, matrix size 96 × 94) were obtained for each participant. The EPI images were acquired in an ascending interleaved order for all participants. The same sequence was used in both sessions. T1-weighted high-resolution anatomical scans (sagittal 175 slices, TR 8.4 ms, TE 3.8 ms, flip angle 8 deg, voxel size: 1 mm^3) were also obtained to reconstruct cortical surfaces of individual participants and to achieve precise co-registrations of EPI images onto individual anatomical spaces.

4.2.3 Design and Procedure

A block design was used in both sessions. Each session took about 1.5 hour during which each participant completed in 7 to 10 runs (depending on setup time and the participants' needs to rest between scans). The run length was 400 s and 368 s for the binocular- and monocular- gloss sessions, respectively. Each run started with four dummy scans to prevent startup magnetization transients and consisted of 16 experimental blocks each lasting 16 s. There were 4 block types (i.e., one for each condition), repeated four times in a run. In each block of the binocular gloss session 10 objects were presented in a pseudo-random order. Stimuli were presented for 1000 ms with 600 ms interstimulus interval (ISI). Participants were instructed to maintain fixation and perform an oddball task for glossiness judgments. Specifically, at the end of each block (indicated by a change in the fixation marker) participants had to indicate if all of the presented objects had the same glossiness (i.e., all matte, or all glossy), or whether one of the presented objects differed in gloss. They had two seconds to make their response before the next block began. They were able to perform this task well (mean d'=2.04; SEM=0.31). This oddball task is different from the 1-back task in Chapter 2 and Chapter 3 because it requires participants to focus on binocular gloss information instead of simply judging on monocular changes (i.e. illumination and object shape). Five 16 s fixation blocks were interposed after the third, fifth, eighth, eleventh, and thirteenth stimulus blocks to measure fMRI signal baseline. In addition, 16 s fixation blocks were interposed at the beginning and at the end of the scan, making a total of seven fixation blocks during one experimental run. An illustration of the scan procedure is provided in **Figure 3**. In the monocular gloss session, stimuli were presented for 500 ms with 500 ms interstimulus interval (ISI) as described in Study 2. Participants were instructed to maintain fixation and perform a 1-back matching task, whereby they pressed a button if the same

image was presented twice in a row. They were able to perform this task well (mean d'=2.03;



SEM=0.10). Other details were the same as for the binocular gloss session.

Figure 3. The stimulus presentation protocol in binocular gloss session for one scan run and a block. In each run there were 23 blocks (16 s + 2 s response time each), including 7 fixation blocks and 16 experimental blocks. During each experimental block, stimuli were presented for 1000 ms with 600 ms interstimulus interval (ISI). Participants were instructed to perform oddball task in glossiness.

4.2.4 Data analysis

Functional MRI data processing. The basic data processing procedures for both the binocular and the monocular gloss sessions are identical to our previous study (Sun et al.,

2015). To summarise the procedure, we computed the global signal variance of the blood oxygenation level dependent (BOLD) signal for each run using the whole-brain average of activity across volumes. If this exceeded 0.23% the scan run was excluded from further analysis to avoid the influence of scanner drifts, physiological noise or other artifacts (Junghöfer et al., 2005). On this basis, 17/146 runs and 6/118 runs across 12 participants for binocular and monocular gloss session respectively were excluded from further analysis.

ROI definition. A total of 15 regions of interest (ROIs) were defined. For all participants V1, V2, V3v, V4 V3d, V3A, V3B/KO (kinetic occipital region), hMT+/V5 (human motion complex), LO (lateral occipital region) and pFs were defined by localizers in a separate session as in previous studies (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013; Sun et al., 2015). For 7 of the 12 participants, higher dorsal areas V7, ventral intraparietal sulcus (VIPS), parieto-occipital IPS (POIPS), dorsal IPS medial (DIPSM), and dorsal IPS anterior (DIPSA) were also defined by a localizer in which random-dot stereogram with 3D structure from motion (SfM) information was contrasted with moving dots without stereogram and SfM information. The detailed information about each localizer and ROI definition can be found in Appendix 1. For the other 5 participants, V7 was identified as anterior and dorsal to V3A and other dorsal areas defined according to Talairach coordinates ([x,y,z] = [30, -78, 27] for right VIPS; [-27, -72, 30] for left VIPS; [24, -75, 45] for right POIPS; [-18, -72, 54] for left POIPS; [18, -60, 63] for right DIPSM; [-15, -63, 60] for left DIPSM; [39, -36, 54] for right DIPSA; [-36, -48, 60] for left DIPSA) and draw around GLM t-value maps that had *t* value larger than 0 for the contrast of "all experiment conditions vs. fixation block" (Dövencioğlu et al., 2013; Murphy et al., 2013; Orban et al., 2003).

Additional fMRI analysis. We used multivoxel pattern analysis (MVPA) to compute predictions accuracies for the experimental conditions. We selected voxels by first computing the contrast "all experimental conditions vs. fixation", and then selecting the top 250 voxels

from this contrast within each ROI of each individual participant. If a participant had fewer than 250 voxels in a particular ROI, we used the maximum number of voxels that had t values greater than 0. After selecting the voxels, we extracted the time series (shifted by 4 s to account for the hemodynamics response delay) and converted the data z-scores. Then, the voxel-by-voxel signal magnitudes for a stimulus condition were obtained by averaging over 8 time points (TRs) (= 1 block) separately for each scanning run. To remove baseline differences in the response patterns between stimulus conditions and scanning runs, we normalized by subtracting the mean for each time point. To perform the multi-voxel pattern analysis (MVPA), we used a linear Support Vector Machine (SVM) implemented in libsvm toolbox (http://www.csie.ntu.edu.tw/~cjlin/libsvm, Chang & Lin, 2011) to discriminate the different conditions in each ROI. In the training phase, 32 response patterns for each stimulus condition were used as a training dataset for those participants that completed 9 runs and 36 response patterns were used for those who completed 10 runs. Then, 4 response patterns for each condition were classified by the trained classifier in the test phase. These training/test sessions were repeated and validated by a leave-one-run-out cross-validation procedure. The ROI-based prediction accuracy for each participant was defined as a mean of these cross-validation classifications. We also used a searchlight classification analysis approach (Kriegeskorte, Goebel, & Bandettini, 2006) whereby we defined a spherical ROI with 8 mm radius, and moved it through the entire volume of cortex. For each location, we recomputed the SVM classification analysis.

4.3 Results

To test for visual responses related to binocular and monocular cues to gloss, we first identified regions of interest within the visual and parietal cortex (**Figure 4**). We then used MVPA to test for responses related to the impression of glossy *vs*. matte surfaces. In particular, we used responses in different experimental conditions to understand how fMRI signals might relate to changes in the material appearance of the viewed object vs. changes in the disparity-defined depth structure. To this end, we concentrated on three main contrasts (**Figure 5A**). First, we tested for responses related to surface gloss, contrasting the mirror and anti-mirror conditions (both perceived as glossy) against the painted (perceptually matte) object. The disparities defining the 3D shape in these conditions all differ, however stereoscopic shape between these conditions is more similar to that of the flat condition. Second, we performed a contrast between the mirror and anti-mirror conditions; the logic of this contrast is that while both appear glossy, the raw disparity composition of the shapes is quite different. Third, we contrasted the painted and flat conditions, which provides the maximal change in 3D shape, while the material appearance in these conditions is equivalent. In the extreme scenario of an cortical region specialized for processing surface material, we would expect to be able to decode glossy vs. matte renderings of the stimuli, but not the difference between mirror and anti-mirror conditions, or the difference the painted and flat conditions. While perhaps computationally attractive, such a pure modular architecture was a priori rather unlikely.

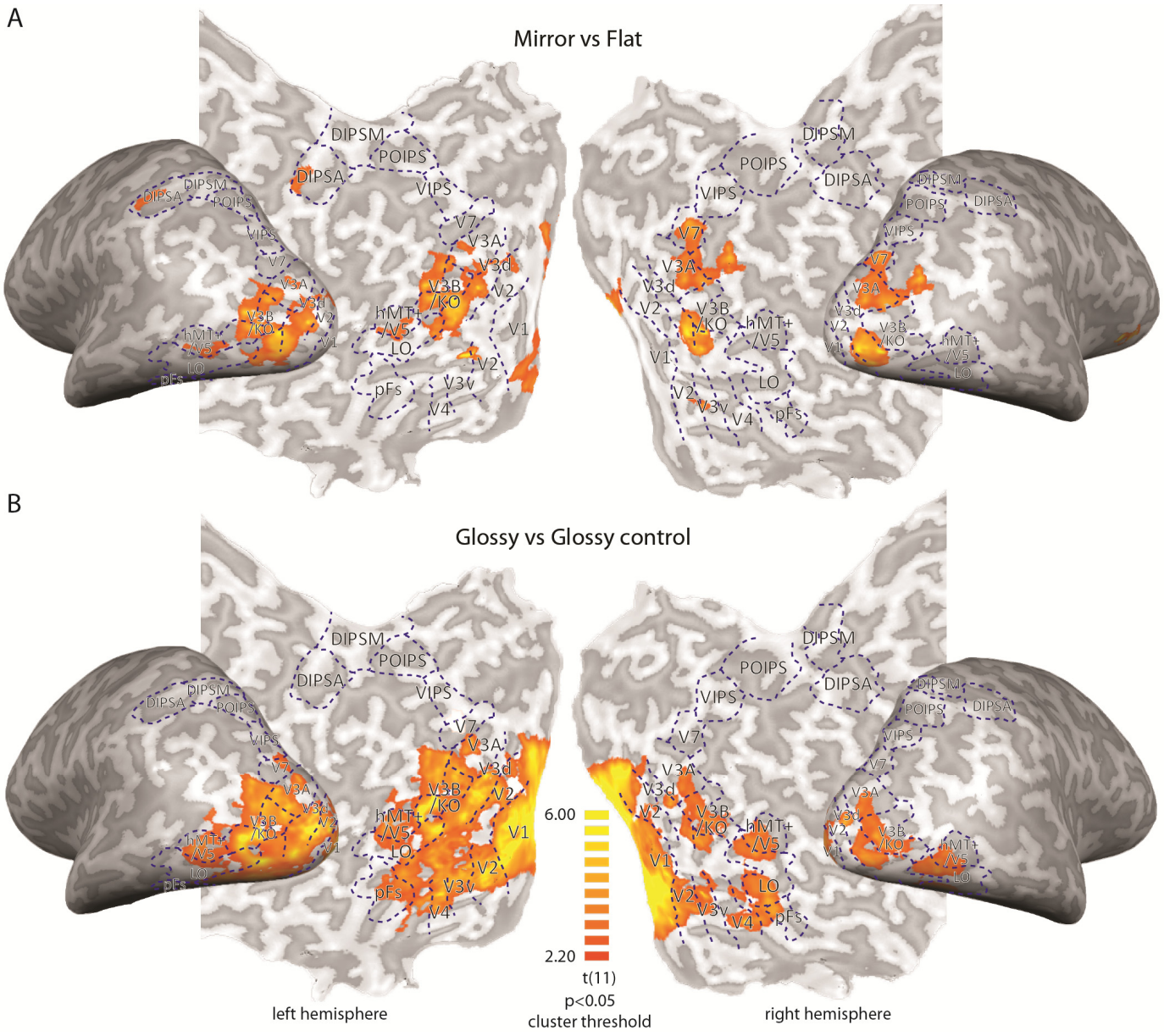


Figure 4. Searchlight classification analysis results for binocular (A) and monocular (B) gloss session. The color code represents significant *t*-value of mirror *vs.* flat and Glossy *vs.* Glossy Control classification accuracies in (A) and (B) respectively testing against chance level (testing against chance level 0.5). Blue dash lines are the ROI boundaries we defined with independent localizers. The significance level was $p<.05$ with cluster-size thresholding 25mm^2. Significant regions are presented on the flat maps of one representative participant. Note that since classification results were averaged across participants and then presented on the flat maps of one representative participant, individual ROI boundaries may not perfectly fit the group level.

We found that we were able to predict the stimulus from the fMRI data at levels reliably above chance in multiple regions of interest (V4, LO, V3d, V3A, V3B/KO, hMT+/V5, V7, VIPS, DIPSM, DIPSA) when contrasting the mirror and anti-mirror conditions against their painted counterparts (**Figure 5A**, black data series). This suggests widespread sensitivity to differences in the material and binocular cues that comprise the stimuli. Considering the differences between the mirror and anti-mirror conditions (**Figure 5A**, gray bars), we were not able to reliably predict the stimuli in any regions of interest. While this might suggest widespread regions of interest respond to glossy material, it seems likely that the relatively subtle deviations in the disparity structure in these two conditions were overcome by the larger change in disparity structure between different shape examples within these categories. Finally, contrasting the painted and flat conditions (**Figure 5A**, white bars) revealed above chance prediction accuracies in V3B/KO, hMT+/V5, V7 and LO. Decoding performance in this condition reveals areas sensitive to changes in the 3D structure of the shapes and agrees with previous work suggesting sensitivity disparity-defined depth in these in these areas (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013).

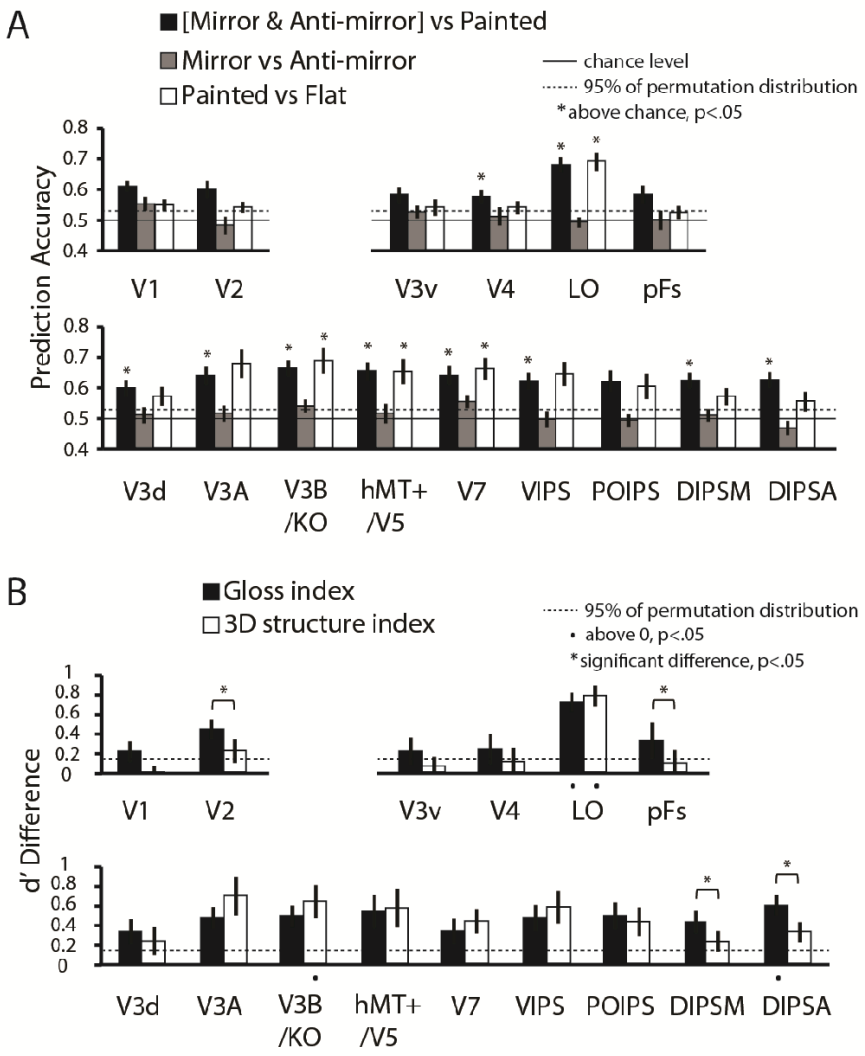


Figure 5. Classification performance of MVPA across 12 participants for (A) [mirror & anti-mirror] *vs.* painted (black bars), mirror *vs.* anti-mirror (gray bars), and painted *vs.* flat (white bars). The bars reflect mean classification accuracy with ± 1 SEM. Solid horizontal lines represent chance performance 0.5 for the binary classification. Dotted horizontal lines represent the upper 95th percentile with Permutation tests (1000 repetitions for each ROI of each participant with randomly shuffling stimulus condition labels per test. The one-tailed, 95% boundaries of accuracy distributions were averaged across all ROIs, which was 52.52% for [mirror & anti-mirror] *vs.* painted, 53.11% for mirror *vs.* anti-mirror, and 53.13% for *vs.* painted flat). Asterisks in the bottom of the bars represent that the accuracies were significantly above the 95% boundaries (single sample *t*-test at p<.05, one-tailed, Bonferroni

corrected). (B) The d' differences between [mirror & anti-mirror *vs.* painted] and [mirror *vs.* anti-mirror] classification were presented as Gloss index. The d' differences between [painted *vs.* flat] and [mirror *vs.* anti-mirror] were presented as 3D structure index. Dotted horizontal lines represent the upper 95th percentile with Permutation tests (1000 repetitions for each ROI of each participant in each contrast with randomly shuffling stimulus condition labels per test. The classification accuracies were transformed into d' for each contrast then the differences between the contrasts were calculated. The one-tailed, 95% boundaries of d' differences distributions were averaged across all ROIs, which was 0.14 for Gloss index and 0.16 for 3D structure index). Black dots in the bottom of the bars represent that the d' was significantly above the 95% boundaries (single sample *t*-test at $p<.05$, one-tailed, Bonferroni corrected). Asterisks above bar pairs represent significant difference between the two indexes (Tukey's HSD post-hoc test at $p<.05$). The bars were arranges in three groups which represent the ROIs in early visual areas, ventral visual areas and dorsal visual areas respectively.

To facilitate comparison of performance between conditions, we calculated a '3D structure index' to examine decoding performance that could be attributed to information about 3D shape. We expressed prediction performance in units of discriminability (d') and contrasted performance for the mirror vs. anti-mirror condition with the painted vs. flat condition based on a simple subtraction. The logic of this contrast is that for both sets of comparisons there is no difference in the material appearance of the shapes, so there contrast reflects differences in the 3D structure of the shapes in both conditions. We also created 'Gloss index' by contrasting performance in the mirror vs. anti-mirror contrast with the [mirror and anti-mirror] vs. painted classification. The logic of this contrast is to compare similarity glossy objects (with different disparity information) against differentially glossy

objects (with different disparity information). We present these data in **Figure 5B** that shows the two indexes across all ROIs. We started by considering whether the indexes were significantly above chance level, using Permutation tests to calculate 95% chance baseline of d' difference for Gloss index (0.14) and 3D structure index (0.16) respectively then testing against it with one sample t-tests (one-tailed, Bonferroni corrected). We found that Gloss index was significantly above the chance baseline in DIPSA (t_{11}=4.4, p<.01) and LO (t_{11}=5.3, p<.01), suggesting that these areas contain signals that distinguish information about gloss. For the 3-D structure index, we found sensitivity significantly above zero in V3B/KO (t_{11}=3.5, p<.05) and LO (t_{11}=4.1, p<.05). These results suggest that LO contain information about both 3D structure and material properties.

We next sought to compare the indices against each other. To this end we ran a 2 (Gloss index and 3D structure index) × 15 (ROIs) repeated-measures ANOVA. This indicated a main effect of ROI ($F_{14,154}$=2.5, p<.01) and importantly a significant interaction with index ($F_{14,154}$=2.8, p<.01). We then used post-hoc contrasts to test the differences between the indices in each ROI. We found a significantly higher Gloss index in V2, pFs, DIPSM and DIPSA, suggesting areas preferentially engaged in the processing of material properties (**Figure 5B**, asterisks indicate p<.05). It is reassuring to note that areas V2 and pFs were previously found to be involved in the processing of information about specular reflectance from monocular cues (Sun et al., 2015; Wada et al., 2014), suggesting that they represent general information about surface gloss regardless of the source. In summary, LO appear to process both surface properties and 3D structure information, while V2, pFs, DIPSM and DIPSA selectively process surface properties. Transfer analysis between [mirror & anti-mirror *vs.* painted] and [flat *vs.* painted] revealed that the processing of surface properties and 3D structure information involves in the same voxels in LO (see **Figure 6**).

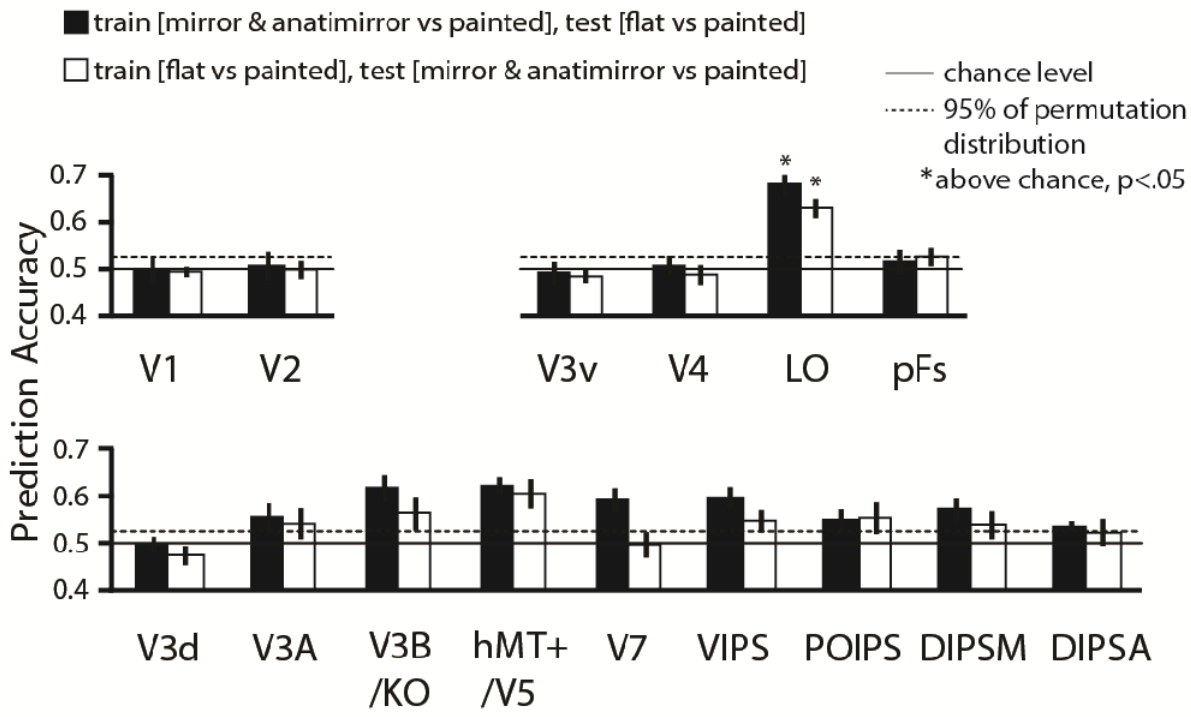
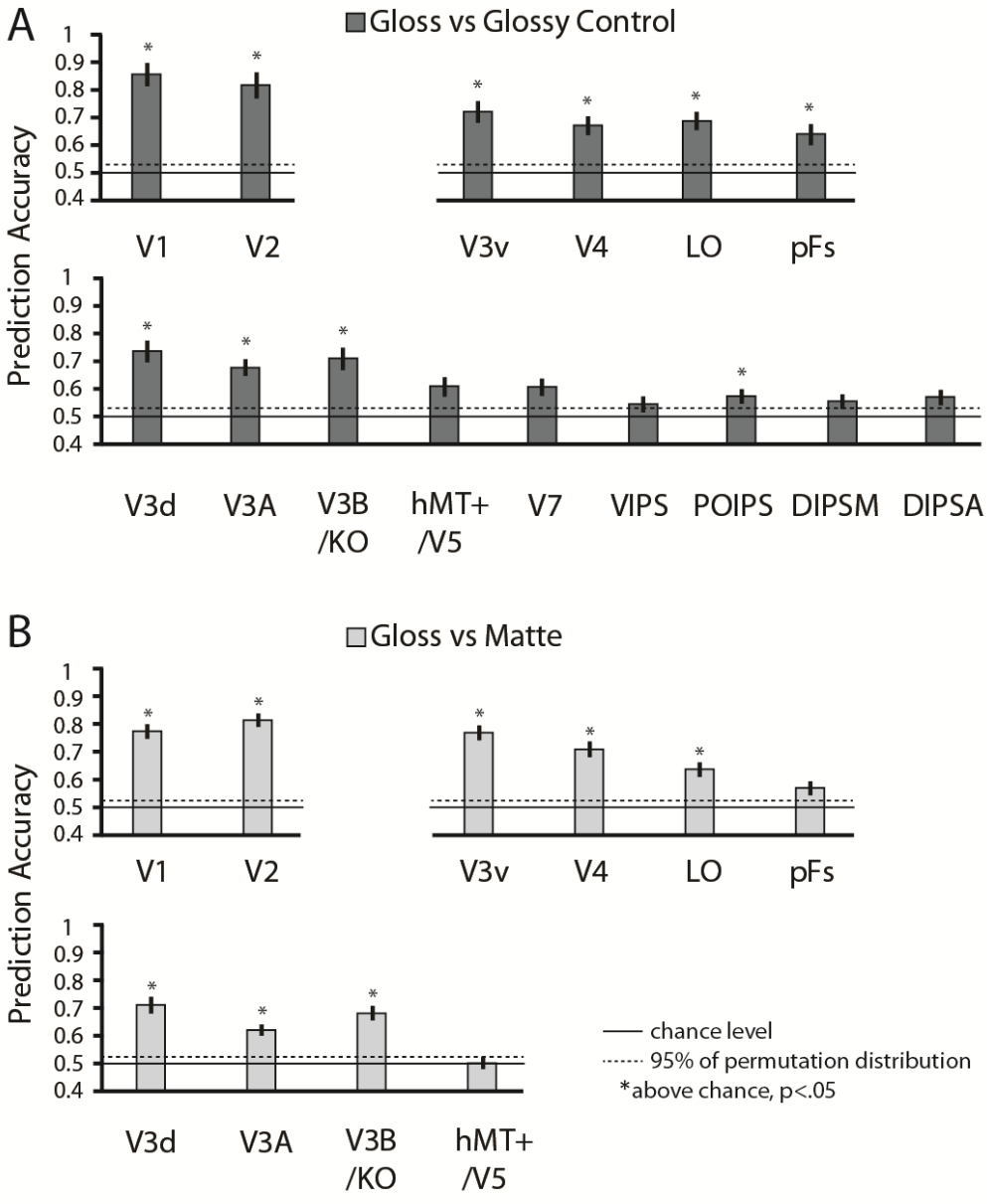


Figure 6. Classification performance of MVPA across 12 participants for transfer analysis between [mirror & anti-mirror *vs.* painted] and [flat *vs.* painted]. We trained the SVM classifier to discriminate mirror & anti-mirror *vs.* painted and tested whether it is distinguishable for flat *vs.* painted (black bars). We also tested the transfer effect in the other way (white bars). The bars reflect mean classification accuracy with ± 1 SEM. Solid horizontal lines represent chance performance 0.5 for the binary classification. Dotted horizontal lines represent the upper 95th percentile with Permutation tests (1000 repetitions for each ROI of each participant with randomly shuffling stimulus condition labels per test. The one-tailed, 95% boundary for black bars was 52.24% and 53.17% for white bars). Asterisks in the top of the bars represent that the accuracies were significantly above the 95% boundaries (single sample *t*-test at p<.05, one-tailed, Bonferroni corrected).

We also performed MVPA to monocular gloss session, testing which areas had dissociable activation patterns between Glossy and Glossy Control conditions. **Figure 7A** shows the classification results of Glossy *vs.* Glossy Control across 12 participants. Classification accuracies were significantly above chance in V1, V2, V3d, V3A, V3B/KO, POIPS, V3v, V4, LO and pFs. These results showed that activity in all the early and ventral visual areas was distinguishable of the monocular gloss cues. Some dorsal areas (i.e. V3d, V3A, V3B/KO, POIPS) were also involved. This result is consistent with the result of our previous study (**Figure 7B**) with the classification contrast of objects with and without



specular highlights (Sun et al., 2015).

Figure 7. Classification performance of MVPA for Glossy *vs.* Glossy Control in monocular gloss session in the current study (A) and in our previous study (Sun et al., 2015) with a group of 15 participants (B). The bars reflect mean classification accuracy with ± 1 SEM. Solid horizontal lines represent chance performance 0.5 for the binary classification. Dotted horizontal lines represent the upper 95th percentile with permutation tests with Permutation tests (1000 repetitions for each ROI of each participant with randomly shuffling stimulus condition labels per test. The one-tailed, 95% boundary in A was 52.79% and 52.39% in B). Asterisks in the top of the bars represent that the accuracies were significantly above the 95% boundaries (single sample *t*-test at $p<.05$, one-tailed, Bonferroni corrected). Higher dorsal areas (V7-DIPSA) were not defined in (B) as parietal localizer was not applied in that study.

Although from the MVPA results it seems that some areas (i.e. V3d, V3A, V3B/KO, V4, LO) support decoding of both monocular and binocular gloss cues, it is not clear whether the same voxels (or their subregions) are involved in the processing of both cues. We examined this issue by performing a transfer analysis. First, we trained the SVM classifier to discriminate between Glossy *vs.* Glossy Control condition in the monocular gloss session and then tested whether the classifier could discriminate between [mirror & anti-mirror] *vs.* painted activation in the binocular gloss session. We also tested the transfer effect in the other direction, that is, we trained the classifier to discriminate between [mirror & anti-mirror] *vs.* painted in the binocular gloss session and then tested whether it discriminates between Glossy *vs.* Glossy Control conditions in the monocular gloss session. Areas that shows transfer effect in both directions are assumed to involve in the processing of both cues. **Figure 8** show the results of classification for the two transfer directions respectively.

Significant transfer effects were found from the monocular to the binocular gloss cues (black bars) in V1, V2, V3d and V3B/KO. However, no significant transfer effects from binocular to the monocular gloss cues were found (white bars). To test whether we might have missed any areas involved, we also performed a searchlight analysis. We found that voxels that had significant transfer effects from the binocular to the monocular gloss cues were found primarily in V3B/KO (**Figure 9**). This result suggests that this subregion may combine monocular and binocular gloss information creating a generalized representation of surface gloss. However, because the binocular gloss classification could not be applied to monocular gloss cues, we can infer that binocular gloss information carries additional information compared to monocular gloss.

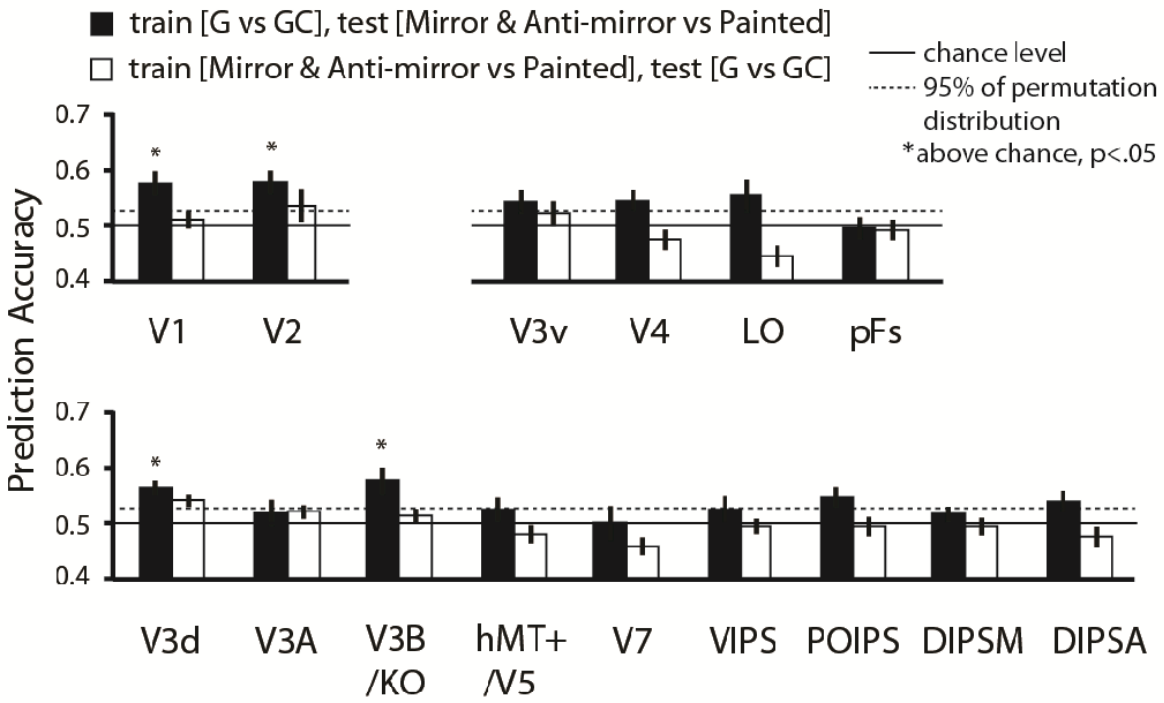


Figure 8. (A) Classification performance of MVPA across 12 participants for transfer analysis between binocular gloss session and monocular gloss session. We trained the SVM classifier to discriminate between Glossy *vs.* Glossy Control conditions in monocular gloss session and tested whether it is distinguishable between [mirror & anti-mirror] *vs.* painted in binocular gloss session (black bars). We also tested the transfer effect in the other way (white

bars). The bars reflect mean classification accuracy with ± 1 SEM. Solid horizontal lines represent chance performance 0.5 for the binary classification. Dotted horizontal lines represent the upper 95th percentile with permutation tests with Permutation tests (1000 repetitions for each ROI of each participant with randomly shuffling stimulus condition labels per test. The one-tailed, 95% boundary for black bars was 52.24% and 53.17% for white bars). Asterisks in the top of the bars represent that the accuracies were significantly above the 95% boundaries (single sample *t*-test at p<.05, one-tailed, no correction was applied).

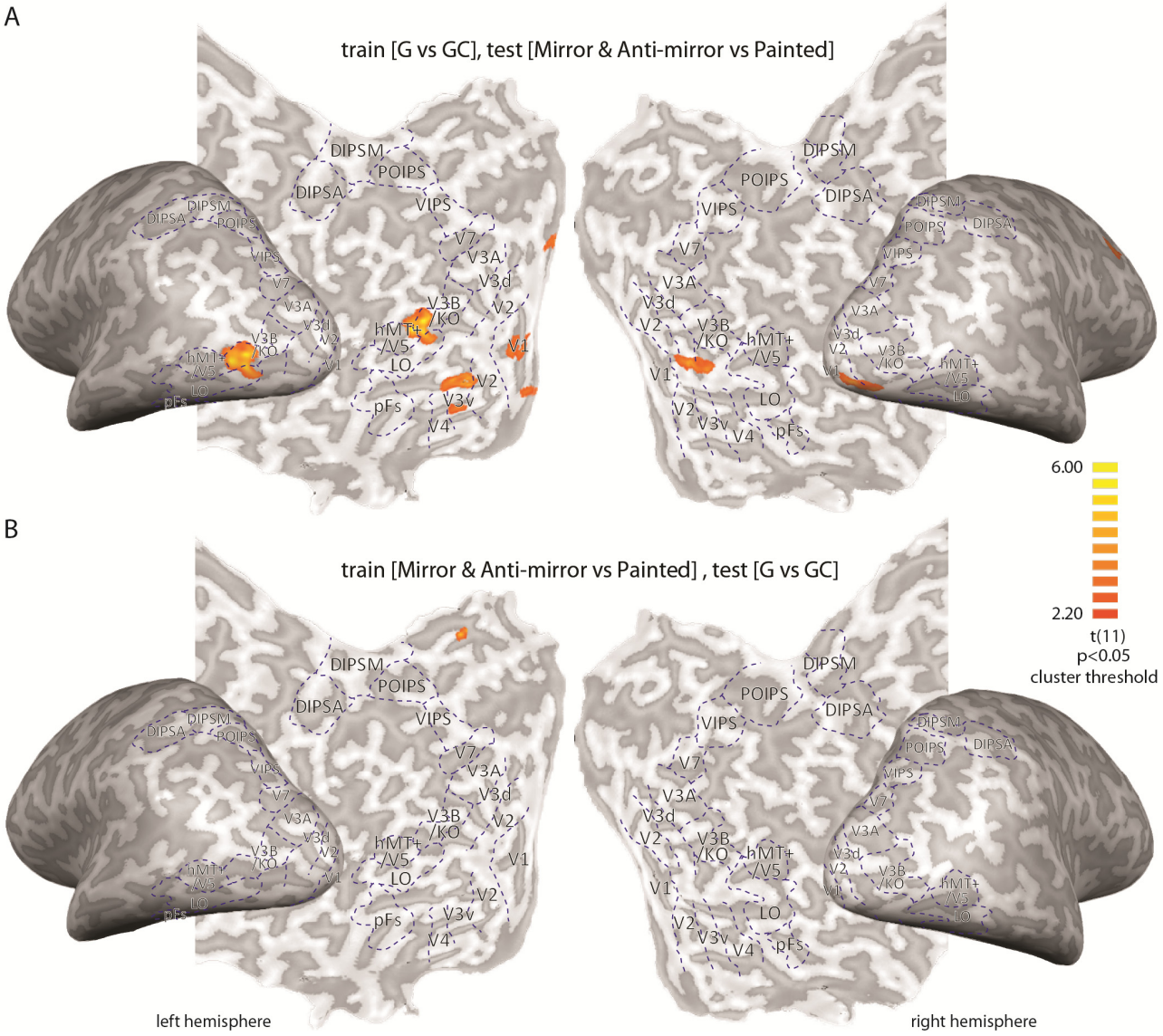


Figure 9. Searchlight transfer analysis results between monocular and binocular gloss sessions. In (A) we trained the SVM classifier to discriminate between Glossy *vs.* Glossy Control conditions in monocular gloss session and tested whether it is distinguishable between [mirror & anti-mirror] *vs.* painted in binocular gloss session. In (B) we tested the transfer effect in the other way. The color code represents the *t*-value of significant classification accuracies testing against chance level (0.5) with cluster-size thresholding 25mm^2. Significant transfer effect was found primarily from monocular to binocular session but not from binocular to monocular session.

4.4 General discussion

Here we sought to test for cortical areas involved in the processing of binocular signals to the impression of surface gloss, drawing a contrast with monocular (pictorial) cues to surface material. We sampled activity across the visual processing hierarchy, and contrasted fMRI responses in conditions that evoked different impressions of surface gloss based on different disparity structures. We found that ventral area LO was sensitive to differences in both the material properties of objects and difference in 3D structure. By contrast we found that differences in gloss evoked more discrminable fMRI responses than differences in disparity-defined shape in dorsal anterior Interparietal Sulcus (DIPSA). By contrasting responses to monocular and binocular- signals to gloss, we found differential involvement of areas within the dorsal and ventral streams. Importantly, V3B/KO appears to be involved in the processing of both types of information. This was supported by a transfer analysis that showed stereoscpically specified gloss could be decoded using an algorithm trained on differences in gloss specified by monocular features. These results point to the involvement of both ventral and dorsal brain areas in processing information related to gloss, with an intriguing confluence in area V3B/KO that has previously been associated with the processing of 3D structure information.

Our approach to investigating binocular cues to gloss was to make subtle modifications to the rendering process so that low-level image statistics were almost identical between different conditions. This allowed us to test for binocular signals to surface reflectance properties, which are likely to interact with image-cues to gloss that have been revealed by previous behavioural studies (e.g. luminance, histogram skew, contrast, spatial frequencies – (Marlow & Anderson, 2013; Marlow et al., 2012; Motoyoshi et al., 2007;

Sharan, Li, Motoyoshi, Nishida, & Adelson, 2008)). To test the impression of gloss from monocular cues, we also introduced a new technique in this study that involved a simple image-editing procedure. Specifically, we manipulated the impression of gloss in the Glossy Control condition by rotating specular highlight components in the image plane. This broke the relationship between surface curvatures specified by the image and the location of reflections (**Figure 1F**) and allowed us to ensure low-level image features were near-identical while the perceptual appearance changed dramatically (Anderson & Kim, 2009; Kim et al., 2011; Marlow et al., 2011). This is a different procedure to that in previous studies that used spatial scrambling, phase scrambling or changing overall luminance (Okazawa et al., 2012; Sun et al., 2015; Wada et al., 2014). It is reassuring that the results of this manipulation (**Figure 7A**) converge with results from our previous work using image scrambling (**Figure 7B**) (Sun et al., 2015) to suggest that monocular gloss cues are processed along ventral areas, and involve dorsal area V3d, V3A and V3B/KO in addition.

In addition to the new stimuli, manipulation and experimental design, we also performed different data analysis approach than before and obtained new findings. We performed both MVPA and searchlight transfer analysis in this study instead of traditional analysis based on intensity of activation (i.e. BOLD signal change) or neuron firing rates (Nishio et al., 2012; Nishio et al., 2014; Okazawa et al., 2012; Sun et al., 2015; Wada et al., 2014). With MVPA we found that early (i.e. V1, V2) and higher dorsal areas (i.e. POIPS) processed information about monocular gloss which was not revealed in previous human fMRI studies with analysis based on percent signal change (Sun et al., 2015; Wada et al., 2014). Except for the ROI-based MVPA, we also exerted searchlight classification through the whole brain. Specifically, we investigated whether the voxels that were trained to discriminate between glossy and matte objects with one gloss cue (binocular or monocular) can discriminate with the other. We located small groups of voxels having transfer effect

successfully by searchlight analysis. They were predominately found in ventral part of left V3B/KO (and small part of V3v, V1 as well, see **Figure 9**), suggesting they are responsible for both binocular and monocular gloss cues. This neural evidence may provide a possible locus that integrates different gloss cues which corresponds to behaviour findings in previous studies (Marlow & Anderson, 2013; Marlow et al., 2012).

For the neural basis of binocular gloss information, we found that both dorsal and ventral areas are involved in the processing of binocular gloss information according to significant above-chance accuracies in the [mirror & anti-mirror] *vs.* painted classification (**Figure 5A**). Such finding is reasonable since information about elaborative stereopsis is assumed to be processed in dorsal areas (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013; Neri et al., 2004; Vanduffel et al., 2002). In addition, since in previous studies it has been found that information about general material perception (e.g. information that allows the classification of metal, bark, glass etc.) is primarily processed in higher ventral areas (Cant & Goodale, 2007, 2011; Cavina-Pratesi et al., 2010a, 2010b; Hiramatsu et al., 2011), the involvement of ventral areas in processing binocular gloss information may be crucial in integrating information into a coherent representation.

We could not find an area whose activation patterns were able to distinguish between mirror and anti-mirror. A possible reason might be that a significant portion of the object shapes we used (**Figure 1A**) provide disparities that likely challenge the visual system in terms of fusibility (Muryy et al., 2014) – so perhaps this is the reason that – despite quite large differences in the disparity content – we could not separate the two conditions. It is also possible that our ability to separate out mirror and anti-mirror resulted from identifying perceived 3D surfaces first, however, the perceived surfaces were discontinuous and broken due to different disparities (Muryy et al., 2013). Thus the heterogeneity with a stimulus class may mean that we could not decode the signals. This neuroimaging result is consistent with

our perceptual judgment that mirror and anti-mirror were perceived the same in glossiness (**Figure S1**). These results suggest that our visual system seems unable to apply the law of reflection precisely to complex reflection and uses simple heuristics instead.

The preference in processing information about binocular gloss and 3D structure is also different across ROIs. We found that V3B/KO, hMT+/V5, V7 and LO were not only responsible for processing binocular gloss information but they are also involved in processing information about 3D structure (**Figure 5A**). The comparison between the Gloss index and the 3D structure index (**Figure 5B**) shows that V2, DIPSM, DIPSA and pFs had better classification performance for decoding binocular gloss information than 3D structure information, indicating that these areas were specialized for processing surface properties rather than objects' 3D structure. Interestingly, V2 and pFs were also found having selectivity for gloss information from specular reflectance in previous studies (Okazawa et al., 2012; Wada et al., 2014) as well as in the current study (**Figure 7**). It seems like that V2 and pFs decoded gloss information from monocular cues much better than from binocular cues. In contrast, LO processes information about binocular gloss and 3D structure equally well (**Figure 5B**) and most importantly, it is the only ROI which shows strong transfer effect between the two kinds of information (**Figure 6**). This result suggests that LO analyses binocular gloss and 3D structure information in a similar way. The reason of this is not clear. One possible explanation is that the processing of binocular gloss and 3D structure influence each other somehow, since previous studies also showed that the disparity of specular highlights affects the perception of 3D structure (Blake & Bülthoff, 1990; Muryy et al., 2013).

The neural basis of monocular gloss information is a little different from binocular gloss. We found that all the earlier (V1, V2, V3d, V3v) and ventral visual areas (V4, LO, pFs) were responsible for processing the monocular gloss cue (i.e. specular reflectance),

consistent with previous findings (Okazawa et al., 2012; Wada et al., 2014). We also replicated the finding that human V3A and V3B/KO are crucial for gloss processing with different control procedure (Okazawa et al., 2012; Sun et al., 2015; Wada et al., 2014). In addition, we found the involvement of higher dorsal area POIPS in decoding monocular gloss information (**Fig. 7A**). The role of POIPS is not clear since no previous evidence has shown their function in processing neither monocular gloss information nor other types of material information. It might be possible that POIPS did not respond to gloss *per se*. Instead, it may respond to the difference of 3D shape (in monocular view) between Glossy and Glossy Control condition. This is because in the Glossy Control condition the rotation of highlights not only destroyed glossiness but also occluded part of shape and shading on the original objects. Also, the rotated highlights themselves also look like a new 'surface' attached on the object, which makes it a more complex shape than the original one. Moreover, the presentation of rotated highlights, which looks unusual, may change attentional distribution when viewing. These factors all have potential influence in the processing of monocular gloss information comparing with previous studies (Okazawa et al., 2012; Sun et al., 2015; Wada et al., 2014).

We found several differences about brain processing between binocular and monocular gloss information. First, earlier areas such as V1 and V2 were found to be selective for processing monocular but not for binocular gloss information. In contrast, dorsal areas (especially for hMT+/V5, V7, VIPS DIPSM and DIPSA) were selective for processing binocular but not for monocular gloss information. This fits the understanding that processing of stereo information needs participation of dorsal visual areas (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013; Neri et al., 2004; Vanduffel et al., 2002). Higher ventral areas such as V4 and LO as well as dorsal area V3d, V3A and V3B/KO were able to distinguish both monocular and binocular gloss information. This evidence is in line with

previous human studies showing that ventral areas and V3A and V3B/KO are process information about specularity (Sun et al., 2015; Wada et al., 2014). Importantly, it also suggests that these areas may represent general information about surface gloss regardless of how it is conveyed. It should be noted that there were slightly different experimental procedures and tasks for the binocular and monocular gloss sessions, which may have affected the difference of SVM classification performance between the two sessions. However, the performance difference across ROIs within each session should not be affected. Moreover, the transfer effect we observed in V3B/KO is less likely due to glossiness-unrelated factors because these factors are very different across sessions (e.g. task, procedures, appearance of stimuli). A direct way to examine whether an area combines monocular and binocular gloss information and represents surface gloss in a general way is to test whether the activities that afford classification evoked by one can be transfer to the classification of the other. Here we trained and tested SVM classifier to discriminate between glossy and matte objects for monocular and binocular gloss information respectively and found transfer effects from monocular to binocular cues in ventral part of left V3B/KO predominately (**Figure 9**). This part may be close to (or overlap with) area LO2 according to Talairach coordinates. LO2 (as well as LO1) was defined by retinotopic mapping in previous studies (Larsson & Heeger, 2006; Wada et al., 2014). However, LO1 was reported to be involved in the processing of surface gloss defined by specular reflectance in the recent human fMRI study (Wada et al., 2014). It might be that LO1 is responsible for monocular gloss cues while LO2 is for binocular gloss cues, and the transfer effect we observe may be due to the intimate interaction between the two areas.

We did not find a transfer effect from binocular to monocular gloss cues. This result suggests that the areas we found distinguishable for both monocular and binocular gloss cues (i.e. V3d, V3A, V3B/KO, V4, LO) may not process the two source of gloss information in

the same way. It is possible that monocular and binocular gloss cues involved in the process of different groups of neurons that are closely located within the ROIs. As a result we could not find transfer effect in these ROIs. Even in some parts of V3B/KO, V3v and V1 which showed significant transfer effect, they may not process the two source of gloss information in the same way, since the transferring occurs in one direction only (monocular to binocular) but not in the other direction (binocular to monocular). The reason might be that binocular gloss involves a more complex processing then monocular gloss and additional activation pattern is needed for the representation of binocular cues. This might be the reason that the classification learned in binocular gloss session could not be applied to monocular gloss session.

It is interesting to ask how the visual system extracts information related to surface gloss from images. For the monocular gloss cues (i.e. specular reflectance), the main computational process is to decompose highlight reflection from its embedded surface (Anderson, 2011). For the binocular gloss cues in the current study, the process appears to be more complicated. First our visual system needs to combine left and right eye images, generating a stereo representation of object structure. Secondly, visual system needs to identify specular reflections on the object and check whether the reflections of environment follow the law of reflections or not, probably by applying some intrinsic heuristics (Kerrigan & Adams, 2013). In addition, the stimuli we used here (potatoes) had complex shapes that lead to the presence of both fusible and unfusible areas in their rendering. Unfusible areas are caused by large vertical or horizontal disparities with magnitude beyond the maximum disparity for fusion (Muryy et al., 2013). These unfusible areas might signal crucial information about surface gloss, since all the regions in matte conditions (painted and flat) are fusible, and unfusible areas are only present in glossy conditions (mirror & anti-mirror).

In summary, here we used systematic manipulation of binocular gloss cues to test for cortical areas that respond to surface material properties. We show the involvement of regions within the ventral and dorsal streams, and draw direct comparisons with cortical responses defined by monocular gloss cues. Our results point to the potential integration of binocular and monocular cues to material appearance in area V3B/KO that showed partial evidence for transfer between different signals.

Research limitations

How the brain decodes binocular gloss information is still not clear in this study. It requires to more studies to examine whether and how the brain make use of fusable/unfusable areas in this processing. The role of POIPS in the processing of monocular gloss is also not clear yet. Previous studies with GLM analysis did not find relevant activations in POIPS. Further studies with MVPA approach are needed to investigate this area in monocular gloss processing. Finally, how the brain integrates binocular and monocular gloss information in V3B/KO and earlier areas is not clear. We only found transfer effect in one direction (monocular to binocular) but not the other direction. The neural interpretation of it needs to be clarified with sophisticated experimental design.

Afterword

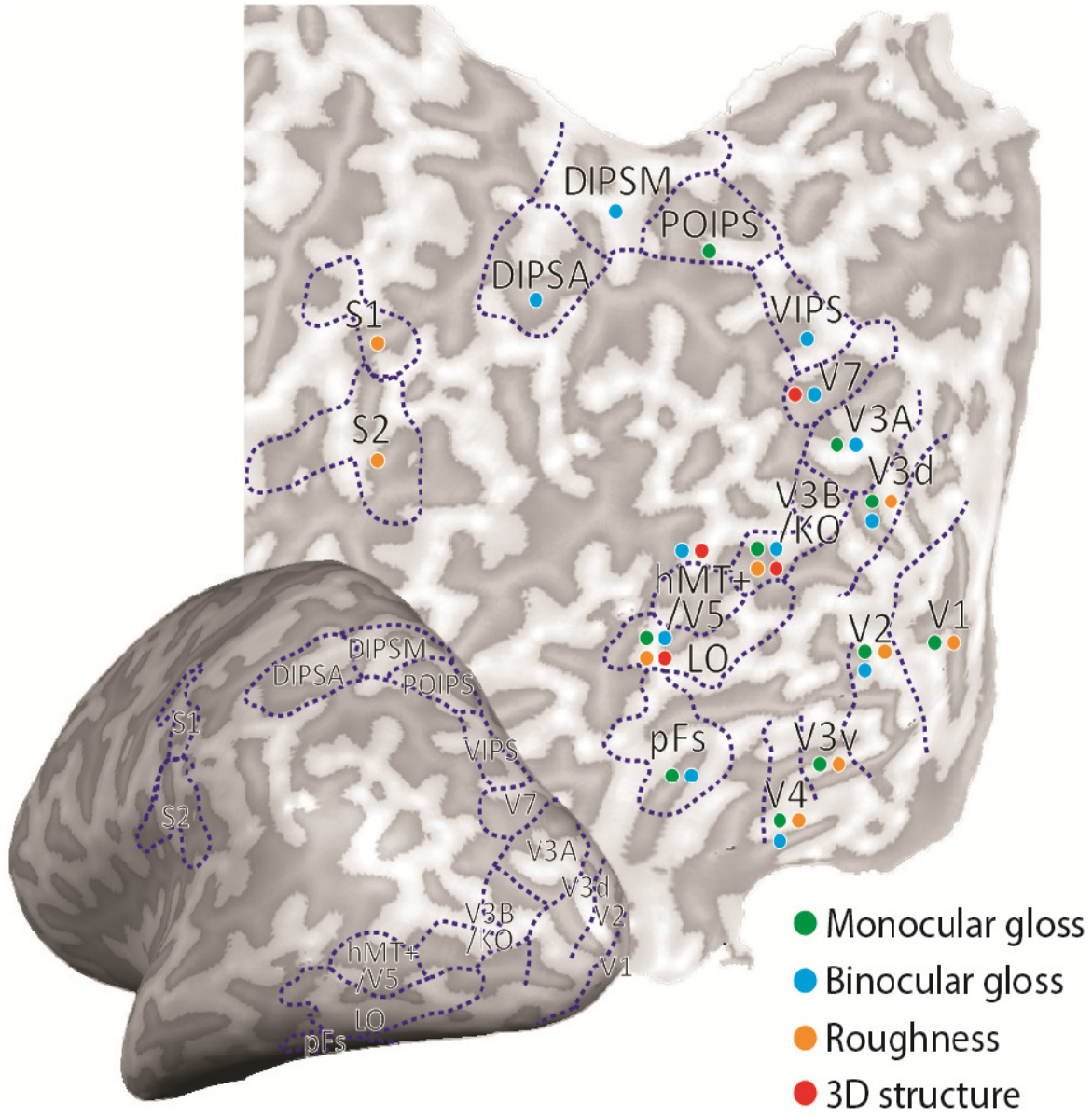
In this study we found that the processing of binocular and monocular gloss information involves similar networks in both dorsal and ventral areas. In particular, earlier and ventral visual areas are more selective to monocualr gloss cue which is consistent with our previous studies, while higher dorsal areas are more selective to binocular gloss cue. Here we found that V3B/KO responded to not only monocualr gloss cue as in previous studies but also responded to binocular gloss cue. Importantly, most of these areas processing binocular and monocualr gloss in independent ways. Only the ventral part of left V3B/KO processes the two cues collectively with additional activation pattern for binocular cues. Overall, our results suggest that the processing of binocular gloss information involves a more complex mechanism that not only includes the areas for monocualr gloss processing found previously but also requires IPS areas.

Chapter 5: Summary

The aim of the three studies was to enhance our understanding about how the brain processes specific material properties (estimation processing). The main material property we selected for investigation is glossiness, due to its importance in providing crucial information which signals objects' status and characteristics. With different ways of manipulation and control procedure of glossiness, the three studies constantly showed that ventral visual areas play a major role in processing gloss information, which is similar to previous studies of general material categorization processing. More specifically, ventral visual area pFs was found to be important for gloss processing throughout the three studies here, while areas related to general material and texture processing were found around fusiform gyrus (FG) and collateral sulcus (CoS) in previous studies (Cant et al., 2009; Cant & Goodale, 2007, 2011; Cavina-Pratesi et al., 2010a, 2010b; Hiramatsu et al., 2011). These areas are located close but distinct in ventral steam according to their Talairach coordinates, suggesting a dissociation between areas underlying material/texture from gloss. This evidence is consistent with other work showed that a patient who had deficits in colour and texture discrimination could judge glossiness correctly (Kentridge et al., 2012). Taken together, the proximity and independence of the areas for glossiness and material processing centres implies their close interrelation and connection. This view is also consistent with other behaviour evidence that estimation and categorization are two interdependent processes that facilitates and interact with each other closely (Fleming, 2014; Fleming et al., 2013).

The processing of surface properties also closely connects with the processing of 3D structure. Previous studies showed that specular highlights influenced the perception of 3D

shape (Blake & Bülthoff, 1990; Fleming et al., 2004; Muryy et al., 2013). In Study 3 we found V3B/KO, hMT/V5, V7 and LO were responsible for processing both binocular gloss and 3D structure information. Importantly, we found transfer effect between the two kinds of information in LO, suggesting that LO processes binocular gloss and 3D structure information in a similar way. In addition, V3B/KO and LO were also found responsible for processing monocular gloss and roughness information in Study 2. This result is consistent with previous finding that V3B/KO integrates different cues related to 3D structure (Ban et al., 2012; Dövencioğlu et al., 2013; Murphy et al., 2013) and LO integrates different cues related to object shape (Kourtzi & Kanwisher, 2000; Larsson & Heeger, 2006). To sum up, V3B/KO and LO might be crucial areas for integrating surface properties and 3D structure information.To further clarify to role of different areas in processing information about surface gloss and 3D structure that summarized from the three studies, I used color coding to indicate the areas that processing the four kinds of information in the following figure: monocular gloss, binocular gloss, roughness, and 3D structure. For monocular gloss (green dots), it is mainly processed in early areas, ventral areas and earlier dorsal areas. Binocular gloss (blue dots) is processed in both dorsal areas and ventral areas except V1. It is more widespread than monocular gloss. For roughness (orange dots), it is mainly processed in early areas, ventral areas, earlier dorsal areas and most importantly--somatosensory areas. For 3D structure (red dots), it is mainly processed in middle areas.



The involvement of gloss processing in each brain area is discussed below in separate sections, from important areas to less important areas:

V3B/KO. Our finding in V3B/KO is very promising and stable across the three studies. Although in the end of Study 1 we doubt the processing of gloss information in V3B/KO since no evidence before showed the involvement of this area in the processing of glossiness or materials until recently (i.e. Wada et al., 2014). In addition, the GCMs show

different information flows between V3B/KO and pFs for gloss processing. Therefore we suspected that in Study 1 V3B/KO may simply respond to the boundaries of highlight in glossy condition rather than responded to surface gloss *per se*. However, in Study 2 and Study 3, even though there was no difference in highlight boundaries between glossy and matte condition, V3B/KO was still able to discriminate between the two conditions, suggesting that it responded more than just highlight boundaries. Another confounding that V3B/KO might respond to is 3D shape, since specular highlights are known to influence the perception of 3D shape (Blake & Bülthoff, 1990; Fleming et al., 2004; Muryy et al., 2013). However, with different control procedures across the three studies, the level of 3D shape effect varied across the three experiments (especially in Study 3, 3D shape is not changed with binocular gloss), and we still obtained very strong and stable results of V3B/KO in discriminating glossiness. Importantly, in Study 3 we found that V3B/KO processes both binocular and monocular gloss cues and we found transfer effect between the two, suggesting that V3B/KO may be the area that integrates both types of information. Therefore, it would be highly possible that V3B/KO indeed responses to surface gloss and it may also act like a 'gloss centre' which integrates gloss cues from a variety information sources. In addition, the involvement V3B/KO might be specific to human only but not other primates since no dorsal areas (V3B/KO or IPS) were found related to glossiness or material processing in macaque studies so far (Goda et al., 2014; Okazawa et al., 2012).

pFs. pFs is another crucial area we found in processing gloss information through the three studies. It constantly showed its selectivity in glossy objects whether the glossiness is defined by monocular (i.e. specular reflectance) or binocular (i.e. disparity of specular reflections) gloss cue. This result suggests that pFs represent surface gloss in general regardless the source of gloss information. Note that no transfer effect was found between the monocular and binocular gloss cues in pFs, suggesting that it processes the two cues in

different way or involving in different voxels. Further studies are needed to examine this issue and to clarify whether pFs can respond to glossiness defined by other gloss cues (e.g. motion cues, illumination). Moreover, it seems like that pFs is particular in processing gloss information rather than other material information, since in Study 2 we found it had significantly higher classification accuracy in decoding surface glossiness than roughness. However, it is not clear whether pFs is exclusively specific for glossiness or it also responds to other material properties (e.g. hardness, coldness, transparency). Further studies are needed to investigate this issue.

LO. Although LO is located very close to pFs, its response patterns differ from pFs throughout the three studies. In Study 1 LO showed weak (not significant) selectivity to gloss objects. In Study 2 we found that LO was able to decode gloss information as pFs, but unlike pFs it decoded roughness information as well. In Study 3 we found that LO was able to decode both monocular and binocular gloss information, but unlike pFs it decoded information about 3D structure as well. These results suggest that comparing with pFs, LO processes object information in a more general way that not just restricts to surface properties but also includes objects's shape and structure. This evidence is in line with previous studies about object-related activities found in LO (Grill-Spector, Kushnir, Hendler, & Malach, 2000; Kourtzi & Kanwisher, 2000; Malach et al., 1995).

Earlier visual areas (V1, V2, V3d, V3v, V4). The activation patterns in these earlier visual areas are quite similar across the three studies. They showed some selectivity in glossiness. In Study 1 the five earlier areas all showed stronger activation for glossy objects than matte objects, however, their activations were much stronger for their scrambled counterparts. This suggests that they primarily deal with low-level image features rather than glossiness. In Study 2 and 3, we found that these earlier areas were distinguishable for

monocular gloss information quite well (around 70% ~ 85% classification accuracies for classifying Glossy *vs.* Glossy Control condition), and they were less sensible in decoding roughness information than glossiness (except V4). In contrast, these earlier areas were not specialized for processing binocular gloss information. Although V3d and V4 showed significant accuracies in decoding binocular gloss information, however, their performance was lower than V3B/KO and LO respectively (which are their closest higher areas). Moreover, although these earlier areas showed marginal transfer effect from monocular gloss cue to binocular gloss cue (Study 3), the magnitude and number of voxvels which showed transfer effect were less than those in V3B/KO according searchlight transfer analysis. Taken together, these results suggest some selectivity for processing monocular gloss in these earlier areas, aligning with previous studies that earlier areas are also related to the processing of glossiness and material (Arcizet et al., 2008; Goda et al., 2014; Hiramatsu et al., 2011; Okazawa et al., 2012; Okazawa et al., 2014; Wada et al., 2014).

V7 and IPS. The role of higher dorsal areas in gloss and material processing has never been discussed in previous studies. In Study 3 we examined these areas in decoding binocular and monocular gloss information. In general we found these areas were able to decode binocular gloss but not monocular gloss, and they usually had better decoding performance for binocular than monocular gloss information. In particular, dorsal IPS area may be specialized for processing binocular gloss that it did not respond to monocular gloss and objects' 3D structure. POIPS is a bit different from the other four higher dorsal areas we defined as it was able to decode monocular gloss but not binocular gloss. Further studies are needed to reveal the function of POIPS in gloss processing. It is also not clear whether these higher dorsal areas respond to binocular gloss information *per se* or other information. For example, they may respond to the small difference of disparity between glossy (mirror & anti-mirror) and matte (painted) condition, since higher dorsal areas were found to be crucial

in processing higher-order disparity (Georgieva, Todd, Peeters, & Orban, 2008; Orban, 2011; Parker, 2007). It is also possible that the processing of higher-order disparity in these areas is necessary in signalling binocular gloss information. It is an open challenge to investigate how the higher dorsal areas involved in processing information about surface gloss.

Somatosensory areas. No studies so far have found visual-haptic bisensory activation about texture information inside somatosensory areas. Here we firstly found that S2 has this bisensory property. With MVPA we successfully revealed that S2 which was defined by sensory inputs (air puffs) to fingers was distinguishable between visually-presented glossy and rough objects. Importantly, this response in S2 cannot be elicited by visual imagery of surface properties, suggesting that visual inputs were necessary in triggering somatosensory responses. S1 seems to have less selectivity to surface properties than S2, suggesting that visual information about surfaces properties is not transformed into tactile information until S2. Overall, we provided neural evidence about S2 in the representation of tactile sensation induced by visual inputs.

STS. Although previous monkey fMRI and single-unit recordings showed that the superior temporal sulcus (STS) responses to glossy objects (Nishio et al., 2012; Okazawa et al., 2012), however, the involvement of human STS seems very weak through the three studies here. We did not presented the results of STS in Study 2 and Study 3 since it there is no much finding as in Study 1. The classification accuracies in all condition in Study 2 and Study 3 were near chance level. As we discussed in the end of Study 1, it might be that the voxel size we used here is too large to detect small responses to glossy stimuli in the human STS (we used almost double size voxel comparing to the monkey fMRI study of Nishio et. al., 2012). It might also be that there are functional differences between human STS and monkey STS. It requires future studies to reveal this issue.

hMT+/V5. Our finding in *hMT+/V5* is similar to what we found in higher dorsal areas. We did not find any evidence about hMT+/V5 in processing monocular gloss information (and roughness information as well) through the three studies. It might take part in processing binocular gloss information, since in Study 3 it showed that in hMT+/V5 were distinguishable between glossy (i.e. mirror & anti-mirror) and matte (i.e. painted) condition defined by binocular gloss. In addition, hMT+/V5 was also responsible for processing information about objects' 3D structure. It is consistent with previous finding that hMT+/V5 responded to stereo-defined 3D shape (Chandrasekaran, Canon, Dahmen, Kourtzi, & Welchman, 2007; Orban, 2011; Peuskens et al., 2004; Welchman, Deubelius, Conrad, Bulthoff, & Kourtzi, 2005). It is still not clear whether hMT+/V5 responds to the disparity difference between glossy (mirror & anti-mirror) and matte (painted) condition instead of binocular gloss information *per se.* It is an open issue to examine the role of hMT+/V5 in processing glossiness information.

In summary, we confirmed the finding of a recent (which is also the first) human fMRI study in gloss perception (Wada et al., 2014) with different manipulation and analysis. Specifically, we separated V3A/B with a special localizer and found that V3B/KO is more selective to glossiness than V3A. We sub-divided ventral areas LO and pFs and found that pFs is more selective than LO in gloss processing. Earlier areas were also found to be less important since they primarily process for low-level image features (i.e. prefer scrambled images). We are also the first study which found that visually-presented information about material properties (glossy and rough objects) is decoded in secondary somatosensory area (S2). This provided neural evidence about the expected sense of touch from visual appearance of objects which might be crucial for further interaction such as planning finger placement and grasping force. We also performed the first human fMRI study to investigate how the brain processed binocular gloss information and compared it with monocular gloss

processing. We revealed that the processing of binocular gloss information involves in a more complex mechanism than the processing of monocular gloss information that needs cooperation of more areas (i.e. IPS) and additional activation patterns for extracting gloss information from binocular cue. Overall, the three studies boost our understanding about the neural basis of surface gloss and material processing in the human brain.

Research limitations

Since fMRI measurement in the three studies had limited time resolution, transient but important activations about gloss processing is not easy to be detected. Relevant EEG or MEG studies with high resolution measurements would be helpful in reveal this issue. Furthermore, high-resolution fMRI measurement would also be useful in finding small sub-regions that involve gloss processing (e.g. STS).

Acknowledgements

I would like to thank my supervisors Dr. Max Di Luca and Dr. Andrew Welchman for their guidance. They have extraordinary scientific insights and I learned a lot from them. I also appreciate all the other co-authors of the three papers for helping in stimuli generation, data analysis to manuscripts revision: Dr. Hiroshi Ban, Dr. Alexander Muryy, Dr. Dorita Chang and Dr. Roland Fleming. I thank Dr. Alan Meeson for providing the script of PSC and MVPA analysis and Dr. Yang Zhang for providing the script of global signal variance analysis. I would like to thank the examiners, Dr. Pia Rotshtein and Dr. Robert Kentridge, and anonymous reviewers for providing critical comments which strengthened the dissertation. Thanks to Nina Salman, Johnny King Lau, Roya Soufi and all the ex lab members who helped in my scans. I also would like to thank my family for their support even they are far away in Taiwan. Finally, I would like thank all my participants. They did good job and without them I would not be able to complete these studies.

Appendix 1

ROI definition for retinotopic visual cortex, V3B/KO, LOC, hMT+/V5, parietal cortex and somatosensory cortex:

Retinotopic visual cortex (areas V1, V2, V3, V3A, V4).

We used standard retinotopic mapping based on rotating wedge stimuli (DeYoe, et al., 1996; Sereno, et al., 1995). This localizer run lasts for 528 s with a rotating wedge (which contains moving patterns in it) moved clockwise in 8 positions for 8 s in each position and repeated for 8 cycles. Two 8 s fixation blocks were interposed at the beginning and at the end of the scan to measure fMRI signal baseline. Participants were instructed to maintain fixation and perform a fixation task, whereby they pressed a button if the fixation changes its orientation (from '+' to '×'). The wedge stimuli evoked travelling waves of activity in retinotopic organized visual areas V1-V4 and the boundaries between these areas were visualized and draw with different temporal phases of activation which correspond to different location in visual field (Larsson & Heeger, 2006).

Lateral Occipital Complex (LOC, includes LO and pFs)

LOC localizer run consisted of 21 blocks, 8 intact-object blocks, 8 scrambled-object blocks and 5 fixation blocks, each lasting 16s (336 seconds in length of a run). Intact- and scrambled-object blocks were showed in random order. In the intact-object blocks, 16 objects were presented in a presudo-random order with a 20 × 20 grid superimposed over the images.

In the scrambled-object blocks, the same images of the 16 objects were randomly relocated within the grid. Stimuli were presented for 500 ms with 500 ms interstimulus interval (ISI). Participants were instructed to maintain fixation and perform a 1-back matching task, whereby they pressed a button if the same image was presented twice in a row. Five 16 s fixation blocks were interposed at the beginning and at the end of the scan, as well as after the fifth, tenth, and fifteenth stimulus blocks to measure fMRI signal baseline. LOC was defined as the set of voxels in lateral occipito-temporal cortex that responded significantly ($p < 10^{-4}$) more strongly to intact than scrambled images of objects (Kourtzi & Kanwisher, 2000). LOC subregions (LO, extending into the posterior inferotemporal sulcus; posterior fusiform (pFs), posterior to mid-fusiform gyrus) were defined based on the overlap of functional activations and anatomical structures, consistent with previous studies (Grill-Spector et al., 2000).

V3B/KO

V3B/KO localizer run consisted of 21 blocks, 8 kinetic-boundaries blocks, 8 transparent-motion blocks and 5 fixation blocks, each lasting 16s (336 seconds in length of a run). Kinetic-boundaries blocks and transparent-motion blocks were showed in random order. In the kinetic-boundaries blocks, black and white dots in a field move in the same direction and the dots in the adjacent field moved in the opposite direction, generating kinetic boundaries (structure from motion). The direction of moving dots changes every 500ms. In the transparent-motion blocks half of the dots in the whole field move in the same direction and the other half move in the opposite direction, generating two transparent surfaces moving and overlapping with each other without any kinetic-boundary. The direction of the two groups of moving dots also changes every 500ms. Participants were instructed to maintain

fixation and perform a fixation task, whereby they pressed a button if the fixation point changes its color. Five 16 s fixation blocks were interposed at the beginning and at the end of the scan, as well as after the fifth, tenth, and fifteenth stimulus blocks to measure fMRI signal baseline. V3B/KO was defined as the set of contiguous voxels that responded significantly more ($p < 10^{-4}$) to kinetic boundaries than transparent motion of a field of black and white dots (Dupont, et al., 1997).

hMT+/V5

hMT+/V5 localizer run consisted of 22 blocks, 10 moving-dots blocks, 10 static-dots blocks (each lasting 16s) and 2 fixation blocks (each lasting 8s), making it 336 seconds in length of a run. Moving-dots blocks and static-dots blocks were showed in random order. In the Moving-dots blocks, dots coherently move toward the centre (contracting) for 1s and then move outward (expanding) for 1s and the cycles repeat until the end of the block. In the static-dots blocks, all the dots are static without any motion. Participants were instructed to maintain fixation and perform a fixation task, whereby they pressed a button if the fixation point changes its colour. Two 8 s fixation blocks were interposed at the beginning and at the end of the scan to measure fMRI signal baseline. hMT+/V5 was defined as the set of contiguous voxels that responded significantly more ($p < 10^{-4}$) to coherent moving dots than static dots (Zeki, et al., 1991).

Higher dorsal areas (V7-DIPSA)

This localizer run consisted of 21 blocks, 8 3D-SfM (3D structure from motion) blocks, 8 moving-dots blocks (each lasting 18s), 3 middle fixation blocks (each lasting 12s),

and 1 initial and 1 final fixation blocks (each lasting 20s), making it 364 seconds in length of a run. 3D-SfM blocks and moving-dots blocks were showed in random order. In the 3D-SfM blocks, black and white dots move in a way which generates the perception of and 3D shape moves laterally on screen. In the moving-dots blocks, dots move in a way that no 3D shape is shown. Participants were instructed to maintain fixation and perform a fixation task, whereby they pressed a button if the fixation point changes its colour. Five fixation blocks were interposed at the beginning and at the end of the scan, as well as after the fifth, tenth, and fifteenth stimulus blocks to measure fMRI signal baseline. V7, VIPS, POIPS, DIPSM and DIPSA were defined as the set of contiguous voxels that responded significantly more ($p < 10^{-4}$) to 3D-SfM than no 3D structure moving dots (Orban, et al., 2006; Orban, Sunaert, Todd, Van Hecke, & Marchal, 1999).

Somatosensory cortex (primarily and secondary somatosensory area)

Somatosensory areas were defined by a somatosensory localizer adapted from a previous study (Huang & Sereno, 2007). This separate localizer session consisted of 20 blocks, 10 air-on blocks and 10 air-off blocks showed alternately, each lasting 16s. In the air-on blocks, air puffs were delivered at the participants' ten fingertips through plastic tubes (6 mm inner diameter) from below a board in cycles (1 s on 1 s off). No air was delivered in air-off blocks. Somatosensory areas were defined by contrasting activations in air-on blocks with air-off blocks. Primary somatosensory area (S1) was defined as the more dorsal portion of the activations around Brodmann area 1-3b and the secondary somatosensory area (S2) was defined with as the more ventral portion of the activations around the parietal opercular areas OP1-OP4.

The boundary of each ROI was draw along the activations of a specific functional localizer as described above with strict p value ($p < 10^{-4}$) and cluster-size thresholding (25 mm^2) for controlling type I error. When the boundaries were not clear (e.g. activation overlapped with each other), Talarach coordinates of the ROIs from previously were used as references and the boundary of two adjacent ROIs was draw in the middle of the two ROI centres. The reference coordinates of each ROI are presented in the table below:

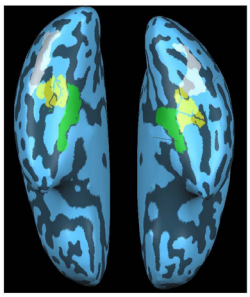
		x	y	z
Left hemisphere	LO	-44	-67	-3
	pFs	-41	-46	-20
	hMT+/V5	-42	-66	2
	V3B/KO	-42	-81	6
	VIPS	-27	-72	30
	POIPS	-18	-72	54
	DIPSM	-15	-63	60
	DIPSA	-36	-48	60
	BA1(for S1)	-44	28	52
	BA 3b (for S1)	-43	23	42
	OP1 (for S2)	-50	-25	18
Right hemisphere	LO	42	-67	-4
	pFs	46	-48	-22
	hMT+/V5	42	-62	6
	V3B/KO	42	-81	6
	VIPS	30	-78	27
	POIPS	24	-75	45
	DIPSM	18	-60	63
	DIPSA	39	-36	54
	BA1(for S1)	50	-27	50
	BA 3b (for S1)	42	-25	43
	OP1 (for S2)	55	-24	18

(Eickhoff, Grefkes, Zilles, & Fink, 2007; Eickhoff, Heim, Zilles, & Amunts, 2006; Eickhoff, et al., 2005; Geyer, Schleicher, & Zilles, 1999; Geyer, Schormann, Mohlberg, & Zilles, 2000; Grill-Spector et al., 2000; Orban et al., 2003; Sunaert et al., 1999)

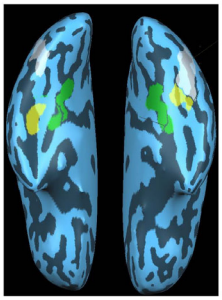
Appendix 2

The locations of LO (white region), pFs (yellow region) and CoS (green region) for 15 participants in Study 1. LO and pFs is defined by a functional localizer (LOC localizer) while CoS is defined by anatomical structure (Weiner, et al., 2014).

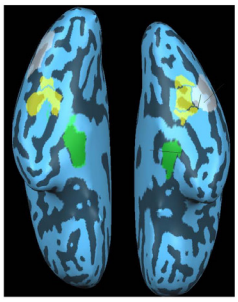
CL



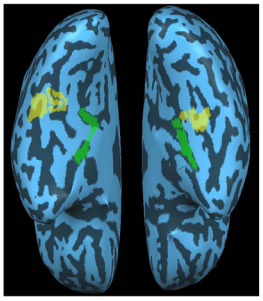
AG



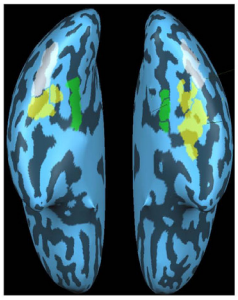
CY



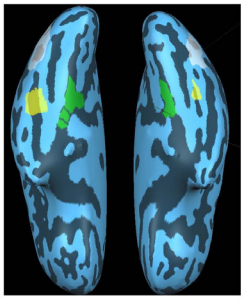
HB



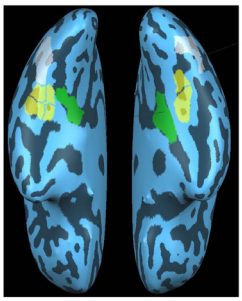
HS



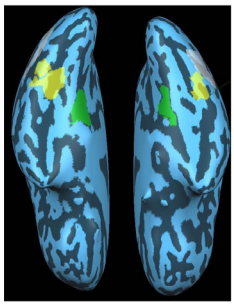
GB



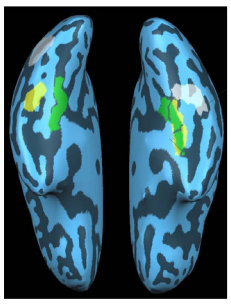
DP



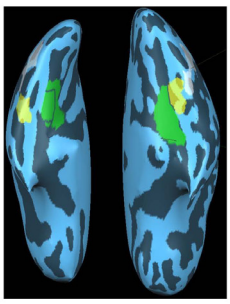
JC



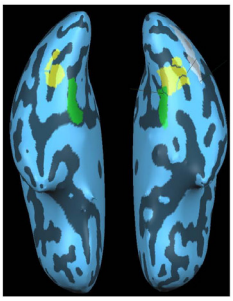
MG



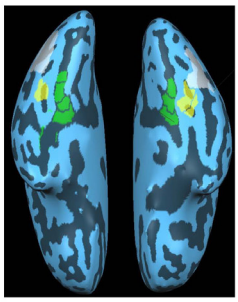
MS



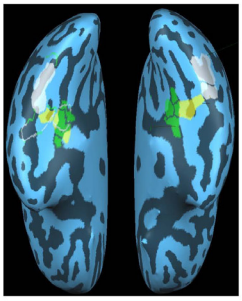
NG



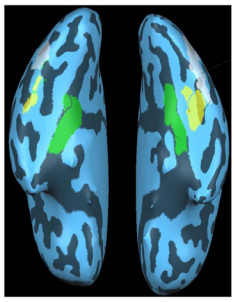
PS



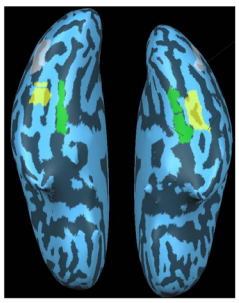
QC



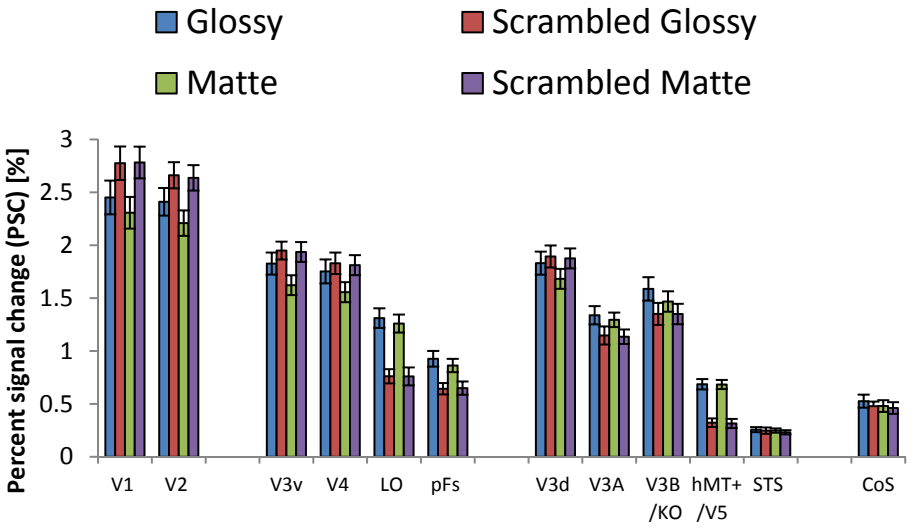
RG



ZG



Appendix 3



The PSC result of Study 1 with anatomically defined CoS under the four conditions: Glossy condition (blue bar), Scrambled Glossy condition (red bar), Matte condition (green bar) and Scrambled Matte condition (purple bar). No significant selectivity was found for glossy objects in CoS.

Reference

Amedi, A., Malach, R., Hendler, T., Peled, S., & Zohary, E. (2001). Visuo-haptic object-related activation in the ventral visual pathway. *Nature Neuroscience, 4*(3), 324-330. doi: 10.1038/85201

Anderson, B. L. (2011). Visual perception of materials and surfaces. *Current Biology, 21*(24), R978-R983. doi: 10.1016/j.cub.2011.11.022

Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision, 9*(11). doi: 10.1167/9.11.10

Anderson, B. L., Kim, J., & Marlow, P. (2011). At what level of representation is surface gloss computed? *Journal of Vision, 11*(11), 396. doi: 10.1167/11.11.396

Anderson, B. L., Marlow, P. J., & Kim, J. (2012). Disentangling 3D Shape and Perceived Gloss. *Journal of Vision, 12*(9), 947. doi: 10.1167/12.9.947

Arcizet, F., Jouffrais, C., & Girard, P. (2008). Natural textures classification in area V4 of the macaque monkey. *Experimental Brain Research, 189*(1), 109-120. doi: 10.1007/s00221-008-1406-9

Ban, H., Preston, T. J., Meeson, A., & Welchman, A. E. (2012). The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature Neuroscience, 15*(4), 636-643. doi: 10.1038/nn.3046

Beck, J. (1972). *Surface color perception*. Ithaca, NY: Cornell University Press.

Bergmann Tiest, W., & Kappers, A. L. (2014). The Influence of Material Cues on Early Grasping Force. In M. Auvray & C. Duriez (Eds.), *Haptics: Neuroscience, Devices, Modeling, and Applications* (Vol. 8618, pp. 393-399): Springer Berlin Heidelberg.

Bergmann Tiest, W. M. (2010). Tactual perception of material properties. *Vision Research, 50*(24), 2775-2782. doi: 10.1016/j.visres.2010.10.005

Bergmann Tiest, W. M., & Kappers, A. M. L. (2007). Haptic and visual perception of roughness. *Acta Psychologica, 124*(2), 177-189. doi: 10.1016/j.actpsy.2006.03.002

Berzhanskaya, J., Swaminathan, G., Beck, J., & Mingolla, E. (2005). Remote effects of highlights on gloss perception. *Perception, 34*(5), 565-575. doi: 10.1068/p5401

Blake, A., & Bülthoff, H. (1990). Does the brain know the physics of specular reflection? *Nature, 343*(6254), 165-168. doi: 10.1038/343165a0

Blakemore, S.-J., Bristow, D., Bird, G., Frith, C., & Ward, J. (2005). Somatosensory activations during the observation of touch and a case of vision–touch synaesthesia. *Brain, 128*(7), 1571-1583. doi: 10.1093/brain/awh500

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10*(4), 433-436. doi: 10.1163/156856897X00357

Buckingham, G., Cant, J. S., & 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. *Journal of Neurophysiology, 102*(6), 3111-3118. doi: 10.1152/jn.00515.2009

Cant, J. S., Arnott, S. R., & Goodale, M. A. (2009). fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Experimental Brain Research, 192*(3), 391-405. doi: 10.1007/s00221-008-1573-8

Cant, J. S., & Goodale, M. A. (2007). Attention to Form or Surface Properties Modulates Different Regions of Human Occipitotemporal Cortex. *Cerebral Cortex, 17*(3), 713-731. doi: 10.1093/cercor/bhk022

Cant, J. S., & Goodale, M. A. (2011). Scratching Beneath the Surface: New Insights into the Functional Properties of the Lateral Occipital Area and Parahippocampal Place Area.

The Journal of Neuroscience, 31(22), 8248-8258. doi: 10.1523/jneurosci.6113-10.2011

Carlsson, K., Petrovic, P., Skare, S., Petersson, K. M., & Ingvar, M. (2000). Tickling Expectations: Neural Processing in Anticipation of a Sensory Stimulus. *Journal of Cognitive Neuroscience, 12*(4), 691-703. doi: 10.1162/089892900562318

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010a). Separate Channels for Processing Form, Texture, and Color: Evidence from fMRI Adaptation and Visual Object Agnosia. *Cerebral Cortex, 20*(10), 2319-2332. doi: 10.1093/cercor/bhp298

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010b). Separate Processing of Texture and Form in the Ventral Stream: Evidence from fMRI and Visual Agnosia. *Cerebral Cortex, 20*(2), 433-446. doi: 10.1093/cercor/bhp111

Chandrasekaran, C., Canon, V., Dahmen, J. C., Kourtzi, Z., & Welchman, A. E. (2007). Neural Correlates of Disparity-Defined Shape Discrimination in the Human Brain. *Journal of Neurophysiology, 97*(2), 1553-1565. doi: 10.1152/jn.01074.2006

Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology, 2*(3), 1-27. doi: 10.1145/1961189.1961199

Connor, C., & Johnson, K. (1992). Neural coding of tactile texture: comparison of spatial and temporal mechanisms for roughness perception. *The Journal of Neuroscience, 12*(9), 3414-3426.

DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., et al. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences, 93*(6), 2382-2386.

Di Luca, M., & Ernst, M. O. (2014). Computational aspects of softness perception. In M. Di Luca (Ed.), *Multisensory softness* (pp. 85-107): Springer.

Doerschner, K., Boyaci, H., & Maloney, L. T. (2010). Estimating the glossiness transfer function induced by illumination change and testing its transitivity. *Journal of Vision, 10*(4). doi: 10.1167/10.4.8

Doerschner, K., Fleming, Roland W., Yilmaz, O., Schrater, Paul R., Hartung, B., & Kersten, D. (2011). Visual Motion and the Perception of Surface Material. *Current Biology, 21*(23), 2010-2016. doi: 10.1016/j.cub.2011.10.036

Doerschner, K., Maloney, L. T., & Boyaci, H. (2010). Perceived glossiness in high dynamic range scenes. *Journal of Vision, 10*(9). doi: 10.1167/10.9.11

Dövencioğlu, D., Ban, H., Schofield, A. J., & Welchman, A. E. (2013). Perceptual integration for qualitatively different 3-d cues in the human brain. *Journal of Cognitive Neuroscience, 25*(9), 1527-1541. doi: 10.1162/jocn_a_00417

Dror, R. O., Willsky, A. S., & Adelson, E. H. (2004). Statistical characterization of real-world illumination. *Journal of Vision, 4*(9). doi: 10.1167/4.9.11

Dupont, P., De Bruyn, B., Vandenberghe, R., Rosier, A. M., Michiels, J., Marchal, G., et al. (1997). The kinetic occipital region in human visual cortex. *Cerebral Cortex, 7*(3), 283-292. doi: 10.1093/cercor/7.3.283

Eck, J., Kaas, A. L., & Goebel, R. (2013). Crossmodal interactions of haptic and visual texture information in early sensory cortex. *NeuroImage, 75*(0), 123-135. doi: 10.1016/j.neuroimage.2013.02.075

Eck, J., Kaas, A. L., Mulders, J. L. J., & Goebel, R. (2013). Roughness perception of unfamiliar dot pattern textures. *Acta Psychologica, 143*(1), 20-34. doi: 10.1016/j.actpsy.2013.02.002

Eickhoff, S. B., Grefkes, C., Zilles, K., & Fink, G. R. (2007). The Somatotopic Organization of Cytoarchitectonic Areas on the Human Parietal Operculum. *Cerebral Cortex, 17*(8), 1800-1811. doi: 10.1093/cercor/bhl090

Eickhoff, S. B., Heim, S., Zilles, K., & Amunts, K. (2006). Testing anatomically specified hypotheses in functional imaging using cytoarchitectonic maps. *NeuroImage, 32*(2), 570-582. doi: 10.1016/j.neuroimage.2006.04.204

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage, 25*(4), 1325-1335. doi: dx.doi.org/10.1016/j.neuroimage.2004.12.034

Ernst, M. O. (2006). A Bayesian view on multimodal cue integration. In G. Knoblich, I. M. Thornton, M. Grosjean & M. Shiffrar (Eds.), *Human Body Perception From the Inside Out* (pp. 105-131): Oxford University Press.

Felleman, D. J., & Van Essen, D. C. (1991). Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex, 1*(1), 1-47. doi: 10.1093/cercor/1.1.1

Fleming, R. W. (2012). Human Perception: Visual Heuristics in the Perception of Glossiness. *Current Biology, 22*(20), R865-R866. doi: 10.1016/j.cub.2012.08.030

Fleming, R. W. (2014). Visual perception of materials and their properties. *Vision Research, 94*(0), 62-75. doi: 10.1016/j.visres.2013.11.004

Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision, 3*(5). doi: 10.1167/3.5.3

Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision, 4*(9). doi: 10.1167/4.9.10

Fleming, R. W., Wiebel, C., & Gegenfurtner, K. (2013). Perceptual qualities and material classes. *Journal of Vision, 13*(8). doi: 10.1167/13.8.9

Gallivan, Jason P., Cant, Jonathan S., Goodale, Melvyn A., & Flanagan, J. R. (2014). Representation of Object Weight in Human Ventral Visual Cortex. *Current Biology, 24*(16), 1866-1873. doi: dx.doi.org/10.1016/j.cub.2014.06.046

Gegenfurtner, K., Baumgartner, E., & Wiebel, C. (2013). The perception of gloss in natural images. *Journal of Vision, 13*(9), 200. doi: 10.1167/13.9.200

Georgieva, S. S., Todd, J. T., Peeters, R., & Orban, G. A. (2008). The Extraction of 3D Shape from Texture and Shading in the Human Brain. *Cerebral Cortex, 18*(10), 2416-2438. doi: 10.1093/cercor/bhn002

Geyer, S., Schleicher, A., & Zilles, K. (1999). Areas 3a, 3b, and 1 of Human Primary Somatosensory Cortex: 1. Microstructural Organization and Interindividual Variability. *NeuroImage, 10*(1), 63-83. doi: 10.1006/nimg.1999.0440

Geyer, S., Schormann, T., Mohlberg, H., & Zilles, K. (2000). Areas 3a, 3b, and 1 of Human Primary Somatosensory Cortex: 2. Spatial Normalization to Standard Anatomical Space. *NeuroImage, 11*(6), 684-696. doi: 10.1006/nimg.2000.0548

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*(6), 278-285. doi: 10.1016/j.tics.2006.04.008

Giesel, M., & Zaidi, Q. (2013). Frequency-based heuristics for material perception. *Journal of Vision, 13*(14), 1-19. doi: 10.1167/13.14.7

Goda, N., Tachibana, A., Okazawa, G., & Komatsu, H. (2014). Representation of the Material Properties of Objects in the Visual Cortex of Nonhuman Primates. *The Journal of Neuroscience, 34*(7), 2660-2673. doi: 10.1523/jneurosci.2593-13.2014

Granzier, J. J. M., Vergne, R., & Gegenfurtner, K. R. (2014). The effects of surface gloss and roughness on color constancy for real 3-D objects. *Journal of Vision, 14*(2). doi: 10.1167/14.2.16

Grill-Spector, K., Kushnir, T., Hendler, T., & Malach, R. (2000). The dynamics of object-selective activation correlate with recognition performance in humans. *Nature Neuroscience, 3*(8), 837-843. doi: 10.1038/77754

Guest, S., Catmur, C., Lloyd, D., & Spence, C. (2002). Audiotactile interactions in roughness perception. *Experimental Brain Research, 146*(2), 161-171. doi: 10.1007/s00221-002-1164-z

Guest, S., & Spence, C. (2003). Tactile dominance in speeded discrimination of textures. *Experimental Brain Research, 150*(2), 201-207. doi: 10.1007/s00221-003-1404-x

Guzman-Martinez, E., Ortega, L., Grabowecky, M., Mossbridge, J., & Suzuki, S. (2012). Interactive Coding of Visual Spatial Frequency and Auditory Amplitude-Modulation Rate. *Current Biology, 22*(5), 383-388. doi: 10.1016/j.cub.2012.01.004

Hiramatsu, C., Goda, N., & Komatsu, H. (2011). Transformation from image-based to perceptual representation of materials along the human ventral visual pathway. *NeuroImage, 57*(2), 482-494. doi: 10.1016/j.neuroimage.2011.04.056

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2006). How direction of illumination affects visually perceived surface roughness. *Journal of Vision, 6*(5). doi: 10.1167/6.5.8

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science, 19*(2), 196-204. doi: 10.1111/j.1467-9280.2008.02067.x

Ho, Y.-X., Maloney, L. T., & Landy, M. S. (2007). The effect of viewpoint on perceived visual roughness. *Journal of Vision, 7*(1). doi: 10.1167/7.1.1

Huang, R.-S., & Sereno, M. I. (2007). Dodecapus: An MR-compatible system for somatosensory stimulation. *NeuroImage, 34*(3), 1060-1073. doi: 10.1016/j.neuroimage.2006.10.024

Järveläinen, J., Schürmann, M., & Hari, R. (2004). Activation of the human primary motor cortex during observation of tool use. *NeuroImage, 23*(1), 187-192. doi: 10.1016/j.neuroimage.2004.06.010

Jones, B., & O'Neil, S. (1985). Combining vision and touch in texture perception. *Perception & Psychophysics, 37*(1), 66-72. doi: 10.3758/bf03207140

Junghöfer, M., Schupp, H. T., Stark, R., & Vaitl, D. (2005). Neuroimaging of emotion: empirical effects of proportional global signal scaling in fMRI data analysis. *NeuroImage, 25*(2), 520-526. doi: 10.1016/j.neuroimage.2004.12.011

Kaas, A. L., van Mier, H., Visser, M., & Goebel, R. (2013). The neural substrate for working memory of tactile surface texture. *Human Brain Mapping, 34*(5), 1148-1162. doi: 10.1002/hbm.21500

Kahrimanovic, M., Bergmann Tiest, W., & Kappers, A. L. (2009). Context effects in haptic perception of roughness. *Experimental Brain Research, 194*(2), 287-297. doi: 10.1007/s00221-008-1697-x

Kentridge, R. W., Thomson, R., & Heywood, C. A. (2012). Glossiness perception can be mediated independently of cortical processing of colour or texture. *Cortex, 48*(9), 1244-1246. doi: 10.1016/j.cortex.2012.01.011

Kerrigan, I. S., & Adams, W. J. (2013). Highlights, disparity, and perceived gloss with convex and concave surfaces. *Journal of Vision, 13*(1). doi: 10.1167/13.1.9

Kerrigan, I. S., Adams, W. J., & Graf, E. W. (2010). Does it feel shiny? Haptic cues affect perceived gloss. *Journal of Vision, 10*(7), 868. doi: 10.1167/10.7.868

Keysers, C., Wicker, B., Gazzola, V., Anton, J.-L., Fogassi, L., & Gallese, V. (2004). A Touching Sight: SII/PV Activation during the Observation and Experience of Touch. *Neuron, 42*(2), 335-346. doi: 10.1016/S0896-6273(04)00156-4

Kim, J., & Anderson, B. L. (2010). Image statistics and the perception of surface gloss and lightness. *Journal of Vision, 10*(9). doi: 10.1167/10.9.3

Kim, J., Marlow, P., & Anderson, B. L. (2011). The perception of gloss depends on highlight congruence with surface shading. *Journal of Vision, 11*(9). doi: 10.1167/11.9.4

Kim, J., Marlow, P. J., & Anderson, B. L. (2012). The dark side of gloss. *Nature Neuroscience, 15*(11), 1590-1595. doi: 10.1038/nn.3221

Kitada, R., Hashimoto, T., Kochiyama, T., Kito, T., Okada, T., Matsumura, M., et al. (2005). Tactile estimation of the roughness of gratings yields a graded response in the human brain: an fMRI study. *NeuroImage, 25*(1), 90-100. doi: 10.1016/j.neuroimage.2004.11.026

Kitada, R., Sadato, N., & Lederman, S. J. (2012). Tactile perception of nonpainful unpleasantness in relation to perceived roughness: Effects of inter-element spacing and speed of relative motion of rigid 2-D raised-dot patterns at two body loci. *Perception, 41*(2), 204.

Klatzky, R., & Lederman, S. (1999). Tactile roughness perception with a rigid link interposed between skin and surface. *Perception & Psychophysics, 61*(4), 591-607. doi: 10.3758/bf03205532

Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience, 2*(9), 635-642. doi: 10.1038/35090055

Köteles, K., De Mazière, P. A., Van Hulle, M., Orban, G. A., & Vogels, R. (2008). Coding of images of materials by macaque inferior temporal cortical neurons. *European Journal of Neuroscience, 27*(2), 466-482. doi: 10.1111/j.1460-9568.2007.06008.x

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience, 20*(9), 3310-3318.

Kriegeskorte, N., Goebel, R., & Bandettini, P. (2006). Information-based functional brain mapping. *Proceedings of the National Academy of Sciences of the United States of America, 103*(10), 3863-3868. doi: 10.1073/pnas.0600244103

Kuehn, E., Trampel, R., Mueller, K., Turner, R., & Schütz-Bosbach, S. (2013). Judging roughness by sight—A 7-tesla fMRI study on responsivity of the primary somatosensory cortex during observed touch of self and others. *Human Brain Mapping, 34*(8), 1882-1895. doi: 10.1002/hbm.22031

Landy, M. S. (2007). Visual perception: A gloss on surface properties. *Nature, 447*(7141), 158-159. doi: 10.1038/nature05714

Larsson, J., & Heeger, D. J. (2006). Two Retinotopic Visual Areas in Human Lateral Occipital Cortex. *The Journal of Neuroscience, 26*(51), 13128-13142. doi: 10.1523/jneurosci.1657-06.2006

Lawrence, M. A., Kitada, R., Klatzky, R. L., & Lederman, S. J. (2007). Haptic roughness perception of linear gratings via bare finger or rigid probe. *Perception, 36*(4), 547-557.

Lederman, S. J., & Abbott, S. G. (1981). Texture perception: studies of intersensory organization using a discrepancy paradigm, and visual versus tactual psychophysics. *Journal of Experimental Psychology: Human perception and performance, 7*(4), 902-915.

Lederman, S. J., & Klatzky, R. L. (2009). Haptic perception: A tutorial. *Attention, Perception, & Psychophysics, 71*(7), 1439-1459. doi: 10.3758/app.71.7.1439

Libouton, X., Barbier, O., Berger, Y., Plaghki, L., & Thonnard, J.-L. (2012). Tactile roughness discrimination of the finger pad relies primarily on vibration sensitive

afferents not necessarily located in the hand. *Behavioural Brain Research, 229*(1), 273-279. doi: 10.1016/j.bbr.2012.01.018

Libouton, X., Barbier, O., Plaghki, L., & Thonnard, J.-L. (2010). Tactile roughness discrimination threshold is unrelated to tactile spatial acuity. *Behavioural Brain Research, 208*(2), 473-478. doi: 10.1016/j.bbr.2009.12.017

Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., et al. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences, 92*(18), 8135-8139.

Marlow, P., & Anderson, B. L. (2013). Generative constraints on image cues for perceived gloss. *Journal of Vision, 13*(14). doi: 10.1167/13.14.2

Marlow, P., Kim, J., & Anderson, B. L. (2011). The role of brightness and orientation congruence in the perception of surface gloss. *Journal of Vision, 11*(9). doi: 10.1167/11.9.16

Marlow, P., Kim, J., & Anderson, Barton L. (2012). The Perception and Misperception of Specular Surface Reflectance. *Current Biology, 22*(20), 1909-1913. doi: 10.1016/j.cub.2012.08.009

Merabet, L., Thut, G., Murray, B., Andrews, J., Hsiao, S., & Pascual-Leone, A. (2004). Feeling by Sight or Seeing by Touch? *Neuron, 42*(1), 173-179. doi: 10.1016/S0896-6273(04)00147-3

Merabet, L. B., Swisher, J. D., McMains, S. A., Halko, M. A., Amedi, A., Pascual-Leone, A., et al. (2007). Combined Activation and Deactivation of Visual Cortex During Tactile Sensory Processing. *Journal of Neurophysiology, 97*(2), 1633-1641. doi: 10.1152/jn.00806.2006

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447*(7141), 206-209. doi: 10.1038/nature05724

Murphy, A. P., Ban, H., & Welchman, A. E. (2013). Integration of texture and disparity cues to surface slant in dorsal visual cortex. *Journal of Neurophysiology, 110*(1), 190-203. doi: 10.1152/jn.01055.2012

Muryy, A. A., Fleming, R. W., & Welchman, A. E. (2012). Binocular cues for glossiness. *Journal of Vision, 12*(9), 869. doi: 10.1167/12.9.869

Muryy, A. A., Fleming, R. W., & Welchman, A. E. (2014). Key characteristics of specular stereo. *Journal of Vision, 14*(14), 1-26. doi: 10.1167/14.14.14

Muryy, A. A., Welchman, A. E., Blake, A., & Fleming, R. W. (2013). Specular reflections and the estimation of shape from binocular disparity. *Proceedings of the National Academy of Sciences, 110*(6), 2413-2418. doi: 10.1073/pnas.1212417110

Nakano, H., Ueta, K., Osumi, M., & Morioka, S. (2012). Brain Activity during the Observation, Imagery, and Execution of Tool Use: An fNIRS/EEG Study. *Journal of Novel Physiotherapies, S1:009*. doi: 10.4172/2165-7025.S1-009

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: a critical link between lower-level and higher level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *An Invitation to Cognitive Science: Visual cognition* (pp. 1-70): M.I.T. Press.

Neri, P., Bridge, H., & Heeger, D. J. (2004). Stereoscopic Processing of Absolute and Relative Disparity in Human Visual Cortex. *Journal of Neurophysiology, 92*(3), 1880-1891. doi: 10.1152/jn.01042.2003

Newman, S. D., Klatzky, R. L., Lederman, S. J., & Just, M. A. (2005). Imagining material versus geometric properties of objects: an fMRI study. *Cognitive Brain Research, 23*(2–3), 235-246. doi: 10.1016/j.cogbrainres.2004.10.020

Nishida, S., Motoyoshi, I., & Maruya, K. (2011). Luminance-color interactions in surface gloss perception. *Journal of Vision, 11*(11), 397. doi: 10.1167/11.11.397

Nishida, S., Motoyoshi, I., Nakano, L., Li, Y., Sharan, L., & Adelson, E. (2008). Do colored highlights look like highlights? *Journal of Vision, 8*(6), 339. doi: 10.1167/8.6.339

Nishio, A., Goda, N., & Komatsu, H. (2012). Neural Selectivity and Representation of Gloss in the Monkey Inferior Temporal Cortex. *The Journal of Neuroscience, 32*(31), 10780-10793. doi: 10.1523/jneurosci.1095-12.2012

Nishio, A., Shimokawa, T., Goda, N., & Komatsu, H. (2014). Perceptual Gloss Parameters Are Encoded by Population Responses in the Monkey Inferior Temporal Cortex. *The Journal of Neuroscience, 34*(33), 11143-11151. doi: 10.1523/jneurosci.1451-14.2014

Norman, J. F., Todd, J. T., & Orban, G. A. (2004). Perception of three-dimensional shape from specular highlights, deformations of shading, and other types of visual information. *Psychological Science, 15*(8), 565–570. doi: 10.1111/j.0956-7976.2004.00720.x

Obein, G., Knoblauch, K., & Viéot, F. (2004). Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision, 4*(9). doi: 10.1167/4.9.4

Okazawa, G., Goda, N., & Komatsu, H. (2012). Selective responses to specular surfaces in the macaque visual cortex revealed by fMRI. *NeuroImage, 63*(3), 1321-1333. doi: 10.1016/j.neuroimage.2012.07.052

Okazawa, G., Koida, K., & Komatsu, H. (2011). Categorical properties of the color term "GOLD". *Journal of Vision, 11*(8). doi: 10.1167/11.8.4

Okazawa, G., Tajima, S., & Komatsu, H. (2014). Image statistics underlying natural texture selectivity of neurons in macaque V4. *Proceedings of the National Academy of Sciences*. doi: 10.1073/pnas.1415146112

Olkkonen, M., & Brainard, D. H. (2010). Perceived glossiness and lightness under real-world illumination. *Journal of Vision, 10*(9). doi: 10.1167/10.9.5

Orban, G. A. (2011). The Extraction of 3D Shape in the Visual System of Human and Nonhuman Primates. *Annual Review of Neuroscience, 34*(1), 361-388. doi: 10.1146/annurev-neuro-061010-113819

Orban, G. A., Claeys, K., Nelissen, K., Smans, R., Sunaert, S., Todd, J. T., et al. (2006). Mapping the parietal cortex of human and non-human primates. *Neuropsychologia, 44*(13), 2647-2667. doi: 10.1016/j.neuropsychologia.2005.11.001

Orban, G. A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., et al. (2003). Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. *Neuropsychologia, 41*(13), 1757-1768. doi: dx.doi.org/10.1016/S0028-3932(03)00177-5

Orban, G. A., Sunaert, S., Todd, J. T., Van Hecke, P., & Marchal, G. (1999). Human Cortical Regions Involved in Extracting Depth from Motion. *Neuron, 24*(4), 929-940. doi: 10.1016/S0896-6273(00)81040-5

Parker, A. J. (2007). Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience, 8*(5), 379-391. doi: 10.1038/nrn2131

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision, 10*(4), 437-442. doi: 10.1163/156856897X00366

Peuskens, H., Claeys, K. G., Todd, J. T., Norman, J. F., Hecke, P. V., & Orban, G. A. (2004). Attention to 3-D Shape, 3-D Motion, and Texture in 3-D Structure from Motion Displays. *Journal of Cognitive Neuroscience, 16*(4), 665-682. doi: 10.1162/089892904323057371

Picard, D. (2006). Partial perceptual equivalence between vision and touch for texture information. *Acta Psychologica, 121*(3), 227-248. doi: 10.1016/j.actpsy.2005.06.001

Podrebarac, S. K., Goodale, M. A., & Snow, J. C. (2014). Are visual texture-selective areas recruited during haptic texture discrimination? *NeuroImage, 94*, 129-137. doi: 10.1016/j.neuroimage.2014.03.013

Porro, C. A., Baraldi, P., Pagnoni, G., Serafini, M., Facchin, P., Maieron, M., et al. (2002). Does Anticipation of Pain Affect Cortical Nociceptive Systems? *The Journal of Neuroscience, 22*(8), 3206-3214.

Porro, C. A., Cettolo, V., Francescato, M. P., & Baraldi, P. (2003). Functional activity mapping of the mesial hemispheric wall during anticipation of pain. *NeuroImage, 19*(4), 1738-1747. doi: 10.1016/S1053-8119(03)00184-8

Porro, C. A., Lui, F., Facchin, P., Maieron, M., & Baraldi, P. (2004). Percept-related activity in the human somatosensory system: functional magnetic resonance imaging studies. *Magnetic Resonance Imaging, 22*(10), 1539-1548. doi: 10.1016/j.mri.2004.10.003

Preston, T. J., Li, S., Kourtzi, Z., & Welchman, A. E. (2008). Multivoxel Pattern Selectivity for Perceptually Relevant Binocular Disparities in the Human Brain. *The Journal of Neuroscience, 28*(44), 11315-11327. doi: 10.1523/jneurosci.2728-08.2008

Pruett, J. R., Sinclair, R. J., & Burton, H. (2000). *Response Patterns in Second Somatosensory Cortex (SII) of Awake Monkeys to Passively Applied Tactile Gratings* (Vol. 84).

Roland, P. E., O'Sullivan, B., & Kawashima, R. (1998). Shape and roughness activate different somatosensory areas in the human brain. *Proceedings of the National Academy of Sciences, 95*(6), 3295-3300.

Sakano, Y., & Ando, H. (2010). Effects of head motion and stereo viewing on perceived glossiness. *Journal of Vision, 10*(9). doi: 10.1167/10.9.15

Sathian, K., Lacey, S., Stilla, R., Gibson, G. O., Deshpande, G., Hu, X., et al. (2011). Dual pathways for haptic and visual perception of spatial and texture information. *NeuroImage, 57*(2), 462-475. doi: 10.1016/j.neuroimage.2011.05.001

Sawabe, M., Yamamoto, S., Nakaguchi, T., Yamauchi, Y., & Tsumura, N. (2010). A study for gloss perception on stereo display using magnitude estimation method. *Journal of Vision, 10*(15), 60. doi: 10.1167/10.15.60

Scheller Lichtenauer, M., Schuetz, P., & Zolliker, P. (2013). Interaction improves perception of gloss. *Journal of Vision, 13*(14). doi: 10.1167/13.14.14

Schroeder, C. E., & Foxe, J. (2005). Multisensory contributions to low-level, 'unisensory' processing. *Current Opinion in Neurobiology, 15*(4), 454-458. doi: 10.1016/j.conb.2005.06.008

Schütz-Bosbach, S., Tausche, P., & Weiss, C. (2009). Roughness perception during the rubber hand illusion. *Brain and Cognition, 70*(1), 136-144. doi: 10.1016/j.bandc.2009.01.006

Sereno, M., Dale, A., Reppas, J., Kwong, K., Belliveau, J., Brady, T., et al. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*(5212), 889-893. doi: 10.1126/science.7754376

Servos, P., Lederman, S., Wilson, D., & Gati, J. (2001). fMRI-derived cortical maps for haptic shape, texture, and hardness. *Cognitive Brain Research, 12*(2), 307-313. doi: 10.1016/S0926-6410(01)00041-6

Sharan, L., Li, Y., Motoyoshi, I., Nishida, S., & Adelson, E. H. (2008). Image statistics for surface reflectance perception. *Journal of the Optical Society of America, 25*(4), 846-865.

Simões-Franklin, C., Whitaker, T. A., & Newell, F. N. (2011). Active and passive touch differentially activate somatosensory cortex in texture perception. *Human Brain Mapping, 32*(7), 1067-1080. doi: 10.1002/hbm.21091

Smith, F. W., & Goodale, M. A. (2015). Decoding Visual Object Categories in Early Somatosensory Cortex. *Cerebral Cortex, 25*(4), 1020-1031. doi: 10.1093/cercor/bht292

Stilla, R., & Sathian, K. (2008). Selective visuo-haptic processing of shape and texture. *Human Brain Mapping, 29*(10), 1123-1138. doi: 10.1002/hbm.20456

Sun, H.-C., Ban, H., Di Luca, M., & Welchman, A. E. (2015). fMRI evidence for areas that process surface gloss in the human visual cortex. *Vision Research, 109*, 149-157. doi: 10.1016/j.visres.2014.11.012

Sunaert, S., van Hecke, P., Marchal, G., & Orban, G. A. (1999). Motion-responsive regions of the human brain. *Experimental Brain Research, 127*, 355-370.

Suzuki, Y., Gyoba, J., & Sakamoto, S. (2008). Selective effects of auditory stimuli on tactile roughness perception. *Brain Research, 1242*(0), 87-94. doi: 10.1016/j.brainres.2008.06.104

Te Pas, S. F., Pont, S. C., & van der Kooij, K. (2010). Both the complexity of illumination and the presence of surrounding objects influence the perception of gloss. *Journal of Vision, 10*(7), 450. doi: 10.1167/10.7.450

Tyler, C. W., Likova, L. T., Kontsevich, L. L., & Wade, A. R. (2006). The specificity of cortical region KO to depth structure. *NeuroImage, 30*(1), 228-238. doi: dx.doi.org/10.1016/j.neuroimage.2005.09.067

Uehara, T., Tani, Y., Nagai, T., Koida, K., Nakauchi, S., & Kitazaki, M. (2013). Effects of retinal-image motion of specular highlights induced by object motion and manual control on glossiness perception. *Journal of Vision, 13*(9), 204. doi: 10.1167/13.9.204

Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1997). The kinetic occipital (KO) region in man: an fMRI study. *Cerebral Cortex, 7*(7), 690-701. doi: 10.1093/cercor/7.7.690

Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J. T., et al. (2002). Extracting 3D from Motion: Differences in Human and Monkey Intraparietal Cortex. *Science, 298*(5592), 413-415. doi: 10.1126/science.1073574

Wada, A., Sakano, Y., & Ando, H. (2014). Human cortical areas involved in perception of surface glossiness. *NeuroImage, 98*(0), 243-257. doi: 10.1016/j.neuroimage.2014.05.001

Weiner, K. S., Golarai, G., Caspers, J., Chuapoco, M. R., Mohlberg, H., Zilles, K., et al. (2014). The mid-fusiform sulcus: A landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. *NeuroImage, 84*, 453-465. doi: 10.1016/j.neuroimage.2013.08.068

Welchman, A. E., Deubelius, A., Conrad, V., Bulthoff, H. H., & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience, 8*(6), 820-827. doi: 10.1038/nn1461

Wendt, G., Faul, F., Ekroll, V., & Mausfeld, R. (2010). Disparity, motion, and color information improve gloss constancy performance. *Journal of Vision, 10*(9). doi: 10.1167/10.9.7

Wendt, G., Faul, F., & Mausfeld, R. (2008). Highlight disparity contributes to the authenticity and strength of perceived glossiness. *Journal of Vision, 8*(1). doi: 10.1167/8.1.14

Whitaker, T. A., Simões-Franklin, C., & Newell, F. N. (2008). Vision and touch: Independent or integrated systems for the perception of texture? *Brain Research, 1242*(0), 59-72. doi: 10.1016/j.brainres.2008.05.037

Wijntjes, M. W. A., & Pont, S. C. (2010). Illusory gloss on Lambertian surfaces. *Journal of Vision, 10*(9). doi: 10.1167/10.9.13

Xiao, B., Jia, X., & Adelson, E. (2013). Can you see what you feel? Tactile and visual matching of material properties of fabrics. *Journal of Vision, 13*(9), 197. doi: 10.1167/13.9.197

Yoshioka, T., Craig, J. C., Beck, G. C., & Hsiao, S. S. (2011). Perceptual Constancy of Texture Roughness in the Tactile System. *The Journal of Neuroscience, 31*(48), 17603-17611. doi: 10.1523/jneurosci.3907-11.2011

Yoshioka, T., Gibb, B., Dorsch, A. K., Hsiao, S. S., & Johnson, K. O. (2001). Neural Coding Mechanisms Underlying Perceived Roughness of Finely Textured Surfaces. *The Journal of Neuroscience, 21*(17), 6905-6916.

Zeki, S., Watson, J., Lueck, C., Friston, K., Kennard, C., & Frackowiak, R. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience, 11*(3), 641-649.